

Université  
de Liège



Université de Liège

Faculté des Sciences

Département des Sciences et Gestion de l'Environnement

## Modélisation spatiale de la production fourragère en zone pastorale nigérienne

**Issa GARBA**

Thèse présentée en vue de l'obtention  
du grade de Docteur en Sciences

**30 Mars 2017**

### **Composition du jury :**

Président : Dr. Pierre. OZER (Université de Liège)

Promoteur : Pr. Bernard TYCHON (Université de Liège)

Co promoteur : Dr. Bakary DJABY (Université de Liège)

Rapporteurs : Pr. Jérôme BINDELLE (Université de Liège)

Dr. Ibra TOURE (CIRAD)

Dr. Pierre HIERNAUX (CNRS, Géoscience Environnement)

Dr. François KAYITAKIRE (JRC – Ispra)

**Année académique 2016-2017**



## Sommaire

SOMMAIRE .....	I
REMERCIEMENTS.....	V
RÉSUMÉ .....	VII
ABSTRACT .....	IX
LISTE DES TABLEAUX.....	XI
LISTE DES SIGLES ET ABRÉVIATIONS.....	XVI
I. INTRODUCTION GÉNÉRALE.....	1
Objectifs de la thèse.....	5
Structure de la thèse.....	6
1.1. Présentation de la zone d'étude.....	6
1.1.1. Situation géographique .....	6
1.1.2. Régime pluviométrique .....	7
1.1.3. Subdivision bioclimatique du Niger .....	8
1.1.4. Évolution de la pluviosité sahélienne de 1900 à 2010 .....	9
1.1.5. Productivité des pâturages.....	10
1.1.5.1. Productivité fourragère des zones bioclimatiques	10
1.1.5.2. Calcul de la production de la phytomasse aérienne basé sur le Dry Matter Productivity (DMP) .....	11
1.2. Évaluation de la phytomasse fourragère au sol (herbacée et ligneuse) .....	12
1.2.1. Généralités sur l'évaluation des pâturages au sol .....	12
1.2.2. Évaluation de la biomasse fourragère sur le terrain par le MEIA .....	13
1.2.3. Distribution spatiale des sites de mesures de phytomasse herbacée du MEIA .....	23
1.2.4 Distribution temporelle des mesures de phytomasse herbacée aérienne .....	28
.....	32
II. VALIDATION DU MODÈLE D'ÉVALUATION DE LA BIOMASSE HERBACÉE DES PARCOURS DU SAHEL (BIOMASAH) AU NIGER .....	33
2.1. Introduction .....	33

2.2. Matériel et méthodes .....	34
<b>2.2.1. Matériel .....</b>	<b>34</b>
<b>2.2.2. Description du Modèle BIOMASAH.....</b>	<b>34</b>
<b>2.2.3. Méthode de validation.....</b>	<b>37</b>
2.2.3.1. Données issues du modèle .....	37
2.2.3.2. Extraction des données simulées par utilisation de la couche vectorielle des sites du MEIA .....	38
<b>2.2.4. Analyse statistique.....</b>	<b>39</b>
2.3. Résultats et discussion.....	42
<b>2.3.1. Résultats .....</b>	<b>42</b>
<b>2.3.2. Discussion partielle.....</b>	<b>46</b>
2.4. Conclusion partielle .....	48
<b>III. ANALYSE DE LA PERFORMANCE DU MODÈLE D'ESTIMATION DE LA BIOMASSE ÉLABORÉ PAR LE MINISTÈRE DE L'ÉLEVAGE ET DES INDUSTRIES ANIMALES (MEIA) DU NIGER.....</b>	<b>49</b>
3.1. Introduction .....	49
3.2. Matériel et méthodes .....	52
<b>3.2.1. Les données de biomasse fourragère du MEIA .....</b>	<b>53</b>
<b>3.2.2. Les images satellitaires utilisées .....</b>	<b>54</b>
3.2.2.1. Images NDVI du satellite SPOT VEGETATION	54
3.2.2.2. Les images Expedited MODIS (eMODIS).....	54
<b>3.2.3. Extraction de l'intégrale et du maximum NDVI par site</b>	<b>55</b>
<b>3.2.4. Analyse statistique des données .....</b>	<b>56</b>
3.4. Résultats et discussion.....	57
<b>3.4.1. Résultats .....</b>	<b>57</b>
3.4.1.1 Caractérisation de la productivité .....	57
3.4.1.2. Analyse à l'échelle globale .....	58
3.4.1.3. Analyse par zone bioclimatique .....	59
3.4.1.4. Analyse suivant les années .....	62
<b>3.4.2. Discussion partielle.....</b>	<b>68</b>
3.5. Conclusion partielle .....	70
<b>IV. AMÉLIORATION DE LA MODÉLISATION DE LA BIOMASSE FOURRAGÈRE EN ZONE SAHÉLIENNE À</b>	

<b>TRAVERS LA MÉTHODE DE RÉGRESSION LINÉAIRE</b>	
<b>MULTIPLE .....</b>	71
4.1. Introduction .....	71
4.2. Matériel et Méthodes.....	72
<b>4.2.1. Matériel .....</b>	74
4.2.1.1. Données agrométéorologiques.....	74
4.2.2.2. Outils .....	75
<b>4.2.2. Méthode .....</b>	79
4.2.2.1. Présence minimum de végétation.....	79
4.2.2.2. Décades d'installation (planting dekads).....	80
4.2.2.3. Longueur de la végétation .....	80
4.2.2.4. Analyse statistique .....	80
4.3. Résultats et discussion.....	84
<b>4.3.1. Résultats .....</b>	84
4.3.1.1. Analyse exploratoire des mesures de masse herbacée de 2001 à 2012 .....	84
.....	86
4.3.1.2. Ajustement global de la productivité fourragère sur l'ensemble de la zone étudiée.....	87
4.3.1.3. Analyse suivant les écorégions .....	92
4.3.1.4. Analyse par Type de sol FAO.....	99
4.3.1.5. Analyse de la productivité fourragère herbacée par faciès .....	101
4.3.1.6. Modèle d'estimation suivant les années .....	103
4.4. Conclusion partielle .....	108
<b>V. PERFORMANCE DE L'ANALYSE DE SIMILARITÉ DANS</b>	
<b>L'ESTIMATION DES RENDEMENTS FOURRAGER AU</b>	
<b>NIGER .....</b>	109
5.1. Introduction .....	109
<b>5.1.1. Contexte.....</b>	109
<b>5.1.2. Revue de littérature sur la similarité .....</b>	110
5.2. Matériel et méthodes .....	113
<b>5.2.1. Généralités sur SPIRITS.....</b>	114
<b>5.2.2. Méthodes .....</b>	115
2.2.2.1. Principe de la méthode de similarité .....	115

5.2.2.2. Préparation des données .....	116
5.2.2.3. Traitement des données sur SPIRITS.....	116
5.2.2.4. Analyse statistique .....	117
5.3. Résultats et discussion.....	120
<b>5.3.1 Résultats .....</b>	<b>120</b>
5.3.1.1. Analyse globale.....	120
.....	127
5.3.1.2. Tests de t et Wilcoxon selon les années .....	128
5.3.1.3. Analyse par écorégion .....	130
5.3.1.4. Analyse par zone bioclimatique.....	132
.....	133
<b>5.3.2. Discussion partielle.....</b>	<b>134</b>
5.4. Conclusion partielle .....	137
<b>VI. COMPARAISON DES MODÈLES .....</b>	<b>138</b>
6.1. Méthode de la comparaison de modèles.....	138
6.2. Résultats de la comparaison des modèles.....	138
<b>VII. DISCUSSION GÉNÉRALE .....</b>	<b>140</b>
<b>VIII. CONCLUSION GÉNÉRALE ET PERSPECTIVES .....</b>	<b>146</b>
<b>IIX. RÉFÉRENCES BIBLIOGRAPHIQUES .....</b>	<b>150</b>
<b>ANNEXES .....</b>	<b>I</b>

## **Remerciements**

La thèse est certes une démarche personnelle, mais ne saurait se réaliser sans le soutien et les conseils avisés du comité et des autres personnes qui nous entourent.

Mes premiers mots de remerciements vont ainsi à mon promoteur de thèse, le Pr. Bernard TYCHON, pour l'opportunité offerte de réaliser ce projet au sein de son équipe et la confiance qu'il m'a accordée. Il m'a octroyé un appui scientifique indéfectible tout au long de ces années de travail. Je l'en remercie chaleureusement.

Mes premiers mots de remerciements vont aussi à mon co-promoteur de thèse, Dr. Bakary DJABY, pour avoir fortement contribué à me donner l'occasion de réaliser cette thèse au sein de l'équipe et la confiance qu'il m'a toujours renouvelée et surtout la patience de me supporter.

Je remercie également Dr. Ibra TOURE pour avoir accepté de faire partie de mon comité de thèse et de mon jury. Ses commentaires critiques, suggestions et conseils ont contribué substantiellement à l'amélioration de la qualité de ce manuscrit.

Merci également à Dr. Pierre OZER pour avoir accepté de présider le jury ; Dr. Pierre HIERNAUX ; Pr. Jérôme BINDELLE et Dr. François KAYITAKIRE pour avoir accepté de faire partie du jury et pour le temps consacré à la lecture du manuscrit.

J'adresse mes remerciements à Dr. Illa SALIFOU, au Pr. Maxime BANOIN et M. Amadou BOUREIMA pour leurs conseils et les multiples lectures du document à chaque étape de son avancement.

Une mention toute spéciale au Pr. BOUAFOU, Directeur Général du Centre Régional AGRHYMET, au Pr. Sanoussi ATTA, Chef de Département Formation Recherche (DFR) et à M. Issifou ALFERI, Cher de Département Information Recherche (DIR) pour leurs soutiens et encouragements très précieux, J'espère que cette collaboration se

poursuivra encore longtemps.

Toute ma reconnaissance à Dr Djime ADOUM, Secrétaire Exécutif (SE) du CILSS, M. Issa IDI, Secrétaire Exécutif Adjoint (SEA) du CILSS et Dr Mahalmoudou HAMADOUN responsable du PRA/SA-LEC-POP DEV.

Un grand merci à l'Unité Eau, Environnement et Développement du département Sciences et Gestion de l'Environnement d'Arlon Campus Environnement, et plus particulièrement à l'équipe BT pour toute l'ambiance chaleureuse et amicale que nous avons eue durant ces années.

Je tiens à remercier l'Université de Liège (ULG) pour son accueil, ses services et ses infrastructures.

J'exprime également ma reconnaissance au projet AGRICAB pour avoir financé en partie ces travaux de thèse, les projets ISACIP et MESA pour les soutiens financiers durant la finalisation de cette thèse.

Je ne saurai terminer sans adresser un merci tout particulier et exprimer ma reconnaissance aux membres de ma famille, mes parents, mes frères et sœurs. Merci pour votre soutien indéfectible et vos conseils très utiles durant ces années.

## Résumé

Ces travaux de thèse ont porté sur la zone pastorale du Niger. Ils ont pour objectif principal de contribuer à l'amélioration des méthodes d'estimation des rendements fourragers au Sahel en général et en particulier au Niger. Il s'agit plus spécifiquement de : valider le modèle BIOMASAHI (BIOMASse SAHélienne) utilisé par le Centre Régional AGRHYMET [AGRonomie Hydrologie et METéorologie] (CRA); tester le modèle du Ministère de L'Élevage et de Industries Animales (MEIA) ; proposer un Modèle par Régression linéaire Multiple (MRM) ; tester la méthode de similarité et enfin comparer ces méthodes d'estimation entre elles. Le travail a été réalisé d'une part avec les mesures de masse herbacée faites au sol de 2001 à 2012 par le MEIA, les observations pluviométriques des stations de la Direction Météorologique du Niger, les variables météorologiques issues du l'institution européenne appelée "European Centre for Medium-Range Weather Forecasts" (ECMWF) et d'autre part, avec les images satellitaires notamment le NDVI de SPOT VEGETATION et MODIS et les pluies estimées dénommées RFE provenant de l'institution américaine "Famine Early Warning Systems NETwork"(FEWSNET). La validation du modèle BIOMASAHI a été faite à l'aide des tests de t et de Wilcoxon pour comparer les moyennes mesurées in situ à celles obtenues du modèle et aussi des tests de corrélation de Pearson, Kendall et Spearman. Quant au modèle MEIA, les performances ont été testées en comparant les résultats inter et intra capteurs SPOT VEGETATION et MODIS en utilisant les  $R^2$  et le RMSE issus des calculs avec l'intégrale et le maximum NDVI comme variables explicatives du rendement fourrager.

Le modèle de référence (MR) a été réalisé par régression linéaire multiple avec la méthode pas-à-pas ascendante pour la sélection de variables basée sur le  $R^2$  ajusté et le RMSE. La validation croisée ‘leave one out’ (LOOCV) a été utilisée pour calculer les  $R^2$  de validation et un diagnostic systématique des résidus pour mieux caractériser le modèle.

La méthode de similarité des profils saisonniers d'indice de végétation a été réalisée en utilisant comme critères le R<sup>2</sup>, le MAD et le RMSE.

Le profil de la période de croissance de la végétation de chaque pixel a été tracé pour toutes les années. Ensuite, le profil de l'année cible a été comparé avec celui des autres années pour identifier l'année similaire. Les résultats de la similarité ont été confrontés aux données observées d'une part avec les tests de corrélations de Pearson, Spearman et Kendall et d'autre part à l'aide de tests de t et de Wilcoxon pour comparer les moyennes. La comparaison des quatre modèles a été faite sur la base des R<sup>2</sup>, des R<sup>2</sup> ajustés et les RMSE.

Le modèle BIOMASAH a donné des moyennes significativement différentes des moyennes observées ( $p < 0,001$ ). Les corrélations de Pearson, Kendall et Spearman sont faibles. En ce qui concerne le modèle MEIA, le meilleur R<sup>2</sup> à l'échelle globale est de 0,56. Il n'y a pas de différence significative à utiliser les NDVI de MODIS ou de SPOT VEGETATION. Le RMSE est de 367 kg.ha<sup>-1</sup>. Les R<sup>2</sup> et le RMSE varient fortement d'une année à l'autre. Le modèle de référence a donné un R<sup>2</sup> ajusté global de 0,69 et un RMSE de 282 kg.ha<sup>-1</sup>, la différence entre le RMSE calculé et celui de la validation est de 2,72 kg.ha<sup>-1</sup>. La comparaison des moyennes de la similarité à celles observées a montré qu'il n'y a pas de différence significative ( $p < 0,001$ ) pour le R<sup>2</sup>. Par contre les différences sont significatives au même seuil pour le MAD et le RMSE.

La comparaison des modèles montre que le Modèle par Régression linéaire Multiple (MRM) est le meilleur, mais il reste perfectible. Ainsi, se propose-t-on de continuer ces recherches avec d'autres indices tels que le LAI et le FAPAR et l'EVI. Aussi, il serait intéressant de prospecter les voies telles que : tenir compte du feuillage des ligneux, ajuster les métriques à la phénologie des herbacées, et à celles des ligneux. Ces travaux permettront d'améliorer la qualité des informations utilisées pour planifier les actions de développement en

faveur de la société nigérienne en vue de la protéger contre les crises pastorales.

Mots clés : NDVI, MEIA, BIOMASA, SPOT VEGETATION, eMODIS, MR, Similarité, Modèle.

## **Abstract**

This work was carried out on the pastoral zone of Niger, the main objective was to contribute to the improvement of the methods of forage yields predicting mainly in the Sahel and especially in Niger. This is specifically to validate the model BIOMASA of ARC; test the MEIA model; to establish a reference model by multiple linear regression; test the similarity method and finally compare the methods. The work was carried out on the one hand with the data measured on the ground by the MEIA from 2001 to 2012, reel rainfall of Niger observations network, meteorological parameters from ECMWF and also with satellite images as SPOT NDVI VEGETATION and MODIS, RFE2 of FEWS NET. Validation of BIOMASA model was made by t and Wilcoxon tests to compare reel biomass and outputs of the model. Pearson, Kendall and Spearman correlation testing was also made. The MEIA model performance was tested by confronting the results between and within SPOT VEGETATION and MODIS sensors, by comparing  $R^2$  and RMSE from the integral and maximum NDVI as a predictor of forage yield. Average comparisons by parametric and nonparametric tests were also made to compare the results. The reference model (RM) was produced by multiple linear regression with stepwise method. The selection of variables was based on adjusted  $R^2$  and RMSE and the LOOCV leave one out cross validation to calculate  $R^2$  for validation, we made also systematic diagnosis of residues for better characterization of the model. The similarity method was performed using the  $R^2$ , MAD and RMSE as a criterion, the profile of the vegetation growth period of each pixel was plotted for all years. Then we compare the profile of the target year with those of other years

to identify the similar year. One hand the results of similarity were compared with actual data with the Pearson correlation test, Spearman and Kendall and secondly using t and Wilcoxon tests to compare means. Comparison of models was made on the basis of  $R^2$ , Adjusted  $R^2$  and RMSE. Model BIOMASAH result on significant difference between average ( $p < 0.001$ ). Pearson correlations, Kendall and Spearman are low. Regarding the MEIA model, globally  $R^2$  (0.56) is best, there's no difference to use MODIS NDVI or SPOT vegetation, the RMSE is 367 kg.ha<sup>-1</sup>.  $R^2$  and RMSE vary greatly from one year to another. On a global scale the multiple linear model gave a good  $R^2$  adjusted (0.69) and RMSE (282 kg / ha) the difference between the calculated and the RMSE of validation is 2.72 kg. Comparing averages of the similarity to the real ones showed that there are no significant differences ( $p < 0.001$ ) for  $R^2$  with the differences are significant against the same threshold for the MAD and RMSE. The Comparison of the models shows that the multiple linear regression one (reference model) is the best. Research should continue with index like LAI, FARAR and EVI.

Key words: NDVI, MEIA, BIOMASAH, SPOT VEGETATION, eMODIS, RM, Similarity, Model

## Liste des Tableaux

Tableau 1.1 : synthèse des décades de début et de fin des mesures in situ de la biomasse herbacée épigée au Niger de 2000 à 2012.....	19
Tableau 1.2 : La répartition des sites de mesure de phytomasse herbacée aérienne selon les écorégions (rendement fourrager de 2001 à 2012) .....	25
Tableau 1.3 : La répartition des sites de mesures de phytomasse herbacée aérienne par type de sol .....	26
Tableau 1.4 : la répartition des sites de mesure de phytomasse herbacée aérienne par zone bioclimatique.....	27
Tableau 1.5 : Disponibilité des mesures de phytomasse herbacée aérienne par site suivant les années .....	29
Tableau 2.1 : Comparaison des moyennes globales suivant les tests de t, Wilcoxon et des signes.....	43
Tableau 2.2 : Comparaison par seuil pluviométrique ( $> 250$ mm) des moyennes globales suivant les 3 tests .....	43
Tableau 2.3 : Comparaison par seuil pluviométrique ( $< 250$ mm) des moyennes globales suivant les 3 tests .....	43
Tableau 2.4 : Comparaison des moyennes par an et suivant les quatre types de test .....	44
Tableau 2.5 : Corrélations paramétrique et non paramétrique sur l'ensemble des données.....	45
Tableau 3.1 : rendement fourrager moyen suivant les années .....	57
Tableau 3.2 : rendement fourrager moyen suivant les zones bioclimatiques .....	57
Tableau 3.3 : Comparaison des $R^2$ des résultats avec SPOT VEGETATION et eMODIS à l'échelle globale .....	59
Tableau 3.4 : Variation du coefficient de détermination ( $R^2$ ) suivant les zones bioclimatiques.....	61
Tableau 3.5 : Variation du RMSE suivant les zones bioclimatiques ..	61
Tableau 3.6 : Variation relative du RMSE suivant les zones bioclimatiques .....	61
Tableau 3.7 : Relation entre l'intégrale de NDVI de SPOT VEGETATION et la biomasse herbacée suivant les années.....	63
Tableau 3.8: relation entre le maximum de NDVI de SPOT VEGETATION et la biomasse herbacée suivant les années .....	63
Tableau3.9 : Relation entre l'intégrale de NDVI d'eMODIS et la biomasse herbacée suivant les années.....	64
Tableau 3.10: relation entre le maximum de NDVI de eMODIS et la biomasse herbacée suivant les années.....	65
Tableau 3.11 : Variation annuelle des $R^2$ issus de la régression entre la biomasse herbacée et l'intégrale ou le maximum NDVI de SPOT	

VEGETATION et eMODIS .....	66
Tableau 3.12 : Variation annuelle des RMSE issues de la régression entre la biomasse herbacée en l'intégrale et maximum NDVI de SPOT VEGETATION et eMODIS .....	66
Tableau 3.13 : Test de Wilcoxon sur les RMSE issues de la régression entre la biomasse herbacée en l'intégrale et maximum NDVI de SPOT VEGETATION et eMODIS (basé sur les rangs).....	67
Tableau 3.14 : Test des signes sur les RMSE issues de la régression entre la biomasse herbacée en l'intégrale et maximum NDVI de SPOT VEGETATION et eMODIS .....	67
Tableau 4.1 : Analyse exploratoire des mesures de masse herbacée de 2001 à 2012 par bootstrap (kg MS.ha-1).....	86
Tableau 4.2 : Quatre meilleurs modèles globaux en fonction du nombre de variables .....	88
Tableau 4.3 : Estimation des coefficients du modèle global à quatre variables .....	89
Tableau 4.4 : Analyse exploratoire bootstrap des résidus .....	91
Tableau 4.5 : Les quatre meilleurs modèles de l'Azaouak en fonction du nombre de variables .....	93
Tableau 4.6 : Estimation des coefficients du modèle de l'Azawak ....	94
Tableau 4.7 : Estimations des coefficients .....	95
Tableau 4.8 : Résumé de l'ajustement .....	96
Tableau 4.9 : Estimations des coefficients .....	96
Tableau 4.10 : Récapitulatif des meilleurs modèles réalisés avec les observations valides enregistrées de 2001 à 2012 par écorégion .....	98
Tableau 4.11 : Récapitulatif des meilleurs modèles réalisés avec les observations valides enregistrées de 2001 à 2012 par type de sol FAO .....	100
Tableau 4.12 : récapitulatif des modèles par faciès .....	102
Tableau 4.13 : Récapitulatif des paramètres des Modèles suivant les années.....	104
Tableau 5.1 : Statistique descriptive de la similarité par utilisation du R <sup>2</sup> .....	121
Tableau 5.2 : Statistiques descriptives de la similarité le RMSE .....	123
Tableau 5.3 : Statistique descriptive de la similarité avec le MAD ..	125
Tableau 5.4 : Corrélations paramétrique et non paramétriques .....	126
Tableau 5.5 : Comparaison des moyennes simulées à celles des mesures de masse herbacées à l'échelle globale suivant les tests de t et de Wilcoxon, selon les mesures de similarité.....	127
Tableau 5.6 : Tests paramétrique et non paramétrique suivant les années.....	129
Tableau 5.7 : Tests paramétrique et non paramétrique selon les	

écorégions .....	131
Tableau 5.8 : Tests paramétrique et non paramétrique par zone bioclimatique .....	133
Tableau 6.1 : Comparaison du R <sup>2</sup> et du RMSE des trois modèles à l'échelle globale.....	139
Tableau 6.2 : Comparaison du R <sup>2</sup> et du RMSE des trois modèles selon les années .....	140

## Liste des Figures

Figure 1.1 : Présentation de la zone d'étude.....	7
Figure 1.2 : Superposition de la moyenne 1971-2000 et celle 2001-2010 .....	8
Figure 1.3 : Zones bioclimatiques du Niger .....	9
Figure 1.4 : Évolution historique de la pluviosité au Sahel.....	9
Figure 1.5 : Estimation de la production moyenne de la phytomasse aérienne (1998-2012) dérivée du DMP.....	12
Figure 1.6 : Situation géographique des sites de relevé de végétation fourragère au Niger.....	14
Figure 1.7 : Schémas des Hazards 1 à 5 et la disposition des plateaux sur les points de mesures.....	17
Figure 1.8 : Exemple de l'année 2012 première étape du filtrage des données .....	21
Figure 1.9 : exemple de l'année 2006 sur la deuxième étape du filtrage des données .....	22
Figure 1.10 : Relation entre la phytomasse aérienne herbacée mesurée et le cumul des pluies estimées par satellite superposée à la relation théorique entre le cumul pluviométrique et la phytomasse .....	23
Figure 1.11 : Les écorégions du Niger .....	24
Figure 2.1 : Exemple de résultat issu du modèle BIOMASA.....	38
Figure 2.1 : Zoom sur les sites A64 et A 23 montrant les surfaces d'extraction des valeurs moyennes de phytomasse herbacée simulée par le modèle .....	39
Figure 2.2 : Variation annuelle des coefficients de corrélation suivant les différents tests .....	45
Figure 3.1 : Modèle schématique d'extraction des métriques du NDVI et du traitement statistique.....	53
Figure 3.2 : Exemple du site A23 matérialisé par la position géographique du centre .....	56
Figure 3.3 : Ajustement de la biomasse herbacée observée avec l'intégrale NDVI SPOT VEGETATION à l'échelle globale .....	58
Figure 3.4 : Ajustement de la biomasse herbacée observée avec le maximum NDVI SPOT VEGETATION à l'échelle globale .....	58
Figure 3.5 : Ajustement la biomasse herbacée observée avec l'intégrale NDVI eMODIS à l'échelle globale .....	59
Figure 3.6 : Ajustement de biomasse herbacée avec le maximum NDVI eMODIS à l'échelle globale.....	59
Figure 3.7 : Ajustement avec l'intégrale NDVI SPOT VEGETATION en zone nord-sahélienne .....	60
Figure 3.8 : Ajustement avec le maximum NDVI SPOT	

VEGETATION en zone nord-sahélienne.....	60
Figure 3.9 : Ajustement avec l'intégrale NDVI SPOT VEGETATION en zone sahélienne .....	60
Figure 3.10 : Ajustement avec le maximum NDVI SPOT VEGETATION en zone sahélienne.....	60
Figure 3.11 : Ajustement avec l'intégrale NDVI SPOT VEGETATION en zone saharienne.....	61
Figure 3.12 : Ajustement avec le maximum NDVI SPOT VEGETATION en zone saharienne.....	61
Figure 4.1 : Schème général de la méthode de régression linéaire multiple conduisant à la prévision des rendements fourragers .....	733
Figure 4.2 : Schéma de métriques calculées sur VAST .....	76
Figure 4.3 : Les valeurs observées en fonction des valeurs prédictes pour l'ajustement global.....	89
Figure 4.4 a : Graphique des résidus par ligne .....	92
Figure 4.4 b : Distribution des résidus .....	92
Figure 4.5 : Les valeurs de production herbacée observées en fonction des valeurs prévues pour le modèle de l'Azaouak.....	94
Figure 4.6 a : Graphique des résidus par ligne .....	95
Figure 4.6 b : Distribution normale des résidus.....	95
Figure 4.7 : Les valeurs observées en fonction des valeurs prévues pour le modèle du Manga2.....	96
Figure 4.8 a : Graphique du résidu par ligne .....	97
Figure 4.8 b : Distribution des résidus .....	97
Figure 5.1 : Schéma général de la démarche .....	114
Figure 5.2 : Principe de l'analyse de similarité appliquée à des séries décadiques de NDVI (en code numérique) .....	115

## Liste des sigles et abréviations

ACMAD	African Centre of Meteorological Applications for Development
AGRHYMET	Agriculture hydrologie et météorologie
AGRICAB	Agriculture Capacity Building
AMS	AgroMetShell
AP3A	Alerte précoce et Prévision des Productions agricoles
AVHRR	Advanced Very High Resolution Radiometer
BIOMASAH	BIOMAsse herbacée au SAHel
CILSS	Comité permanent Inter-Etats de lutte contre la Sécheresse dans le Sahel
CIRAD	Centre de Coopération Internationale en Recherche Agronomique pour le Développement
CRA	Centre Régional AGRHYMET
ECMWF	European Centre for Medium-Range Weather Forecasts
eMODIS	enhanced MODIS
ENVISAT	ENVIronmental SATellite - Advanced Synthetic Aperture Radar
ENVISAT-MERIS	ENVISAT Medium-spectral Resolution, Imaging Spectrometer
ESA	European Space Agency
EVI	Enhanced Vegetation Index
FAO	Food and Agriculture Organization of the United Nations
FAPAR	Fraction of Absorbed Photosynthetically Active Radiation
Fcover	Fractional cover
FEWSNET	Famine Early Warning Systems NETwork
GeoWRSI	Geo-spatial Water Requirement Satisfaction Index.
GMFS	Global Monitoring for Food Security
GPP	Gros Primary Productivity

JRC	Joint Research Centre
LAI	Leaf Area Index
LOOCV	Leave one out Cross Validation
LULC	Land use land Cover
MAD	Mean Absolute deviation
MAE	Mean Absolute Error
MARS	Monitoring Agricultural ResourceS
MEIA	Ministère de l'élevage et des Industries animales
METEOSAT	METEOrological SATellites
METOP	Series of three polar orbiting meteorological satellites
MODIS	MOderate Resolution Imaging Spectroradiometer
MRM	Modèle par Régression linéaire Multiple
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NOAA	National Oceanic and Atmospheric Administration
NOAA-CPC	NOAA-Climate Prediction Center
OMM	Organisation Météorologique Mondiale
RADAR	RAdio Detection And Ranging
RFE	Rain fall estimated
RMSE	Root Mean Square Error (ou erreur quadratique moyenne)
RRMSE	Relative RMSE
SIG	Système d'Information Géographique
SIM	SIMilarity
SNDVI	Standardized NDVI
SPI	Standardized Precipitation Index
SPIRITS	Processing and Interpretation of Remotely Sensed Image Time Series
SPOT	Satellite Probatoire pour l'Observation de la Terre
VEGETATION	VEGETATION

SPOT VGT	Satellite Probatoire pour l'Observation de la Terre — VEGETATION
STEP	The System for Terrestrial Ecosystem
UBT	Parameterization Unité Bétail Tropical
ULg	Université de Liège
VAST	Vegetation Analysis in Space and Time
VCI	Vegetation Condition Index
VITO	Vlaamse Instelling voor Technologisch Onderzoek
WMO	World Meteorological Organization
WSI	Water Satisfaction Index

## **I. Introduction générale**

Le Sahel est une large entité biogéographique s'étendant du Sénégal à l'Éthiopie (UNESCO, 1981). Cette zone naturelle fait face à des défis sécuritaires qui engendrent des pertes énormes en vies humaines et animales au Mali (terrorisme et rébellion armée), en Libye (terrorisme et guerre tribale), et dans la zone du Lac Tchad (Boko Haram au Nigeria, au Niger, au Tchad et au Cameroun) ainsi qu'à des défis climatiques et démographiques avec des conséquences sur la dégradation des ressources naturelles et la sécurité alimentaire et nutritionnelle des populations. Les communautés pastorales sont particulièrement sensibles et affectées par ces défis (DNEP, 2013). Ces dernières années plusieurs initiatives ont été prises en faveur de ces populations à l'échelle continentale et au niveau régional :

- le programme MESA « Monitoring of Environment for Security in Africa » financé par l'Union Européenne (UE), contribue à la mise en œuvre du cadre stratégique de l'Union Africaine (UA) pour le pastoralisme en Afrique. Ce cadre vise à assurer, protéger et améliorer la vie, la subsistance et les droits des éleveurs pastoraux africains (UA, 2013) ;
- le Projet Régional d'Appui au Pastoralisme au Sahel (PRAPS), financé par la Banque Mondiale, qui est une concrétisation de la déclaration de Nouakchott du 29 octobre 2013 ratifiée par les chefs d'États et de gouvernements des six pays sahéliens (Burkina Faso, Niger, Sénégal, Mauritanie et Tchad). Cette déclaration est un engagement qui vise à sécuriser le mode d'existence des pasteurs et à accroître le produit brut des activités des éleveurs de 30% (DNP, 2013).

Le régime pluviométrique du Sahel est essentiellement lié à la dynamique de mousson ouest africaine (AMMA, 2002). La faune et la flore de cet espace naturel ont dû s'adapter au cours du Quaternaire à des fluctuations climatiques entre des climats tropicaux humides et arides, voire hyperarides. La distribution des précipitations au cours de

la saison des pluies modulée par leur redistribution par ruissellement à la surface du sol constitue des facteurs déterminants de la diversité et de la production du couvert végétal (Hiernaux et Le Houérou, 2006). L'accroissement de la variabilité climatique dont les conséquences commencent à être perceptibles dans les années 1970 suite à la baisse des normales pluviométriques 1941-1960/1961-1990 (Hulme, 1992) 1971-2000/1981-2010 (Lona, 2014) a probablement joué un rôle dans les changements observés de la couverture végétale actuelle. Suivant les gradients Nord-Sud et Est-Ouest, une forte variabilité interannuelle de la production fourragère, associée à de grandes fluctuations temporelles de la disponibilité fourragère en quantité et en qualité est observée.

Il est alors nécessaire de mettre en place une méthode d'estimation de ces productions fourragères en fin de saison des pluies, en vue d'établir un Système d'Alerte Précoce Pastorale (SAPP) efficace et efficient pour les neuf mois de la saison sèche qui suit. Ce SAPP doit être capable de fournir à temps des informations sur la disponibilité du fourrage à l'échelle spatiale (locale, zonale, nationale et régionale) aux acteurs du secteur (éleveurs, décideurs politiques, ONG, techniciens, etc.).

L'espace pastoral du Niger, vaste et diversifié avec plus de 62 millions hectares d'espace pâturable (Rhissa, 2010) constitue une des forces majeures au développement du secteur de l'élevage. La diversité du couvert végétal permet à des espèces animales différentes de se côtoyer et de prospérer sur le même territoire. En effet, le pays dispose non seulement d'un atout important pour le développement socio-économique, mais aussi d'un avantage comparatif naturel indéniable à produire des bovins et des petits ruminants exportables vers les pays côtiers et l'Afrique Centrale. Le cheptel est composé pour l'essentiel de bovins, ovins, caprins, camelins, équins et asins. L'effectif est estimé selon les résultats du Recensement Général de l'Agriculture et du Cheptel 2005/2007 à 31 039 041 têtes de bétail toutes espèces confondues soit : 11 238 268 têtes de caprins (36.2%), 9 192 017 têtes

d'ovins (29,6%), 7 336 088 têtes de bovins (23,6%), 1 477 073 têtes d'asins (4,7%), 1 565 420 têtes de camelin (5%) et 230 174 d'équins (0,7%) (RGAC, 2008). Pour une valeur totale de près de 2 000 milliards de Fcfa, le Niger possède, avec son élevage, un atout important pour son développement socio-économique (Rhissa, 2010). Les activités d'élevage participent pleinement à la sécurité alimentaire et à la lutte contre la pauvreté de la majorité des ménages. Selon les statistiques disponibles, les productions animales contribuent pour près de 11,8 % à la constitution du Produit Intérieur Brut (PIB) et à 35 % du PIB agricole. Il représente une source importante de devises pour l'État et les collectivités territoriales, mais contribue également, de manière significative, au budget des ménages. L'enquête budgets-consommation des ménages indique que le secteur élevage contribue à plus de 15 % au budget des ménages. Quant à sa contribution à la satisfaction des besoins alimentaires, elle se situerait à hauteur de 25 % (OMD1, 2011). Aujourd'hui, le pastoralisme fait face à plusieurs contraintes au nombre desquelles on peut citer :

- la réduction de l'espace pastoral due à l'extension du front pionnier d'exploitation des terres (cultures, aménagements des aires protégées...). Cette extension du front pionnier est due à l'augmentation de la population (urbanisation et des besoins du marché des produits agricoles et élevage),
- des obstacles directs (conflits et difficultés de passages aux frontières) et indirects (incertitudes sur l'existence et les conditions d'accès à la ressource),
- le foncier car les politiques foncières pèsent largement sur les ressources fourragères et leur gestion,
- l'accès aux marchés,
- problèmes de santé, de sécurité, de changement climatique et la récurrence des épisodes de sécheresse viennent s'ajouter à cette liste de contraintes (Touré et al., 2011).

Il paraît urgent de disposer d'une méthode appropriée d'estimation des

ressources fourragères pour contribuer à la gestion des pâturages, à la prise de décision pour anticiper et gérer les crises pastorales ; autrement dit, de disposer d'un outil d'évaluation de pâturage, capable de procéder à une appréciation de la production fourragère et de leur réaction aux facteurs comme le broutage, le piétinement, etc.

Au Sahel, Les premiers travaux d'estimation de la biomasse fourragère à l'aide d'images satellitaires ont été réalisés au Sénégal, avec le capteur NOAA-AVHRR (Tucker et al. 1983, Tucker et al. 1985) où les auteurs ont mis en évidence une corrélation significative entre les valeurs issues des mesures au sol de biomasse et le NDVI "Normalized Difference Vegetation Index". Depuis, des corrélations similaires ont été obtenues dans d'autres régions sahéliennes, notamment au Niger (Maïdagi et al., 1987 ; Justice et Hiernaux, 1986 ; Wylie et al., 1991), au Mali (Hiernaux et Justice, 1986), au Burkina Faso (Groten et al., 1993 ; Masielli et al., 1992) et dans d'autres zones du monde (Tucker et Sellers, 1986 ; etc.). L'analyse des résultats de ces travaux de recherche indique qu'il y a encore des améliorations à apporter au niveau de la performance des modèles.

Le Centre Régional AGRHYMET(CRA), le Ministère de l'Élevage et des Industries Animales (MEIA) du Niger et le Centre de Suivi Écologique (CSE) du Sénégal réalisent depuis 40 ans des estimations de masse herbacée, dans le cadre de l'alerte précoce afin d'aider les autorités publiques à prendre des décisions pertinentes pour assurer une meilleure sécurité alimentaire tout en garantissant une meilleure gestion des ressources. Le MEIA et le CSE s'appuient sur les données mesurées *in situ* et l'indice de végétation à différence normalisée (NDVI) dont les premiers travaux datent des années 1970 (Rouse et al., 1973). Le NDVI a été exploité par plusieurs auteurs (Lobell et Asner, 2004; Huete et al., 2002; Justice et al., 1985) qui ont montré ses performances, mais aussi ses limites dans le suivi et la caractérisation de la végétation à l'échelle globale. En outre, dans le cadre du projet Alerte précoce et Prévision des Productions agricoles (AP3A), le CRA a développé le modèle

BIOMASAH d'estimation de la biomasse fourragère au Sahel basé sur le bilan hydrique. Ce modèle s'appuie sur la carte des sols (texture) et le cumul de pluies estimées par satellite. Toutefois, les résultats issus de ce modèle n'ont jamais été validés à l'aide d'une série d'observations sur 12 ans. Aussi, paraît-il impératif d'évaluer ces modèles en vue d'en proposer un plus performant.

## **Objectifs de la thèse**

Ces travaux de recherche qui s'inscrivent dans le cadre des activités du Projet AGRICAB "Agriculture Capacity Building", financé par l'Union Européenne (UE) coordonné par VITO "Vlaamse Instelling voor Technologisch Onderzoek" ont pour objectif général de contribuer à l'amélioration des méthodes d'estimation des rendements fourragers (herbacée) annuels à l'échelle nationale, tenant compte de la variabilité spatio-temporelle au Sahel. Ainsi, avant de proposer toute amélioration, il apparaît nécessaire d'une part, de tester puis valider le modèle BIOMASAH développé par AGRHYMET, d'autre part, d'analyser la performance du modèle d'estimation de la biomasse élaboré par le Ministère de l'Élevage et des Industries Animales (MEIA) du Niger. Au regard, des améliorations importantes enregistrées dans le domaine de la prévision des rendements agricoles à l'aide des méthodes statistiques combinant les rendements observés, les images satellitaires (métriques phénologiques) et des paramètres agrométéorologiques dérivés de certains modèles, il est utile de réaliser comme pour les rendements agricoles, un Modèle de Régression linéaire Multiple (MRM) d'estimation du rendement fourrager à partir de variables explicatives agrométéorologiques et d'indices de végétation provenant de la télédétection. Aussi, compte tenu de la profondeur historique des données disponibles, à l'instar de la méthode des analogues développée par les météorologistes, n'est-il pas nécessaire de tester la méthode de similarité pour estimer le rendement fourrager ? Et enfin, pour avoir un modèle de référence, il est primordial de comparer les trois modèles

notamment celui du MEIA, la Régression linéaire Multiple (MRM) et celui obtenu par l'analyse de similarité. Deux résultats majeurs sont attendus de ce travail de recherche. Il s'agit, d'une part de la mise en évidence des forces et les limites de chacun des modèles d'estimation des rendements fourragers (herbacée) et d'autre part, la mise à disposition d'un modèle amélioré qui pourra être considéré comme une nouvelle référence pour l'estimation des productions fourragères.

## **Structure de la thèse**

La thèse est articulée autour de six parties : la première (I) est relative à l'introduction générale; la deuxième (II) traite de la validation du modèle d'AGRHYMET appelé BIOMASAH ; ensuite la troisième (III) est consacrée à l'analyse la performance du modèle d'estimation de la biomasse élaboré par le Ministère de l'Élevage et des Industries Animales (MEIA) du Niger ; la quatrième partie (IV) propose une amélioration de la modélisation de la biomasse fourragère en zone sahélienne à travers la méthode de régression linéaire multiple; la cinquième (V) traite de la performance de l'analyse de similarité dans l'estimation des rendements fourragers au Niger ; et enfin la sixième (VI) procède à une comparaison des modèles.

### **1.1. Présentation de la zone d'étude**

#### **1.1.1. Situation géographique**

La zone d'étude correspond à la zone pastorale du Niger définie sur les cartes de l'Atlas «Elevage et potentialités pastorales sahéliennes» (IEMVT, 1987). Elle s'étend entre le 13° et le 16° de latitude Nord et entre 2° et 12° de longitude Est (Figure 1.1). Le choix de cette zone du Sahel pour la validation des différents modèles de biomasse fourragère est essentiellement lié à la disponibilité des données in situ. À l'instar des autres régions sahéliennes, cette zone est caractérisée par la dynamique de la mousson ouest africaine (AMMA, 2002) avec une forte variabilité spatiotemporelle des précipitations (Sivakumar et al.,

1993). Le climat est du type aride avec une pluviométrie normale variant entre 150 et 300 mm (Touré et al., 2012). La durée de la saison des pluies varie en moyenne de 60 à 120 jours pour le sahel central et occidental. Elle est en moyenne de 40 jours dans le Sahel septentrional et oriental (AGRHYMET, 2002).

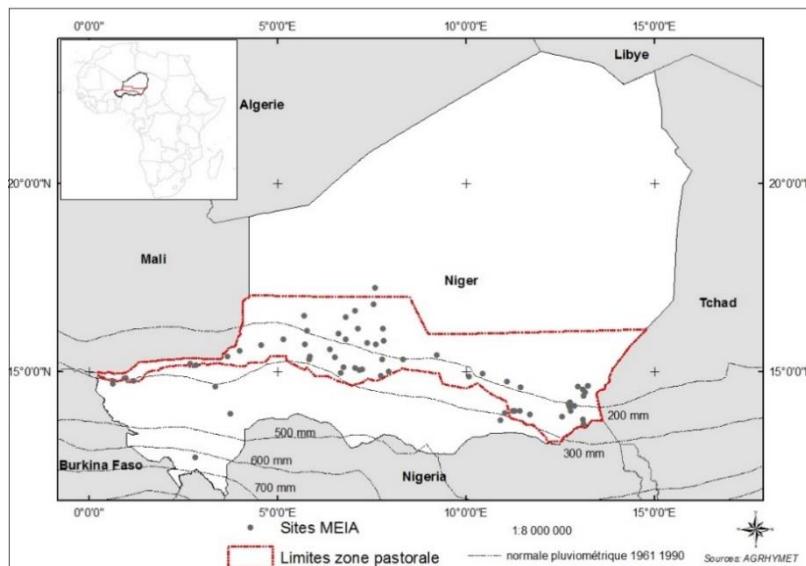


Figure 1.1 : Présentation de la zone d'étude

### 1.1.2. Régime pluviométrique

Le Niger est un vaste pays sahélien situé dans le domaine climatique de transition compris entre les zones sahariennes au Nord et Soudanienne au Sud. Le Sahel qui signifie rivage en arabe est une entité biogéographique enregistrant une pluviosité moyenne annuelle comprise 150 et 600 mm. Il est subdivisé conventionnellement en trois sous-zones : nord sahélienne (150-300 mm), sahélienne typique (300-450 mm) et sud sahélienne (450-600 mm) (Justice et Hiernaux, 1986). Au cours des quatre dernières décennies, le Sahel a connu plusieurs déficits pluviométriques à l'origine de crises de sécheresse majeure (1968-1974, 1983-1984, 2002-2003, 2005, 2009) qui ont lourdement affecté les populations et leur bétail (Touré et al., 2012). La répartition

spatiotemporelle des précipitations et leur redistribution à la surface du sol par le ruissellement sont les facteurs déterminants de la diversité du couvert végétal et de sa production interannuelle (Hiernaux et Le Houérou, 2006). L'examen de la moyenne pluviométrique 1971-2000 superposée à celle de 2001 à 2010 montre un retour des précipitations dans les zones Nord. Elle est matérialisée par une remontée des isohyètes 150, 300 et 450 mm. Cette évolution est plus marquée dans les régions septentrionales du Niger (Figure 1.2).

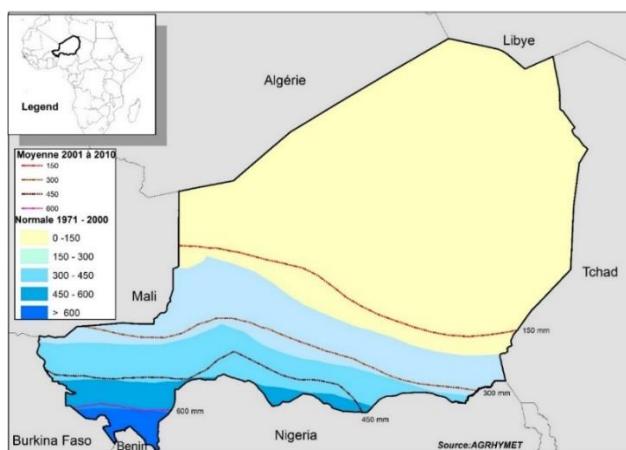


Figure 1.2 : Superposition de la moyenne 1971- 2000 et celle 2001 - 2010

### 1.1.3. Subdivision bioclimatique du Niger

De 1933 à nos jours, plusieurs auteurs ont subdivisé la zone sahélienne sur la base des moyennes pluviométriques en considérant la forte relation qui existe entre les hauteurs pluviométriques et la structure de la végétation ( Aubréville, 1949 ; Boudet, 1977 ; Le Houérou, 1980 ; Nicholson, 1981). Une subdivision a été réalisée sur la base de la moyenne 1981-2010 et en se basant sur la subdivision de (Justice et al., 1986), soit par les isohyètes 150, 300, 450, 600 mm (Figure 1.3).

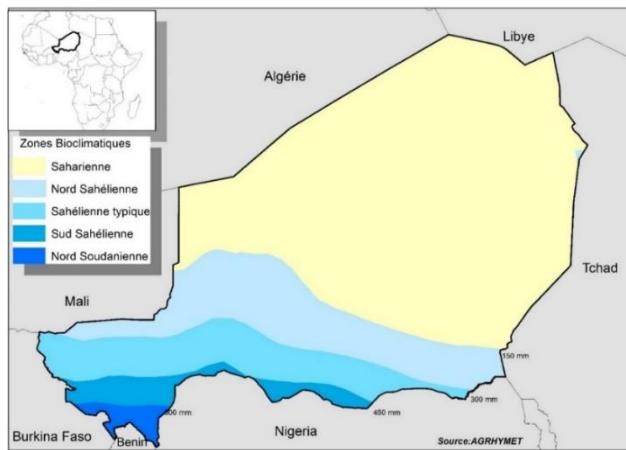
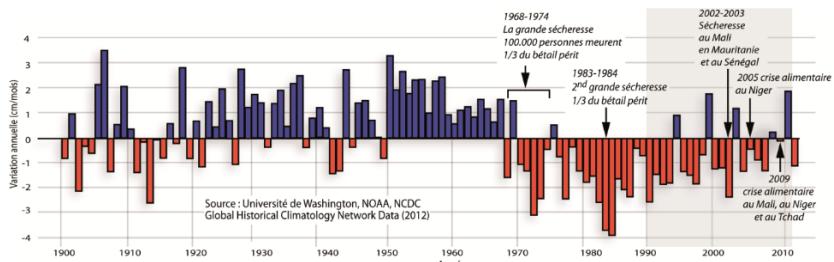


Figure 1.3 : Zones bioclimatiques du Niger

#### 1.1.4. Évolution de la pluviosité sahélienne de 1900 à 2010

En se référant aux travaux du "Joint Institute for the Study of the Atmosphere and Ocean" JISAO (2012), les variations annuelles des précipitations (cm/mois) de la zone sahélienne toute entière entre 1900 et 2010 oscillent en dents de scie montrant des périodes humides et sèches. La période allant de 1900 à 1950 est marquée par une alternance de 3 à 4 années humides suivies d'une année sèche. De 1951 à 1969, on observe une persistance d'années humides, de 1970 à 1993 une succession d'années sèches. Par contre, la période allant de 1994 à 2011 est caractérisée par une alternance d'une année humide suivie de 3 à 4 années sèches (Figure 1.4).



Source : [http://jisao.washington.edu/data\\_sets/sahel/](http://jisao.washington.edu/data_sets/sahel/)

Figure 1.4 : Évolution historique de la pluviosité au Sahel : variation annuelle en cm par mois de 1900 à 2011.

### **1.1.5. Productivité des pâturages**

#### **1.1.5.1. Productivité fourragère des zones bioclimatiques**

En se référant aux travaux de (Hiernaux et Justice, 1986; Boudet, 1977; OZER et al., 2010; Lona I., 2014), la zone saharienne qui totalise en moyenne des précipitations annuelles inférieures à 150 mm ne permet que le développement des plantes à cycle court (éphémérophytes) mais aussi un peuplement ouvert d'herbacées pérennes en particulier les graminées *Panicum turgidum* et *Aristida sieberiana*, et aussi des dicotylédones pérennes *Cornulaca monacantha*, des ligneux épars mais notable avec la concentration des eaux de ruissellement comme dans les vallées de l'Aïr, le long des Oued de l'Azaouak. Malgré le caractère marginal et très irrégulier de la production fourragère, les éleveurs nomades l'exploitent en se déplaçant avec leurs troupeaux de camelins, caprins, ovins, asins, et bovins au gré de la disponibilité des pâturages et des points d'eau. La sous zone nord-sahélienne comprise entre 150 et 300 mm dispose d'un couvert ligneux ne dépassant guère 2 % et une masse herbacée pouvant atteindre 400 kg de MS.ha-1 (Boudet, 1977). Il est important de noter que localement la concentration des eaux amène des peuplements plus denses le long des écoulements, autours des mares, des dépressions inter dunaires, et des fourrés de brousse tigrée. Dans ces cas la masse herbacée peut être beaucoup plus élevée que 400 kg MS. ha-1 sur sols sableux et peut atteindre 1 000 voire 2 000 kg MS.ha-1 avec 200 mm de pluie. Cette zone est actuellement convoitée à la fois par les éleveurs nomades et transhumants. La sous-zone sahélienne typique (300-450 mm) est caractérisée par une végétation très différenciée suivant les principales unités géomorphologiques. Par exemple, sur le terrain sablonneux, le couvert ligneux ne dépasse guère en moyenne 5 %. Il est à noter que ce recouvrement dépend aussi des sols et du ruissellement. La productivité augmente du Nord vers le Sud, elle va de 500 à 2 000 kg MS. ha-1. Vers l'isohyète 300 mm la capacité de charge annuelle est estimée à 17 UBT.Jours. ha-1 et 57 UBT.Jours.ha-1 vers l'isohyète 400 mm. La

sous-zone sud-sahélienne est plus arrosée (450-600 mm) avec un taux de recouvrement ligneux moyen variant du nord vers le sud de 5 à 30 % même si cela dépend aussi de la texture du sol, du ruissellement, et de l'occupation des sols. La masse herbacée varie de 1000-1500 kg MS.ha<sup>-1</sup> correspondant à une capacité de charge de 50 à 70 UBT.jours.ha<sup>-1</sup>(Boudet, 1977).

### **1.1.5.2. Calcul de la production de la phytomasse aérienne basé sur le Dry Matter Productivity (DMP)**

La disponibilité fourragère en quantité et en qualité constitue un des facteurs importants pour comprendre la pratique du pastoralisme au Sahel. L'intensité et la répartition spatiale des pluies déterminent la croissance végétative en saison des pluies et donc le stock de fourrage disponible pour les 9 mois de la saison sèche qui la suit. En fonction de la situation, les éleveurs anticipent ou retardent la principale transhumance pour optimiser leur système de production. Les zones d'accueil doivent faire face à un risque de concentration des animaux avec pour corollaire, les risques d'épidémie, les conflits entre agriculteurs et éleveurs, la prolifération des espèces non appétées concourant à une dégradation de l'environnement. Dès lors, au niveau régional, Il est indispensable, de renforcer le dispositif d'évaluation de la production fourragère, d'informer et de situer les déficits fourragers dans les zones pastorales sahéliennes. La productivité de matière sèche (DMP: Dry Matter Productivity) (Swinnen et al., 2015) offre la possibilité d'estimer la production de phytomasse aérienne de l'année en cours et de la comparer à une moyenne de référence sur une série temporelle de 1998 à 2012. La quantité de matière sèche (ou production de phytomasse aérienne, PPA) de la saison de pluie est calculée à partir du cumul des décades de la saison pluvieuse (juin à octobre):  $PPA = 0.01 \times \sum_{décade i}^{décade f} DMP_{décade i}$  où PPA (kg MS.ha<sup>-1</sup>) = production de phytomasse aérienne ; 0.01 = valeur de conversion en kg ; décade i =

première décade du mois de juin ; décade f = troisième décade du mois d'octobre. Le résultat ainsi obtenu est une estimation de la production aérienne totale sur l'année considérée, exprimée en kg MS.ha<sup>-1</sup>. (Figure 1.5)

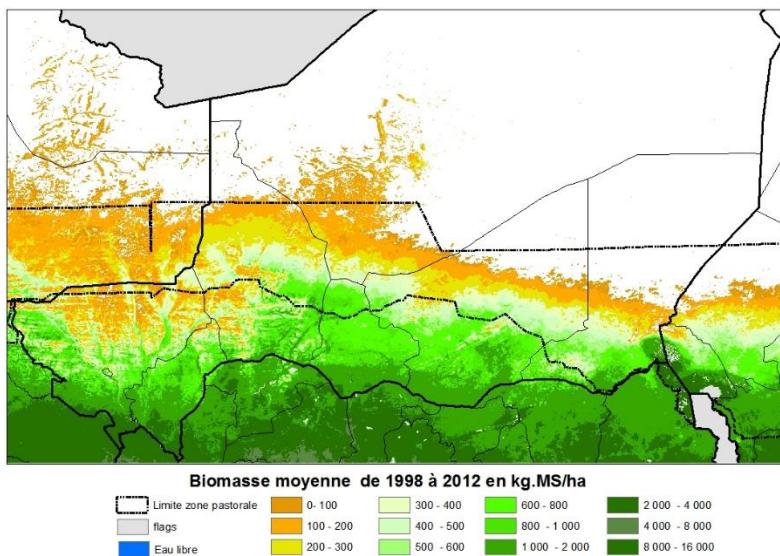


Figure 1.5 : Estimation de la production moyenne de la phytomasse aérienne (1998-2012) dérivée du DMP.

## 1.2. Évaluation de la phytomasse fourragère au sol (herbacée et ligneuse)

### 1.2.1. Généralités sur l'évaluation des pâturages au sol

La méthode d'évaluation de la phytomasse annuelle utilisée par le Ministère de l'élevage et des industries animales (MEIA) tire son origine des travaux menés au Niger et au Sénégal (Wylie et al., 1991; Diallo et al., 1991). Elle s'articule autour de trois étapes : La mesure de la masse aérienne totale in situ (sites en zone pastorale) qui se fait par la technique du double échantillonnage c'est-à-dire par coupe et par estimation visuelle ; le traitement des données d'observation de la terre qui consiste à calculer l'intégrale de la courbe de NDVI extraite de la série d'images provenant soit de NOAA AVHRR, de SPOT

VEGETATION ou de MODIS au cours de la période de croissance végétative; l'analyse statistique à travers la création d'un modèle de régression utilisant la méthode des moindres carrées linéaires. Les données de biomasse mesurées *in situ* constituent la variable dépendante (y), l'intégrale de la courbe de NDVI de la saison de croissance est considérée comme variable indépendante (x). Une équation de premier ordre du type  $y = ax + b$  obtenue est utilisée pour estimer la production.

Au Niger, la mesure de la phytomasse herbacée est réalisée en zone pastorale à l'aide de la méthode appelée de double échantillonnage qui peut être aussi qualifiée de double mesure : une mesure par fauchage (destructive) et deuvièmement (non destructive) par estimation visuelle. La caractérisation des sites est la première étape et est réalisée pour chaque site dans une aire de 3 km de côté. Cette superficie correspond à 9 pixels de résolution spatiale de 1 km<sup>2</sup>. Sur cette surface de 9 km<sup>2</sup>, les éléments clés de la physionomie du paysage sont observés et notés. Ces éléments clés sont le type de sol, la topographie, la végétation, la flore, les aménagements et mode d'exploitation par l'homme. L'ensemble de ces informations permet de définir les unités paysagères qui seront la base de la stratification du site en unités paysagères. Un échantillonnage stratifié permet statistiquement d'obtenir une variance plus faible comparativement à celle d'un échantillonnage aléatoire simple (FAO, 1981). Le rendement moyen fourrager par site est estimé par la moyenne des rendements moyens fourrager des unités paysagères composant le site.  $p_m = \frac{(p_1+p_2+p_3+\dots+p_n)}{N}$

Où : pm représente le rendement moyen de chaque unité paysagère pondérée par leur surface ; p1, p2,..., pn les rendements par échantillon ; N le nombre d'unités paysagères.

### **1.2.2. Évaluation de la biomasse fourragère sur le terrain par le MEIA**

La méthode de collecte des données *in situ* par le MEIA tire son origine

des études menées de 1986 à 1988 sur le suivi de la production fourragère par satellite et les mesures au sol de (Wylie et al., 1991). La présente étude est réalisée dans la partie pastorale du centre-nord du Niger. Afin d'intégrer l'ensemble de la diversité des zones agro-écologiques du pays, les cadres du MEIA ont installé des sites complémentaires de suivi sur toute la partie pastorale définie ci-dessous (Figure 1.7).

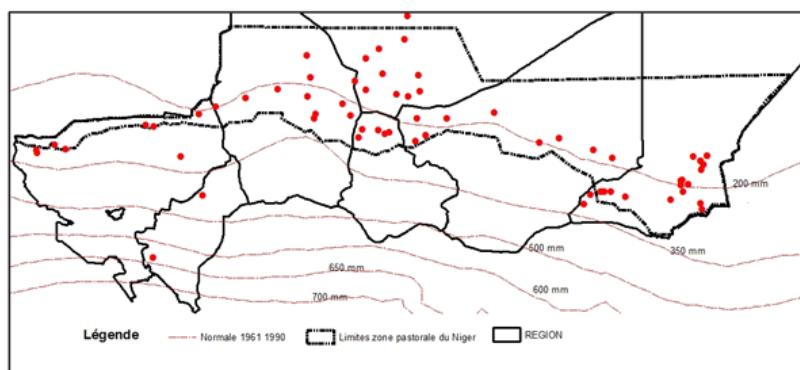


Figure 1.6 : Situation géographique des sites de relevé de végétation fourragère au Niger

#### • Évaluation de la strate herbacée

Dans la zone pastorale sur chaque site de mesure in situ, il y a 12 points de mesures qui sont placés à des distances précises suivant la ligne du Hazard. Le Hazard est le chemin suivi par l'équipe d'évaluation de la biomasse herbacée sur un site. Les itinéraires sont effectués en véhicule tout terrain, les sites d'observation et mesure sont des arrêts basés sur le compteur kilométrique du véhicule. Il y a 5 types de Hazards qui sont appliqués en fonction des caractéristiques du site (Figure 1.8 a, b, c, d, e). Au niveau de chaque point (p) de mesures 5 quadrats ou placeaux de 0,5 m<sup>2</sup> distants de 5 pas sont placés, les trois sur la ligne du Hazard et les deux sur la droite imaginaire perpendiculaire à ce Hazard et passant par le quadrat du milieu (Figure 1.8 f). La décision de procéder à la coupe est prise à travers un tirage sans remise. Au total, 60 mesures

par ligne de Hazard sont réalisées avec une moyenne de 9 à 12 coupes par site. Le rendement moyen est obtenu par une simple moyenne arithmétique des poids observés obtenus.

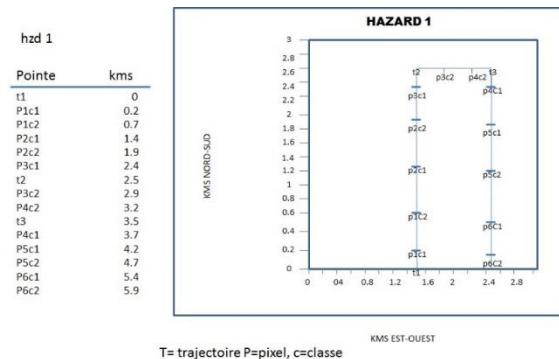


Figure 1.7 a : Schéma de Hazard 1 (source : MEIA)

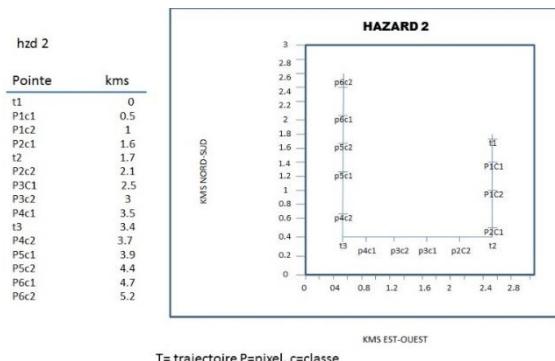
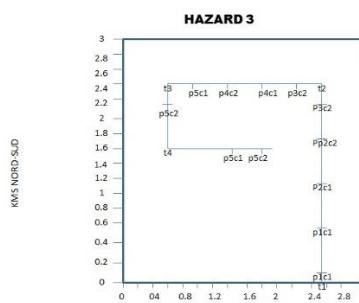


Figure 1.7 b : Schéma de Hazard 2 (source : MEIA)

**hzd 3**

Pointe	kms
t1	0
P1c1	0.1
P1c2	0.6
P2c1	1.2
P2c2	1.7
P3c1	2.2
t2	2.5
P3c2	2.7
P4c2	3.1
P4c1	3.6
P5c1	4.2
T3	4.5
P5c2	4.7
t4	5.5
P6c1	6.3
P6c2	6.8

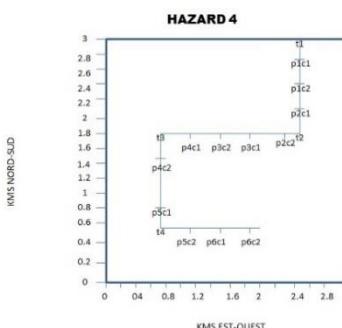


T= trajectoire P=pixel, c=classe

Figure 1.7 c : Schéma de Hazard 3(source : MEIA)

**hzd 4**

Pointe	kms
t1	0
P1c1	0.2
P1c2	0.7
P2c1	1.1
t2	1.5
P2c2	1.6
P3c1	2.2
P3c2	2.7
P4c1	3.1
T3	3.5
P4c2	3.6
P5c1	4.3
t4	4.5
P5c2	4.8
P6c1	5.2
P6c2	5.7

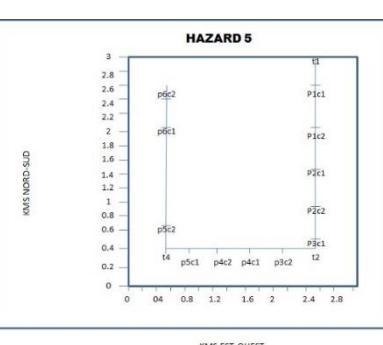


t= trajectoire; P=pixel; c= classe

Figure 1.7 d: Schéma de Hazard 4(source :MEIA)

**hzd 5**

Pointe	kms
t1	0
P1c1	0.1
P1c2	0.6
P2c1	1.1
P2c2	1.6
P3c1	2.3
t2	2.5
P3c2	2.8
P4c1	3.2
P4c2	3.7
P5c1	4.2
t4	4.5
P5c2	4.7
P6c1	6.2
P6c2	6.7



t= trajectoire; P=pixel; c= classe

Figure 1.7 e : Schéma de Hazard 5 (source : MEIA)

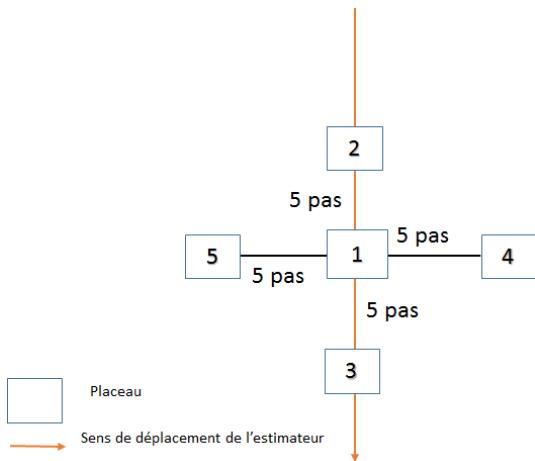


Figure 1.7 f : schémas des Hazards 1 à 5 et la disposition des plateaux sur les points de mesures

- **Évaluation de la strate ligneuse**

La strate ligneuse est prise en compte surtout dans les enclaves pastorales qui ne font pas partie des sites exploités. Un comptage direct des ligneux est réalisé dans des aires circulaires ou rectangulaires d'une superficie de 1/8, 1/4, 1/2 ou 1 ha (en fonction la densité du peuplement). Par exemple, dans le cas de l'aire circulaire, une corde de 56,40 m est tendue (aire d'un cercle de 56,4 m de rayon = 1 ha) puis on tourne progressivement pour faire un cercle complet dans lequel les individus ligneux sont décomptés. Selon, le nombre de personnes disponibles, un observateur est placé à 1 m environ devant le collecteur de données, de façon à ce que celui-ci ait l'observateur sous son contrôle. Toutes les essences ligneuses rencontrées par l'observateur sont comptabilisées en évitant le double comptage. Le calcul de la masse foliaire des ligneux est faite en utilisant la méthode de la plante de référence. Cette méthode consiste à choisir par espèce ligneuse un individu test, selon les espèces, il peut s'agir d'un arbre, un arbuste ou un buisson (moyen), ce ligneux de référence est caractérisé (taille, densité des feuilles, forme, etc.). La masse des feuilles est exprimée par

$m^2$  de houppier, ensuite, le nombre de ligneux équivalents à l'individu de référence est déterminé dans le cercle. Enfin, la récolte des feuilles et des branchettes consommables est réalisée sur les individus de référence.

- **Les conditions de mesure de la biomasse herbacée épigée par le MEIA**

Les sites de relevés de phytomasse fourragère épigée du Niger ne sont pas mis en défens car ils font l'objet d'une exploitation naturelle par les animaux au même titre que les autres espaces pâturés du pays. Aussi, le prélèvement réalisé par les animaux n'est pas pris en compte lors du calcul de rendement moyen, ce qui peut constituer un biais qu'il paraît important de souligner. En effet, dans une étude réalisée dans le sahel malien, Hiernaux et Justice (1986) indiquaient que l'impact de la pâture par les animaux peut en partie expliquer les valeurs élevées de biomasse potentielle observées comparées à la biomasse mesurée. Par conséquent, nous pensons que c'est la biomasse herbacée résiduelle qui est mesurée. Un second biais proviendrait de la gestion des missions d'évaluation réalisées par les services techniques : celles-ci sont en réalité effectuées selon la disponibilité des ressources financières. Par ailleurs, il ressort de l'examen du Tableau 1.1 que les mesures sont réalisées *in situ* entre la deuxième décade du mois de septembre et la fin du mois d'octobre, or celles-ci ne correspondent pas toujours à la période idéale de mesure. En effet, suite à une étude menée au Niger, Justice et Hiernaux (1986), ont clairement montré que la date du maximum de croissance de la végétation se situe entre deuxième décade du mois d'août et la deuxième décade du mois de septembre (Annexe 1.1). Par conséquent, le fait que les dates de mesures réalisées par le MEIA ne coïncident pas toujours avec la période du maximum de végétation, peut donc constituer un autre biais.

Tableau 1.1 : synthèse des décades de début et de fin des mesures in situ de la biomasse herbacée épigée au Niger de 2000 à 2012

Années	Début des mesures	Fin des mesures
2000	1 <sup>ere</sup> d sep	3 <sup>e</sup> d sep
2001	1 <sup>ere</sup> d sep	3 <sup>e</sup> d oct
2002	2 <sup>e</sup> d sep	3 <sup>e</sup> d oct
2003	3 <sup>e</sup> d sep	3 <sup>e</sup> d oct
2004	1 <sup>ere</sup> d sep	3 <sup>e</sup> d oct
2005	1 <sup>ere</sup> d sep	3 <sup>e</sup> d oct
2006	3 <sup>e</sup> d sep	3 <sup>e</sup> d oct
2007	2 <sup>e</sup> d sep	3 <sup>e</sup> d oct
2008	2 <sup>e</sup> d aou	3 <sup>e</sup> d sep
2009	2 <sup>e</sup> d sep	3 <sup>e</sup> d oct
2010	3 <sup>e</sup> d sep	2 <sup>e</sup> d oct
2011	3 <sup>e</sup> d sep	2 <sup>e</sup> d oct
2012	3 <sup>e</sup> d sep.	2 <sup>e</sup> d oct

d : décade ; aou ; Août ; sep : Septembre ; oct :Octobre.

- **Filtrage des données aberrantes**

En statistique ou dans toute étude scientifique, la fiabilité des résultats dépend de la qualité des données utilisées pour faire les analyses. Aussi, avant toute analyse statistique, il est essentiel de s'assurer de la qualité des données. Grâce à un examen minutieux de la série de données ou de l'échantillon, afin de rechercher d'éventuelles données aberrantes, de manière à les éliminer systématiquement. Les données aberrantes ne doivent pas être toujours éliminées. Elles peuvent représenter des exceptions, dont il est bon d'en tenir compte.

Il paraît cependant fondamental de s'accorder sur la définition d'une donnée aberrante, puis de s'assurer qu'elle l'est réellement avant de procéder à son élimination.

Selon Anderson et al (2011), une donnée est dite aberrante, quand elle

diffère singulièrement des autres de la série considérée comme normale, soit parce qu'elle est exagérément grande (donnée aberrante majeure) ou petite (donnée aberrante mineure). Généralement, dans un tableau ou sur une courbe, elle attire l'attention du chercheur, car elle s'éloigne des autres données. Le plus souvent, elle est liée à des erreurs de mesure, de saisie, de calcul, etc.

De nombreuses techniques statistiques existent pour identifier les données aberrantes. La plus connue est celle qui est basée sur la détermination des limites supérieure et inférieure de la série de données statistiques en s'appuyant sur les écarts inter quartiles. Les données sont aberrantes quand elles sont inférieures à  $Q1 - 1,5 \times IQR$  ou supérieures à  $1,5 \times IQR + Q3$  où  $Q1$  représente le quartile 25% ;  $Q3$  le Quartile 75% et  $IQR$ , l'Écart Inter Quartile (Sullivan, 2013 ; Brase et Brase, 2009). Toutes les valeurs qui vont au-delà de ces limites sont considérées comme aberrantes.

Des tests existent aussi pour l'identification des données aberrantes :

- Le test de Cochran qui permet de détecter les valeurs aberrantes en termes de dispersion, il s'applique sur l'écart type des mesures (Burke, 2001);
- Le test de Grubbs qui détecte les valeurs aberrantes en termes de dispersion de moyenne, le principe de ce test est de comparer la valeur absolue des écarts réduits (Grubbs, 1969).

En plus de ces évaluations quantitatives, il est indiqué d'effectuer des évaluations qualitatives de la donnée. Celles-ci nécessitent de connaître les conditions dans lesquelles elles ont été réalisées.

La première opération a permis d'examiner qualitativement les données du MEIA, année par année, avec le concours de la personne responsable de la collecte de ces données depuis une trentaine d'années. Ainsi, les dates de mesures de ces données ont été examinées, année par année et site par site.

Pour le filtrage des données, il a été retenu d'effectuer un seuillage sur base du maximum du NDVI. Ce choix s'explique par les études

réalisées par Wylie *et al.* (1991) qui ont montré la forte corrélation ( $r^2=0,86$ ) entre la masse fourragère mesurée *in situ* et le maximum du NDVI de la saison de croissance de la végétation de la même année. En plus, le NDVI est indicateur d'état de la végétation, donc il exprime directement son état. Par ailleurs, les données pluviométriques sont souvent utilisées pour l'évaluation de la végétation (Ndiaye et al., 2015). Le paramètre pluie est certes un facteur limitant de la production végétale dans toute la zone sahélienne, mais reste et est un indicateur de cause.

Pour éliminer objectivement les données qui s'écartent des données normales, nous avons donc utilisé le maximum de NDVI de la saison de croissance de la végétation de chacun des sites et suivant les années. Une représentation graphique sous forme de nuage de points des rendements fourragers en fonction du maximum NDVI (compte numérique, CN) a été réalisée par la suite. De manière empirique, l'équipe identifie les points particuliers du graphique, les discute avant de décider de l'éliminer ou pas. Exemple des Figure1.9 a et b illustrent les cas de l'année 2012. Sur les 584 mesures de masse fourragère enregistrées par le MEIA, sur la période allant de 2000 à 2012, la première opération de filtrage a éliminé 26% des données.

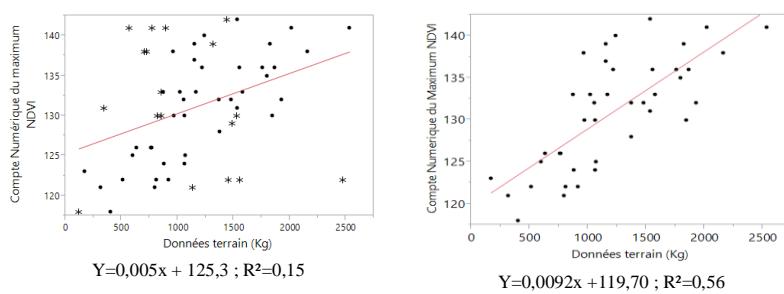


Figure1.8: Exemple de l'année 2012 première étape du filtrage des données

La deuxième opération de filtrage est basée sur l'analyse du graphique des résidus en fonction des valeurs prédictes (Figures 1.10). Il est procédé à l'élimination année par année des observations (points représentés par une étoile, \*), dans la Figure 1.10 a, dont les résidus sortent des limites inférieure et supérieure à l'équivalent de l'écart type des valeurs mesurées sur le terrain. La Figure 1.10 b indique l'exemple de l'année 2006 après le filtrage.

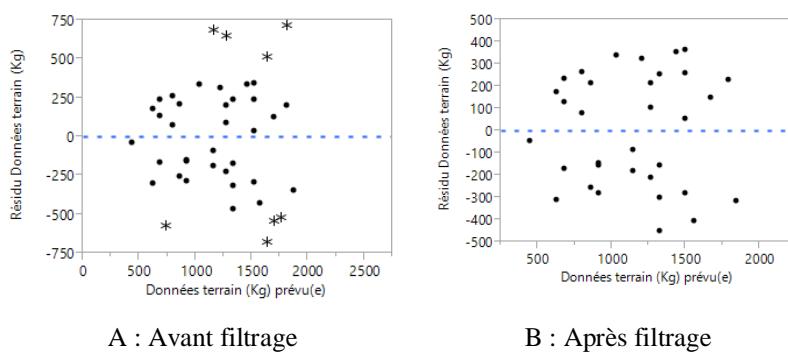


Figure 1.9 : Exemple de l'année 2006 sur la deuxième étape  
du filtrage des données

Ce qui a permis d'éliminer 29% des 452 enregistrements restants. En résumé, 49 % des enregistrements de la série de départ ont été éliminés. Le filtrage des données aberrantes a ainsi permis d'obtenir une base de données contenant 319 observations.

- Relation entre les cumuls pluviométriques annuels et la phytomasse aérienne herbacée mesurée**

Le coefficient de détermination ( $R^2$ ) de la relation entre la phytomasse aérienne herbacée mesurée et le cumul de pluies estimées par satellite (RFE2) est de 0,16. Ce qui signifie qu'avec ces données, le cumul pluviométrique n'explique que 16 % de la variation du rendement fourrager dans la zone pastorale du Niger. Selon Breman et de Ridder (1991), l'eau, le phosphore et l'azote constituent les principaux facteurs

limitants à la production de phytomasse au Sahel. Le cadre théorique de la relation entre le cumul de pluies et la phytomasse aérienne établi par Breman H. (1982), comparé à la relation entre le cumul de pluies estimées par satellite et la phytomasse mesurée (Figure 1.11) indique que les données observées sont en général sous-estimées de 33% et 40% respectivement au niveau des isohyètes 250 mm et 400 mm. Cependant, les observations qui présentent des valeurs supérieures au cadre théorique sont au nombre de 30 soit environ 8 % du total. Le non-respect de la période optimale de mesure (maximum de la végétation) et l'impact de la pâture éventuelle des animaux avant les mesures peuvent expliquer cette tendance à la sous-estimation du rendement. Par ailleurs, les pluies utilisées sont estimées par satellite, donc elles présentent quand même des limites par rapport à la réalité du terrain.

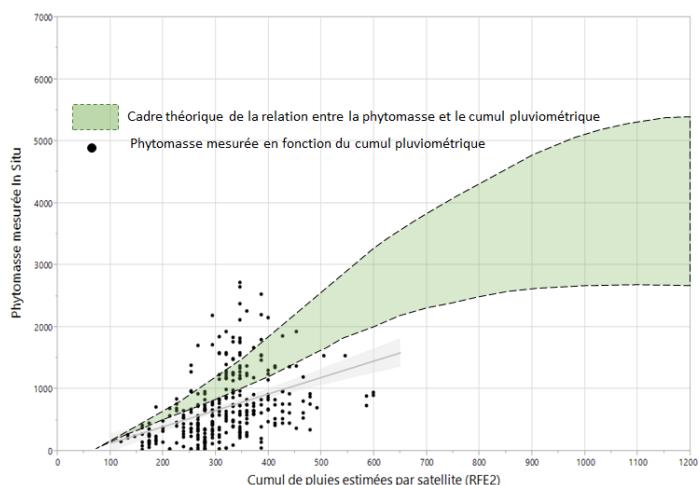


Figure 1.10 : Relation entre la phytomasse aérienne herbacée mesurée et le cumul des pluies estimées par satellite superposée à la relation théorique entre le cumul pluviométrique et la phytomasse

### 1.2.3. Distribution spatiale des sites de mesures de phytomasse aérienne herbacée du MEIA

Pour apprécier la représentativité des sites de mesure de la biomasse fourragère du Niger, il est nécessaire d'analyser la variation inter

saisonnier de la productivité. Les critères sur lesquels le ministère se base pour choisir les sites au Niger sont mal connus pour le moment. Dans les différents tableaux de synthèse, nous avons les proportions par zone agro écologique (Tableau 1.2) montrant que 80 % des sites sont situés dans l’Azaouak et le Manga (Figure1.12). Suivant les types de sol (Tableau 1.3), on remarque que 48 % des sites sont situés sur les Arenosols cambics (Qc7-1a) et les Arenosols luvics (Ql1-1a) sur les 16 types présents. Selon les zones bioclimatiques (tableau 1.4) la zone nord sahélienne et la zone sahélienne typique contiennent 90 % des sites. Il ressort de l’analyse de ces tableaux que nous avons 68 sites dont un a été éliminé à cause d’une erreur de positionnement géographique.

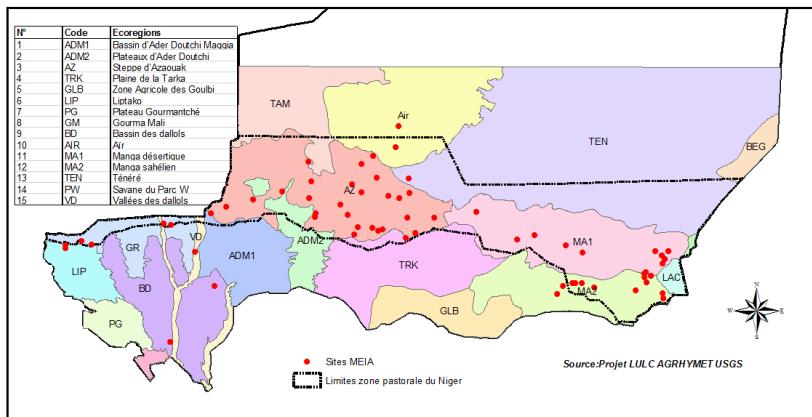


Figure 1.11 : Les écorégions du Niger

Tableau 1.2 : La répartition des sites de mesure de phytomasse herbacée aérienne selon les écorégions  
(rendement fourrager de 2001 à 2012)

Écorégion	Nombre de Site	Pourcentage de sites	Min (kg.ha <sup>-1</sup> )	Max (kg.ha <sup>-1</sup> )	Moyenne (kg.ha <sup>-1</sup> )	ET
ADM1	1	1 %	358	358	358	0
Air	2	3 %	423	960	692	380
AZ	27	39 %	153	2644	1474	722
BD	2	3 %	662	1532	1097	615
GR	3	4 %	205	657	361	256
LAC	1	1 %	519	519	519	0
LIP	3	4 %	730	2197	1230	838
MA1	14	20 %	205	1298	758	329
MA2	14	20 %	170	1334	757	382
TEN	1	1 %	647	647	647	0
VD	1	1 %	1366	1366	1366	0

ADM1 :Ader Doutchi Magia1 ; AZ :Azaouak ; BD : Bassin du Dallol ;GR :Gourma ;LP : Liptako ;MA1 : Manga1 MA2 : Manga2 ;TEN : Ténéré ; Vallée du Dallol ; ET : Écart Type ; Min :minimum masses herbacées ; max : Maximum Masses herbacées ;

Tableau 1.3 : La répartition des sites de mesures de phytomasse herbacée aérienne par type de sol

Sol FAO	Nombre Sites	Pourcentage de sites	Min (kg.ha <sup>-1</sup> )	Max (kg.ha <sup>-1</sup> )	Moyenne (kg.ha <sup>-1</sup> )	ET
Be28-1a	2	3 %	657	2197	1427	1089
Bv7-a	1	1 %	1366	1366	1366	0
Ge5-1a	4	6 %	1500	2260	1854	385
Je33-1/3a	1	1 %	1334	1334	1334	0
Qc1	7	10 %	170	1306	809	403
Qc1-1a	1	1 %	898	898	898	0
Qc7-1a	15	22 %	205	1245	654	285
Ql1-1a	18	26 %	153	2644	1445	752
Ql10	1	1 %	221	221	221	0
Ql10-1a	2	3 %	358	741	550	271
Ql11-1a	2	3 %	662	1532	1097	615
Ql6	2	3 %	730	762	746	23
Ql6-1a	1	1 %	205	205	205	0
Re35-a	8	12 %	647	2606	1190	647
Vc15	1	1 %	237	237	237	0
Water	1	1 %	519	519	519	0
Y4-1a	2	3 %	423	960	692	380

Définition des codes sol FAO : [www.fao.org/docrep/019/as357e.pdf](http://www.fao.org/docrep/019/as357e.pdf); ET : Écart Type

B = Cambisols ; G= Gleysols ; J = Fluvisols ; Q =Arenosols ; R =Regosols ; V= Vertisols; Y = Yermosols ; L =Luvic

Tableau 1.4 : la répartition des sites de mesure de phytomasse herbacée aérienne par zone bioclimatique

Zone	Nombre de sites	Pourcentage de sites	Min (kg.ha <sup>-1</sup> )	Max (kg.ha <sup>-1</sup> )	Moyenne (kg.ha <sup>-1</sup> )	E T (kg.ha <sup>-1</sup> )
Saharienne	12	17,39 %	205	1245	720	319
Nord sahélien	50	72,46 %	153	2644	1110	705
Sahélienne typique	6	8,70 %	662	2197	1091	598
Sud-sahélienne	1	1,45 %	1532	1532	1532	-

ET : Écart Type

### **1.2.4 Distribution temporelle des mesures de phytomasse herbacée aérienne.**

Les opérations de filtrages des mesures de phytomasse herbacée aérienne, ont permis d'obtenir une base de données de 319 mesures (tableau 1.5). L'analyse de la distribution temporelle des enregistrements de cette base de données indique qu'aucun site ne présente en continu, de mesures propres sur les 12 ans, soit parce qu'il n'y a pas eu du tout de relevé sur le site ou parce que la donnée mesurée est aberrante. Les sites A30 et A44 sont les plus fréquents dans la base de données, ils présentent chacun 9 mesures. Le tableau 1.5 indique aussi une variation du nombre de mesures suivant les années. Les années 2008, 2010 et 2011 présentent plus de mesures avec respectivement 36, 35 et 40. L'année 2003 ne dispose que de 12 mesures ce qui pose le problème du nombre minimum de sites nécessaires pour établir un modèle représentatif de l'année dans un territoire pastoral aussi vaste que celui du Niger.

Les codes des sites ont été attribués par le MEIA pour faciliter le traitement des données. Les coordonnées et les localités correspondantes qui sont associées à ces codes sont consignées dans le tableau de l'annexe 1.2, permettent de localiser facilement les sites.

Tableau 1.5 : Disponibilité des mesures de phytomasse herbacée aérienne par site suivant les années

Code Sites	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	Total
A105		1			1	1		1			1		5
A106	1				1				1		1		4
A11	1	1				1	1				1		5
A110			1		1	1					1	1	5
A111				1	1			1	1				4
A112					1				1				2
A113			1						1			1	3
A114					1	1			1				3
A115			1			1	1						3
A117			1		1	1		1		1		1	6
A118							1				1		2
A121										1	1		2
A122								1		1			2
A13	1	1		1		1		1		1	1	1	8
A15	1	1		1		1		1			1	1	7
A16	1			1	1	1		1	1		1	1	8
A17							1	1	1	1	1	1	6
A19								1		1	1		3
A20				1	1	1		1		1	1		6

Code Sites	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	Total
A22	1			1		1					1	1	5
A23		1				1	1				1		4
A24	1			1	1	1	1	1	1			1	8
A26	1	1		1				1		1			5
A3	1	1			1			1	1	1	1	1	8
A30			1	1	1	1	1	1		1	1	1	9
A32				1			1	1			1	1	5
A33	1	1			1	1	1			1	1	1	8
A40		1		1			1			1	1		5
A41			1		1		1	1					4
A42		1	1			1	1						4
A43		1			1	1	1			1			5
A44	1			1			1		1	1	1	1	7
A45	1	1			1	1		1	1	1	1	1	9
A46		1			1	1		1	1	1	1		7
A47	1					1	1			1	1		5
A48					1		1		1		1	1	5
A49	1	1		1	1					1		1	6
A5	1						1	1	1	1	1		6
A50			1		1			1		1		1	5

Code Sites	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	Total
A51				1	1				1		1	1	5
A52	1			1	1		1		1	1		1	7
A61					1		1			1	1		4
A62				1	1			1			1		4
A64	1	1				1		1			1		5
A65				1		1		1		1	1		5
A66		1					1	1	1	1	1	1	7
A67		1		1						1			3
A68	1	1		1	1			1	1	1			7
A69		1			1			1		1			4
A7					1	1			1	1	1		5
A70		1		1			1	1	1		1	1	7
A72	1												1
A73	1				1			1	1	1		1	6
A77	1	1		1	1	1		1				1	7
A80										1		1	2
N82			1				1						1
N85		1	1							1			3
n87		1	1								1		3
n88			1	1	1			1		1	1		6

Code Sites	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	Total
N93			1										1
n95			1		1			1		1	1	1	6
n96		1		1						1	1		4
n97				1	1			1					3
n98		1		1	1			1		1	1		6
P08	1	1									1		3
Total	22	28	12	26	31	24	23	36	17	35	40	25	319

## **II. Validation du modèle d'évaluation de la biomasse herbacée des parcours du Sahel (BIOMASAH) au Niger**

### **2.1. Introduction**

En Afrique Subsaharienne, l'estimation de la production des pâturages a été réalisée par plusieurs programmes, plus particulièrement dans les systèmes d'alerte précoce (AGRHYMET, 1992). Les approches mises au point comprennent les méthodes de mesure au sol combinées ou non à l'imagerie satellitaire. Ces techniques ont permis l'évaluation à grande échelle de la production des parcours (Justice et al., 1989). La disponibilité croissante des données d'observation de la terre de résolutions moyenne et basse, depuis les satellites de type NOAA, SPOT VEGETATION, MODIS et METOP associée à la disponibilité des données de masse herbacée mesurées *in situ* devrait faciliter la modélisation de cette estimation sur la base de modèles simples. L'absence de données de terrain ou la difficulté d'obtenir des données sur des régions entières et sur plusieurs années a conduit à l'élaboration au Sahel d'un modèle d'estimation spatialisée des productions herbacées (BIOMASAH) par le Centre Régional AGRHYMET (CRA). Ce modèle qui s'appuie sur les approches développées dans le cadre du programme « Productivité des Pâturages Sahéliens » au Mali (PPS) n'a jamais fait l'objet d'une validation par manque de données mesurées sur le terrain. Néanmoins, les produits dérivés de ce modèle ont été utilisés dans l'évaluation des campagnes agricoles et l'étude des zones à risques (Andrea et al., 2002). Cette approche a été rendue possible grâce à l'intégration des informations fournies par le modèle de simulation de la production BIOMASAH dans un système d'information géographique (SIG). Le modèle de production est, à l'origine, basé sur des relations empiriques entre le bilan hydrique, le bilan azoté et la production de matières sèches des herbacées dans les pâturages. Les intrants du SIG se composent des cartes numériques de l'Atlas pastoral de l'Institut d'Élevage et de Médecine Vétérinaire des

pays Tropicaux (IEMVT, 1987) qui ont été numérisées et des données de pluviométrie annuelle spatialisées disponibles dans la base de données du Centre Régional AGRHYMET CRA. Les données de phytomasse aérienne herbacée sont collectées par les services du Ministère de l'Élevage et des industries Animales du Niger (MEIA) depuis 1989. Les résultats du modèle sont finalement validés grâce à ces données historiques du terrain. L'objectif de ce chapitre est d'utiliser les observations au sol du MEIA pour valider les sorties du modèle biomasse BIOMASAH sur la période 2001 à 2011 à travers des tests de comparaison de moyennes et des calculs de coefficients de corrélation paramétrique et non paramétrique.

## **2.2. Matériel et méthodes**

### **2.2.1. Matériel**

L'étude a été réalisée en utilisant comme données d'entrée, les pluies estimées par satellite (RainFall Estimates version 2 RFE2) décennales de 3 km de résolution et la table de texture des sols. Les sorties sont comparées aux mesures de biomasse collectées par le Ministère de l'Élevage et des Industries Animales (MEIA). La comparaison de moyennes se fait à l'aide de tests paramétriques et non paramétriques.

### **2.2.2. Description du Modèle BIOMASAH**

L'approche utilisée s'appuie sur les résultats du projet Productivité des Pâturages Sahéliens (Penning et Djitéye, 1982) . Les auteurs de cette étude ont proposé trois échelles de travail pour l'évaluation de la production herbacée :

1. *La méthode d'évaluation globale* de la production fourragère basée sur une quantité limitée d'informations et sans travail préalable sur le terrain. Elle est basée sur l'exploitation des informations :
  - Climatiques : pluviosité, rayonnement solaire, température et humidité atmosphérique ;

- Informations biologiques des plantes : annuelles/pérennes ; C3/C4 ;
2. *L'évaluation semi-détaillée* qui s'effectue sur la base d'observations de terrain, ainsi que sur l'estimation théorique de la production moyenne par unité de paysage et sur son taux d'azote. Les observations permettent d'ajuster et compléter l'estimation théorique ;
  3. *L'évaluation détaillée* similaire à la *semi-détaillée*, mais différant sur l'échelle portant sur une région plus petite (Terroir) avec des données de terrain qui ont les mêmes contributions que dans le cas précédent mais beaucoup plus fines.

L'application développée dans le cadre du modèle BIOMASAH est intermédiaire. Elle utilise à la fois des éléments de deux méthodes (globale et semi-détaillée). L'unité de base de calcul de la biomasse herbacée est définie par le croisement des entités de potentialités pastorales de la carte des potentialités pastorales de l'Institut d'Élevage et de Médecine Vétérinaire de pays Tropicaux (IEMVT) et les pixels de l'image de la pluviosité annuelle estimée (5 km de résolution). Le calcul du bilan hydrique est réalisé dans chaque unité pastorale préalablement classée en unités géomorphologiques correspondant à des ensembles relativement homogènes : détritique, sablonneux, fluviatile et lacustre (Breman et De Ridder, 1991), codifiée sur leur texture définie dans l'étude des sols des pays du CILSS (AGRHYMET, 2001). Un coefficient de ruissellement est ensuite attribué à chaque unité composée des unités pastorales et des unités géomorphologiques. L'infiltration dans chaque unité est calculée suivant la formule :  $I = P(1 - CR)$ , où I représente la quantité d'eau infiltrée (mm/an), P la pluviosité annuelle (mm/an), et CR le coefficient de ruissellement. Les coefficients de ruissellement utilisés sont des valeurs annuelles moyennes par type de sol.

Dans la méthode, le facteur limitant pour la croissance des végétaux est

l'eau quand l'infiltration est inférieure à 250 mm (Penning et al., 1982).

En revanche, l'azote et le phosphore sont déterminants quand l'infiltration est supérieure à ce seuil de 250 mm.

Intégrée dans un SIG, l'application calcule la biomasse herbacée dans chaque unité sur la base des relations ci-dessous fournies en clair et/ou sous forme de graphique dans le manuel sur les pâturages tropicaux de Breman et de Ridder (1991): Pour les zones où le facteur limitant est l'eau, on utilise l'équation 1 ou 2 en fonction du coefficient de ruissellement :

$$\text{Éq.1 (CR=0)} : B_1 = 5.11 * \text{Pluviométrie totale annuelle} - 48.28 \quad (B_1)$$

$$\text{Éq.2 (CR=0.5)} : B_2 = 2.37 * \text{Pluviométrie totale annuelle} - 216 \quad (B_2)$$

Où B1 : biomasse en condition de ruissellement (CR= 0) et B2 : biomasse en condition de ruissellement (CR= 0.5). Pour toutes les autres valeurs de coefficient de ruissellement, la production de biomasse (B) (Kg MS. ha<sup>-1</sup>) est déduite par interpolation :

$$B = B_1 \frac{CR * (B_1 - B_2)}{0,5}$$

Pour les zones où le facteur limitant est l'azote, les pertes annuelles du système sont évaluées en fonction de la pluviosité moyenne (Penning et al., 1982). De même, les quantités moyennes d'azote (Nb en kg) dans la biomasse herbacée aérienne sont évaluées suivant des normes établies. On estime qu'au Sahel, 1 mm de pluie apporte 8,3 g.ha<sup>-1</sup> d'azote aux pâturages. Pour la fixation biologique, elle est de 0,02 kg N. ha<sup>-1</sup> par pourcent de légumineuses présentes. Les bactéries associées aux graminées et les bactéries libres apportent respectivement 0,013 kg et 0,025 kg d'azote par kg de matière sèche de biomasse. Avec toutes ces contributions et dans l'hypothèse que les légumineuses contribuent au moins pour 5 %, l'équation pour Nb est : Nb = 0,0083 \* I / (f - 0,13), f représentant les pertes annuelles (l'annexe 2.1) et I l'infiltration. Le

taux d'azote est déterminé par interpolation des situations d'équilibre et de ruissellement données par les équations suivantes où P représente la pluviosité annuelle considérée :

$$A_1(CR = 0) = 4,119 + \frac{15,48}{1 - 0,01546 * P} ; A_2(CR = 0,5) = 9,913 + \frac{217}{1 - 0,00473 * P}$$

Le taux d'azote A

$$A = A_1 + \frac{CR*(A_2 - A_1)}{0,5}$$

où  $A_1$ : taux d'azote dans la biomasse en condition de ruissellement nul CR = 0 ;  $A_2$ : taux d'azote dans la biomasse en condition de ruissellement CR = 0,5

La connaissance de la quantité moyenne d'azote dans la biomasse aérienne et du taux d'azote de la strate herbacée à la fin de la saison de croissance permet une estimation de la biomasse produite en kg.MS<sup>-1</sup>.ha<sup>-1</sup> selon l'équation : B=1000 Nb/A.

### **2.2.3. Méthode de validation**

#### **2.2.3.1. Données issues du modèle**

L'application du calcul de la biomasse a été développée sous le programme Avenue dans le logiciel ArcView GIS 3.2. Un script est aussi disponible sous le logiciel visualcarte<sup>1</sup>. Ce programme, écrit sous forme d'extension a été conçu de manière à ce que les données de pluies estimées soient les seules qui changent chaque année, les autres paramètres d'entrée sont fixés (sols et unités pastorales). Nous avons utilisé les pluies estimées (RFE2) RainFall Estimate version2

---

<sup>1</sup> A.A.V.V. — Pj AP3A (2001) — VisualCarte — Système de Gestion de la Cartographie Thématique Version Beta (CD\_ROM. ISBN : 88-900502-8-4.

développées par l'institution Climate Prediction Center (NOAA-CPC), (Novella et Thiaw, 2012). Ces données présentent l'avantage d'être générées avec le même algorithme sur la période concernée par la validation (2001-2011) et elles sont gratuitement téléchargeables à partir du site internet de FEWS NET<sup>2</sup>. Les images décadiques RFE2 sont caractérisées par une résolution spatiale de 3 km x 3km. Le prétraitement a été fait à l'aide du logiciel libre Quantum GIS (QGIS), cette opération consiste à faire les cumuls annuels des images décadiques correspondant à la saison des pluies (mai à octobre). Ensuite, ramener à la taille du pixel (5 km) compatible avec le programme et enfin sauvegarder le fichier en format GIS (ERDAS imagine). Les images de 2001 à 2011 du cumul annuel de pluies estimées sont utilisées dans le modèle (Figure 2.1).

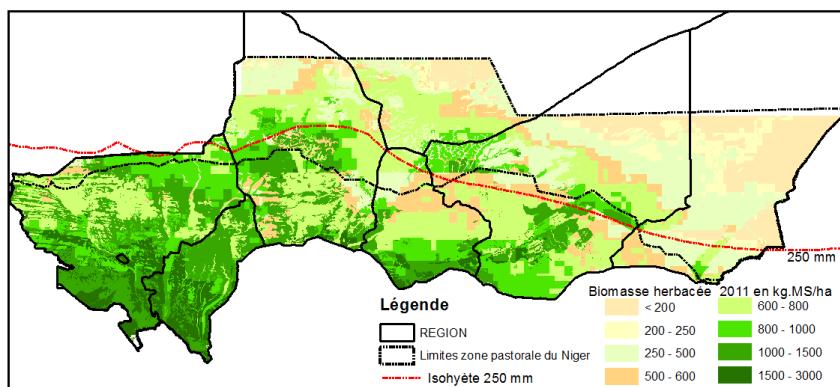


Figure 2.1 : Exemple de résultat issu du modèle BIOMASAH

### **2.2.3.2. Extraction des données simulées par utilisation de la couche vectorielle des sites du MEIA**

Les sites du MEIA sont géo-référencés, la table attributaire de la couche vectorielle de ces sites contient : les noms des sites, les années de

---

<sup>2</sup> <http://earlywarning.usgs.gov/fews/africa/>

mesures, les rendements fourragers, les types de sol, les écorégions et les zones bioclimatiques.

Pour extraire les données issues du modèle BIOMASAH devant servir à la validation, une zone tampon 3 km de côté a été réalisée à l'aide de cette couche vectorielle, sur chacun des sites (Figure 2.2. L'outil de statistique zonale (zonal statistic) du logiciel QGIS est utilisé pour extraire les moyennes de la biomasse aérienne herbacée correspondant à chacun des sites et suivant les années. Ainsi, on obtient dans la table des attributs de la couche vectorielle initiale, et en plus des champs initiaux, un nouveau champ contenant les mesures de phytomasse simulée suivant les années (2001 à 2011).

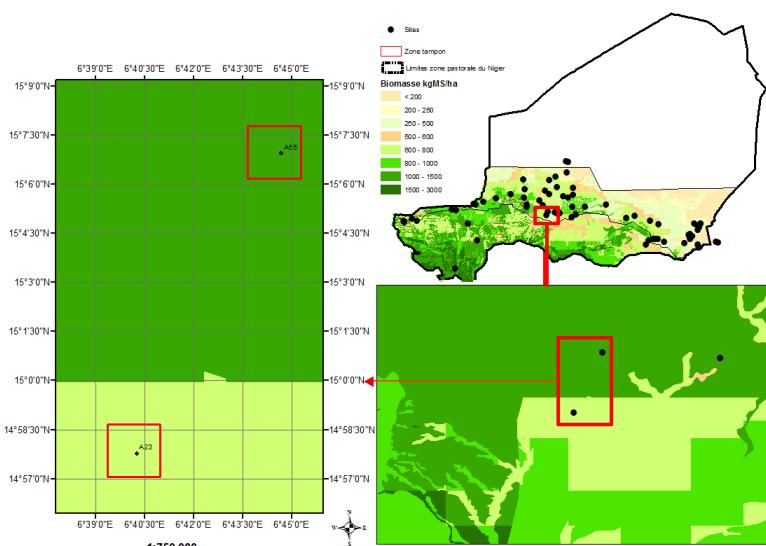


Figure 2.2 : Zoom sur les sites A65 et A 23 montrant les surfaces d'extraction des valeurs moyennes de phytomasse herbacée simulée par le modèle

#### 2.2.4. Analyse statistique

La comparaison des deux séries a été réalisée à l'aide de tests paramétriques et non paramétriques. Il s'agit notamment : du test de t apparié qui nécessite des conditions de validité (des observations appariées, l'indépendance des observations, l'échantillonnage aléatoire,

la distribution normale pour les différences, l'homogénéité des variances) ; du test de Wilcoxon et le test des signes qui ne nécessitent pas d'hypothèse préalable sur les formes de distribution (Dagnelie, 2013, Johnson et Bhattacharyya, 2010).

- **Test de t**

Le test de t a été choisi, car les deux séries de biomasses végétales à comparer sont celles des sites. Pour chaque site, il y a la variable issue du modèle et celle issue des mesures terrain. La statistique t dans le cas apparié est calculée suivant la formule suivante :  $t = \frac{M_d}{SE_d}$  où  $M_d$  est la différence entre les deux moyennes,  $SE_d$  l'erreur standard de la différence des deux moyennes. Avant ces analyses, les distributions des variables ont été étudiées, de même que l'égalité des variances qui est une des conditions d'applicabilité des tests de comparaison de moyennes.

- **Test des signes**

C'est un test qui s'applique dans le cas d'échantillons appariés. Le test des signes consiste à remplacer les observations plus grandes que  $M_0$  par un signe + et celles qui lui sont inférieures par un signe -. Si l'hypothèse nulle ( $H_0$ ) est vraie alors le nombre de signes +, soit  $n+$ , doit être 'proche' du nombre de signes -, le nombre de signes +, qui va constituer la statistique de test, possédant une distribution binomiale de paramètres  $n$  et  $1/2$  (ceci est également vrai pour  $n-$ , le nombre de signes -). Si l'alternative est unidirectionnelle de la forme  $H_1$  : la médiane est supérieure à  $M_0$  (Gilbert, 2004)

$$H_0: \Pr(x \geq 0) = \frac{1}{2} \text{ contre } H_1: \Pr(x \geq 0) \neq \frac{1}{2}$$

- **Test de Wilcoxon**

Selon cette méthode, on classe les observations par paires. Ce qui permet d'obtenir un compte du signe des différences par paires (comme dans le test des Signes) et les rangs de ces différences.  $Y (+)$  désigne la somme des rangs des différences positives ;  $Y (-)$  désigne la somme des rangs des différences négatives. Le principe est :  $Y (+) + Y (-) = n(n+1)$

/2 où n est le nombre de paires. En moyenne, si les deux échantillons proviennent d'une même population, Y (+) et Y (-) valent tous deux la moitié de cette valeur soit : n (n+1)/4 (Rousson, 2013; Good et al., 2012; Weiers et Heinz, 2011; Lejeune, 2010)

- **Les corrélations**

Les relations entre les données du modèle et celles du terrain ont été analysées avec différents types de coefficient de corrélation : r de Pearson,  $\rho$  de Spearman,  $\tau$  de Kendall, D de Hoeffding (Brase et Brase, 2012; Weiers et al., 2011; Rakotomalala, 2010). L'évolution interannuelle des coefficients de corrélation a été étudiée. La corrélation de Pearson est donnée par le coefficient de corrélation r de Pearson. Elle exprime la liaison qui existe entre les variables  $r_p = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$  où  $\sigma_{xy}$  désigne la covariance entre les variables x et y, et  $\sigma_x$   $\sigma_y$  leur écart type. La Corrélation de Spearman calcule un coefficient de corrélation entre les rangs des valeurs des deux variables. Cette corrélation est utilisée lorsque les distributions des variables sont asymétriques (skewness en anglais). L'interprétation est identique à celle de la corrélation de Pearson. Le coefficient de corrélation de Spearman est défini par :

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \text{ d'où } di = RXi - RYi \text{ R étant le rang}$$

Pour ce qui est de la corrélation de Kendall, le coefficient de corrélation des rangs de Kendall (Kendall  $\tau$ ) est une mesure de corrélation non paramétrique. Il sert à déterminer la relation qui existe entre deux séries de données. Il se calcule par la formule :

$$\tau = 1 - \frac{4Q}{n(n^2 - 1)} \text{ où}$$

$Q$  est le nombre d'inversions nécessaires parmi les valeurs de  $Y$  pour obtenir le même ordre (croissant) que celui des valeurs de  $X$  ; n est le nombre de paires d'observations.

La corrélation D de Hoeffding représente la mesure de la relation de

dépendance entre deux variables utilisant une formule basée sur les rangs (Wilding et Mudholkar, 2008).

## 2.3. Résultats et discussion

### 2.3.1. Résultats

Il y a globalement et dans les classes supérieure et inférieure au seuil pluviométrique (250 mm), une différence très significative au seuil de 1 pour 10 000 ( $P < 0,0001$ ) entre la productivité observée exprimée en Kg de MS.ha<sup>-1</sup> et la productivité potentielle issues du modèle biomasse (Tableaux 2.1, 2.2 et 2.3). Les comparaisons année par année font ressortir que l'hypothèse nulle de différences entre les valeurs mesurées et prédictes ne peut être rejetée avec une probabilité de 0,05 pour les années 2007 et 2010 (Tableau 2.4). Ceci permet d'affirmer que le modèle surestime globalement la production de biomasse. L'étude des corrélations paramétrique et non paramétrique de l'ensemble des données montre également des relations hautement significatives (Tableau 2.5). On note, suivant les années, des résultats intéressants car, même si, les coefficients de corrélation  $r$  ne dépassent guère 0,19, la relation reste significative ( $p < 0,001$ ). Ce qui permet de rejeter l'hypothèse nulle qui consistait à dire qu'il n'y a pas de relation entre les données mesurées et les données du modèle (Tableau 2.5). Les corrélations annuelles montrent que les relations varient suivant les années pour toutes les méthodes (Figure 2.2).

Tableau 2.1 : Comparaison des moyennes globales suivant les tests de t, Wilcoxon et des signes

	Nombre d'observations	Moyenne	Écart-type	Test de T	Test de Wilcoxon	Test des signes
Production simulée	319	983,17	348,36	t=12,38	s=16394,5	M=96
Masse mesurée	319	591,17	521,19	NDDL=305		
				p<, 0001	p <, 0001	p<, 0001

Tableau 2.2 : Comparaison par seuil pluviométrique ( $> 250$  mm) des moyennes globales suivant les 3 tests

	Nombre d'observations	Moyenne	Écart type	Test de T	Test de Wilcoxon	Test des signes
Production simulée	261	1042,93	332,53	t=10,93	s=11435,5	M=76,5
Masse mesurée	261	641,81	541,46	NDDL=260		
				p<, 0001	p <, 0001	p<, 0001

Tableau 2.3 : Comparaison par seuil pluviométrique ( $< 250$  mm) des moyennes globales suivant les 3 tests

	Nombre observations	Moyenne	Écart type	Test de T	Test de Wilcoxon	Test des signes
Production simulée	55	666,63	227,16	t=9,66	s=510,5	M=19,5
Masse mesurée	55	297,48	216,95	NDDL=44		
				p<, 0001	p <, 0001	p<, 0001

Tableau 2.4 : Comparaison des moyennes par an et suivant les quatre types de test

Année	Nb d'obs	Biomasse modèle		Biomasse observée		P		
		Moy	ET	Moy	ET	Test de T	Test de Wilcoxon	Test des signes
2001	24	1136,38	63,00	576,50	578,55	0,0003	0,0001	0,0001
2002	24	1097,08	448,15	446,89	298,29	<, 0001	<, 0001	<, 0001
2003	10	1327,65	383,09	833,89	443,53	<, 0433	<, 0273	0,10
2004	25	808,48	258,82	304,61	211,08	<, 0003	<, 0003	<, 0029
2005	34	1051,31	279,29	643,58	490,98	0,0003	0,0003	0,0029
2006	32	932,17	289,62	684,92	598,89	0,0355	0,0243	0,0037
2007	32	1035,86	319,15	664,22	526,92	0,0708	0,1417	0,3449
2008	40	867,66	302,25	412,97	429,03	0,0001	0,0001	0,0001
2009	20	792,18	366,31	382,09	274,70	<, 0001	<, 0001	<, 0001
2010	35	1013,76	297,52	921,136	779,56	0,80	0,58	0,17
2011	40	963,62	293,97	481,44	578,55	<, 0001	<, 0001	<, 0001

Nb : nombre ; obs : observation ; moy : moyenne ; ET : Écart-Type

Tableau 2.5 : Corrélations paramétrique et non paramétrique sur l'ensemble des données

Corrélation	Coefficient	Probabilité
r de Pearson	0,19	<,0001*
$\rho$ de Spearman	0,22	<, 0001*
$\tau$ de Kendall	0,13	<, 0001*
D de Hoeffding	0,01	<, 0001*

\*significatif au seuil de 1 pour 10000

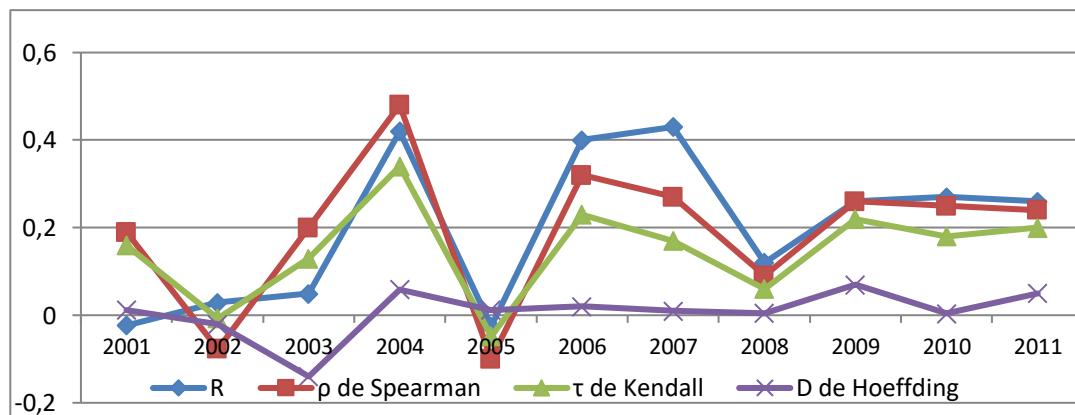


Figure 2.2 : Variation annuelle des coefficients de corrélation suivant les différents tests

### **2.3.2. Discussion partielle**

Le modèle BIOMASAH considère que l'eau est le principal facteur limitant dans les zones en dessous de l'isohyète 250 mm. Quatre-vingt pourcents des sites d'observations sont localisés dans cette zone. La comparaison des moyennes indique une différence d'un facteur 2 entre les valeurs du modèle et celles mesurées. Dans la zone au-dessus de l'isohyète 250 mm, le modèle considère que l'azote et le phosphore constituent les facteurs limitants à la production de biomasse. À ce niveau aussi, la comparaison des moyennes donne une différence très significative avec les mêmes tendances que dans l'isohyète inférieure. Les différences observées entre les observations du modèle BIOMASAH et les observations du terrain s'expliquent par plusieurs raisons. En effet, le Niger dispose d'un effectif de cheptel dépassant les 31 039 041 millions de têtes (RGAC,2008), localisé essentiellement dans la zone pastorale (Rhissa, 2010). Selon l'IPCC, à l'échelle globale, plus de 20 % des pâturages sont dégradés par les troupeaux, par le surpâturage (un excès de pression sur le pâturage par les animaux c'est-à-dire une surexploitation de la ressource destinée à leur alimentation entraînant sa dégradation quantitative et qualitative), le compactage, et l'érosion. Ce pourcentage serait plus élevé dans les zones arides (Nori et Davies, 2007). Malheureusement les affirmations ne sont basées sur aucune analyse scientifique. Cette pression peut contribuer à baisser significativement la productivité observée, car le prélèvement du bétail sur ces sites (ouverts toute l'année), n'est malheureusement pas pris en compte dans le modèle. Le Tableau 1.1 indique que les mesures de la phytomasse sur le terrain interviennent, le plus souvent, à plus d'un mois après la date du maximum de végétation considérée comme période optimale de mesure. Ces mesures tardives de la phytomasse herbacée est un facteur qui peut contribuer à minorer la production. La distribution intra saisonnière des séquences sèches peut aussi expliquer en partie cette différence. En effet, ces dernières années, on assiste de plus en plus à une forte variabilité temporelle des pluies, notamment

une fréquence élevée des phénomènes extrêmes notamment les sécheresses et inondations pouvant avoir des conséquences sur la productivité (Steinfeld et al., 2009). Par exemple, on peut enregistrer un cumul de pluie élevé avec un nombre de jours de pluie très limité, ces précipitations ne seront pas très favorables pour la bonne croissance de la végétation, car le ruissellement sera fort (avec comme conséquences des pertes énormes en terre en quantité et en qualité). Les sécheresses provoquent un ralentissement de la croissance des plantes ou même leur perte réduisant la capacité des pâturages d'exprimer au mieux leur potentiel. Une séquence sèche qui survient pendant un stade critique de la croissance peut avoir des effets catastrophiques sur le rendement, même si les précipitations sont globalement abondantes (Aguiar, 2009). Il est important de rappeler que, le modèle n'a pas été conçu pour gérer les séquences sèches. Le modèle peut être limité par le fait que le Coefficient de Ruissellement (CR) peut être négatif lorsque l'unité paysagère bénéficie du ruissellement. En plus, le ruissellement est largement fonction du volume et de l'intensité des événements pluvieux. Un coefficient annuel suppose donc une distribution des pluies par volume et intensité. Il est noté aussi l'inexistence d'un bilan hydrique entre les unités paysagères associées dans un bassin versant. Les forts cumuls de pluies auront tendance à surestimer la production potentielle même si les ruissellements seront aussi sous-estimés lorsque calculé par an. Au vu des résultats de cette validation qui montrent des différences importantes entre les données du terrain et les données du modèle, il est impérieux de mener de recherches pour trouver un moyen plus intéressant d'estimation de la production fourragère. Car la disponibilité des données du terrain et des séries d'images satellitaires constitue une belle opportunité pour proposer un modèle combinant les données du terrain et les images satellitaires.

## **2.4. Conclusion partielle**

La comparaison des moyennes réalisée entre les phytomasses aériennes herbacées mesurées *in situ* et celles simulées sur la période s'étalant de 2001 à 2011, met en évidence une différence significative entre ces deux types approches. Le modèle BIOMASAH surestime la masse herbacée, comme le montre les tests paramétriques (test de t) et non paramétriques (de Wilcoxon et des signes). Les résultats de ce modèle correspondent en moyenne au double des données mesurées, avec un R<sup>2</sup> ne dépassant guère 0,15. En conclusion, ces résultats indiquent que les sorties du modèle BIOMASAH ne permettent pas d'expliquer la variabilité contenue dans les données du terrain. N'est-il pas possible d'améliorer le modèle BIOMASAH, en prenant en compte la distribution journalière ou décadaire des pluies pour le bilan hydrique du sol ? La disponibilité de séries assez longues de données mesurées *in situ*, d'images d'indice de végétation et d'autres données agrométéorologiques ouvre toutefois de nouvelles perspectives pour produire un modèle plus fiable.

### **III. Analyse de la performance du modèle d'estimation de la biomasse élaboré par le Ministère de l'Élevage et des Industries Animales (MEIA) du Niger**

Chapitre tiré de l'article :

GARBA I., B. Djaby, I. Salifou, A. Boureima, I. Touré, B. Tychon (2015). Évaluation des ressources pastorale au sahel Nigérien à l'aide des données NDVI issues de SPOT-VEGETATION et MODIS. *Photo interprétation European Journal of Applied Remote Sensing*, N°2015/1, 13-26.

#### **3.1. Introduction**

L'élevage joue un rôle économique très important au Niger, il est pratiqué par 80 % des ménages et constitue la 2<sup>e</sup> recette d'exportation avec 31 039 041 millions de têtes toutes espèces comprises bovins, ovins, caprins, camelins, azins (MEIA, 2012). Le pastoralisme est pratiqué de façon traditionnelle et est basé essentiellement sur l'exploitation des fourrages naturels soumis à de fortes variabilités inter et intra saisonnières. Ceci nécessite pour les acteurs (Éleveurs, Techniciens, Décideurs politique, ONG...) de disposer d'outils appropriés de suivi et de gestion de pâturage permettant de : (i) anticiper les crises pastorales ; (ii) favoriser une exploitation judicieuse du disponible fourrager ; (iii) orienter les politiques nationales et communautaires (Politiques foncières, commerciales des produits de l'élevage et des intrants, politiques douanières, banquières ; la politique de développement des collectivités locales : infrastructures sociales adaptées pour la santé, l'éducation, l'eau et l'électricité...). L'utilisation de l'imagerie satellitaire combinée aux mesures *in situ* constitue un moyen important d'évaluation quantitative et qualitative des pâturages (Swets et al., 1999; Wylie et al., 1991; Diallo et al., 1991). Au Sahel, le Centre Régional AGRHYMET(CRA), le Ministère de

l’Élevage et des Industries Animales (MEIA) du Niger et le Centre de Suivi Écologique (CSE) du Sénégal réalisent annuellement des estimations de biomasse pour aider les pouvoirs publics à prendre des décisions sur la sécurité alimentaire et la gestion des ressources naturelles. Le modèle du MEIA s’appuie sur l’utilisation de données mesurées au sol combinées à l’intégrale des images d’indice de végétation à différence normalisée (NDVI pour son acronyme en anglais) dont, plusieurs auteurs (Huete et al., 2002, Justice et al., 1986, Hiernaux et Justice, 1986) ont montré les performances mais aussi les limites dans le suivi et la caractérisation de la végétation à l’échelle globale.

Le NDVI fait partie des nombreux indices dérivés des mesures de réflectances spectrales dans différentes bandes du spectre électromagnétique. Grâce au pigment chlorophyllien des feuilles vertes la végétation absorbe fortement dans le rouge (0,6 à 0,7 µm) pour la photosynthèse et réfléchit dans le proche infrarouge (0,7 à 1,1 µm) grâce à la structure des cellules foliaires. Dès le lancement des premiers satellites d'observation de la terre (NOAA AVHRR), avec l'acquisition d'images dans le visible et le proche infrarouge, une exploitation a été faite pour déterminer la distribution spatiale de la végétation. Dans le cas du satellite NOAA, les deux canaux utilisés pour le calcul du NDVI sont le rouge (0,58 à 0,68 µm) et le proche infrarouge (0,73 à 1,10 µm)(Rouse et al., 1973).  $NDVI = \frac{PIR-R}{PIR+R}$  où :

PIR = Réflectance dans le proche Infra-Rouge et R = Réflectance dans le Rouge. Dans les capteurs VEGETATION de SPOT-4 et 5, il s'agit des longueurs d'onde (0,61 - 0,68 µm) rouge et (0,78 - 0,89) µm pour le proche infrarouge. Pour ce qui est de MODIS, la bande du canal rouge se trouvent entre 0,62 et 0,67 µm et celle du proche infrarouge entre 0,841 -0,876 µm.

Les acquisitions journalières de NDVI sont en général traitées à travers l'algorithme du « maximum values compositing » pour réduire l'impact des interférences atmosphériques, et d'angle de visée et produire des

images décennales (Holben et Fraser, 1984). Le NDVI est devenu très vite un outil précieux de suivi de la végétation. En effet, à l'aide des indices de végétation mesurés par satellite, on peut surveiller, pendant la saison de croissance, les ressources végétales importantes comme les cultures, les pâturages et les forêts à une échelle spatiale appropriée. Dans le cadre du suivi et de la quantification de la phytomasse, plusieurs études ont été réalisées dans le monde pour le suivi et la quantification des productions primaires. Les séries d'images d'indice de végétation dérivées de MODIS, SPOT VEGETATION, NOAA etc. ont été largement utilisées à travers le monde pour le suivi de la phénologie des cultures (Hmimina et al., 2013, Soudani et al., 2012, Butt et al., 2011, Beck et al., 2007, Wittemyer et al., 2007, Wagenseil et Samimi, 2006, Lüdeke et al., 1996) et l'estimation des rendements agricoles qui a été faite principalement en combinant les valeurs de NDVI aux rendements mesurés *in situ* (Mkhabela et al., 2011, Fontana et al., 2007, Kastens et al., 2005, Bozzini et Maselli, 2002, Potdar et al., 1999, Rasmussen, 1998). Des estimations de rendements des cultures ont été également réalisées avec succès au Niger et au Burkina Faso à l'aide de séries de NDVI (NOAA AVHRR) croisées aux rendements mesurés *in situ* (Maselli et al., 1993, Grotel, 1993). Les indices de végétation provenant de multiples capteurs ont été aussi utilisés pour la prévision des rendements agricoles (Mkhabela et al., 2011; Kastens et al., 2005; Bozzini et al., 2002). Au Sahel, pour le suivi et l'estimation de la production primaire à l'aides des images satellitaires, des études ont été réalisées dès les années 1980 (Gaston et al., 1987; Tucker et al., 1986; Tucker et al., 1985; Tucker et al., 1981). Les études qui ont été à la base du modèle du MEIA sont celles réalisées avec succès au Niger et au Sénégal ayant établi une corrélation significative entre l'intégrale de la série des images décennales de NDVI de la saison des pluies et la phytomasse mesurée au sol (Wylie et al., 1991; Diallo et al., 1991). Ce large spectre d'utilisation des images NDVI permet d'affirmer que cet indice est incontestablement le plus utilisé pour estimer les

productions des cultures et des pâturages.

Au regard de l'importance de l'estimation de la biomasse fourragère pour la planification des interventions d'un pays à vocation pastorale comme le Niger, de la quantité importante de données mesurées disponibles dans la base de données du MEIA, de la disponibilité des séries d'images NDVI issues de SPOT VEGTATION et de MODIS, une analyse critique de la méthode du MEIA s'avère nécessaire pour en déterminer les forces et les faiblesses en vue de proposer des pistes d'amélioration.

Ainsi, l'objectif général de cette étude est de contribuer à l'amélioration de la qualité des informations utilisées pour la planification des actions dans le domaine de l'élevage. Il s'agit plus spécifiquement de :

- contribuer à l'amélioration de la méthode d'estimation de la production fourragère au Niger ;
- comparer d'une part les performances de l'intégrale et du maximum NDVI, d'autre part celles de données SPOT VEGETATION et eMODIS pour l'estimation de production fourragère au Niger.

Trois hypothèses sont formulées :

- il est possible d'utiliser le NDVI de eMODIS à la place du NDVI de SPOT VEGETATION ;
- l'intégrale du NDVI et le Maximum du NDVI peuvent avoir des performances différentes suivant les années ;
- les différences entre zones bioclimatiques ne sont pas prises en compte en réalisant un modèle unique pour tout le pays.

### **3.2. Matériel et méthodes**

Les données mesurées au sol par le MEIA dont la méthode de collecte a été décrite au paragraphe 1.6.1 sont associées aux images NDVI de SPOT VEGETATION et de MODIS pour extraire les métriques (intégrale et maximum NDVI) par site. On a ensuite procédé à des ajustements pour établir des équations linéaires du type  $y = ax + b$  où  $y$  représente les données observées ,  $x$  l'intégrale ou le maximum NDVI

de la saison de croissance de la végétation et des tests non paramétriques pour évaluer les niveaux de signification de la différence entre les paramètres des modèles (Figure 3.1).

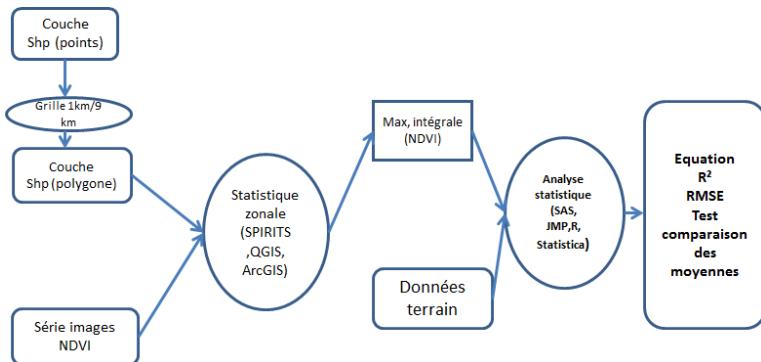


Figure 3.1 : Modèle schématique d'extraction des métriques du NDVI et du traitement statistique

### 3.2.1. Les données de biomasse fourragère du MEIA

Rappelons que les données de biomasse observée utilisées proviennent de la division pastorale du MEIA. Elles sont mesurées chaque année pour établir le bilan fourrager sur lequel se fonde la planification des interventions dans le domaine pastoral au Niger. Ces données sont collectées sur le terrain à travers un dispositif de collecte et de traitement de la production fourragère herbacée conformément à la méthode décrite plus haut. Une base de données est créée et comprend les numéros des sites, les informations attributaires telles que la production annuelle, la position géographique. Pour apprécier la représentativité des sites de mesure de la biomasse fourragère du Niger, nous avons examiné la production et l'évolution temporelle du nombre et de sites de 2001 à 2012.

### **3.2.2. Les images satellitaires utilisées**

Le travail a été réalisé à partir de deux sources satellitaires d'images d'indice de végétation : les images NDVI SPOT VEGETATION provenant du VITO et celles d'eMODIS mises à la disposition des utilisateurs par FEWSNET.

#### **3.2.2.1. Images NDVI du satellite SPOT VEGETATION**

Les images NDVI utilisées dans le cadre de cette étude proviennent des capteurs SPOT VEGETATION (SPOT4 lancé en 1998 et SPOT5 mis en orbite en 2002). Elles ont été traitées et mises à la disposition des utilisateurs des pays en développement sur le web grâce au projet DevCoCast. L'amélioration de la qualité des images NDVI décadaires a été faite suivant la méthode Maximum Value Composite (MVC). Cette technique consiste à faire une synthèse décadaire S10 des images journalières en attribuant à chaque pixel la valeur d'indice maximale enregistrée par le pixel au cours de cette décennie. Cette technique permet de minimiser les contaminations atmosphériques et les effets off-nadir(Holben et al., 1986). Les images SPOT VEGETATION sont téléchargées du site en code numérique, c'est-à-dire, codées entre 0 et 255 Digital Number (DN). Pour convertir le DN en Valeur NDVI de type décimal, on applique la formule suivante :  $NDVI = 0,004 * DN - 0,1$ .

#### **3.2.2.2. Les images Expedited MODIS (eMODIS)**

Les images EMODIS proviennent du satellite Américain (TERRA, lancé en 1999) Moderate Resolution Imaging Spectroradiometer (MODIS). Elles sont développées par l'United States Geological Survey (USGS) et le centre Earth Resources Observation and Science (EROS) pour le suivi de la végétation verte en temps quasi réel. La série d'images eMODIS est disponible depuis 2001. Ce sont des synthèses de 5 et 10 jours corrigées des contaminations atmosphériques. Pour

améliorer la qualité, la série d'images NDVI subit un lissage temporel par la méthode développée par Swets et al. (1999). Les images EMODIS sont téléchargées du site : (<http://earlywarning.usgs.gov/fews/africa>) en code numérique, c'est-à-dire, codées entre 0 et 255 Digital Number (DN). Pour convertir le DN en vraies valeurs NDVI, on applique la formule suivante :  $NDVI = 0,01 * DN - 1$ .

### **3.2.3. Extraction de l'intégrale et du maximum NDVI par site**

Les données de production de biomasse mesurée *in situ* (68 sites) ont été récupérées auprès du Ministère de l'Elevage et des Industries Animales (MEIA) en format Excel, elles contiennent essentiellement quatre colonnes : le Numéro du Site, la Latitude, la Longitude et la Production en kg.ha<sup>-1</sup>. Une couche géographique ponctuelle de type shapefile a été créée avec les coordonnées des 68 sites. Puis on a générée une grille de polygones de 1 km de côté qui se superpose parfaitement aux pixels des images NDVI de SPOT VEGETATION à l'aide de l'outil fishnet du logiciel ArcGIS ; et superposé la couche des sites sur celle des grilles pour extraire les neufs polygones contigus de 1km<sup>2</sup> correspondant à la surface de chaque site ; on a procédé à la fusion des 9 carrés de 1km de côté de chaque site pour en faire un seul de 3 km de côté ; la jointure spatiale est réalisée pour affecter les attributs de chaque site au polygone correspondant. La couche de ces polygones contenant les informations sur les mesures *in situ* c'est-à-dire les productivités observées des années 1998 et 2011 est ensuite superposée sur les images NDVI (prétraitées) de mai à octobre des années respectives pour en extraire les intégrales et les maximums des pixels correspondant à chaque site (**Figure 3.2**). La table issue de cette opération est ensuite utilisée pour faire les analyses statistiques entre la masse aérienne des herbacées mesurée *in situ* et le maximum ou l'intégrale de NDVI pour chaque année. La fonction overlay a été utilisée pour relever les types de sol, les zones agro-écologiques et les zones bioclimatiques correspondant à chaque site.

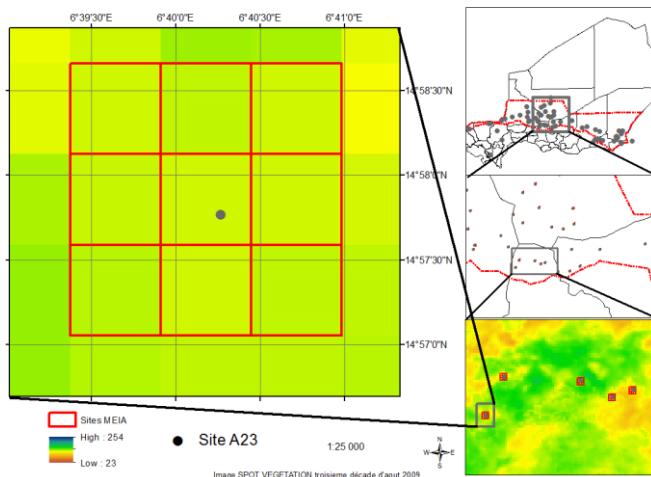


Figure 3.2 : Exemple du site A23  
matérialisé par la position géographique du centre

### 3.2.4. Analyse statistique des données

Les analyses réalisées ont consisté à étudier l'évolution interannuelle des coefficients de régression et leur évolution spatiale. Les coefficients issus des régressions sont le R<sup>2</sup>, le RMSE. Les images NDVI de SPOT VEGETATION et EMODIS ont été traitées avec les logiciels SPIRITS (*Software for the Processing and Interpretation of Remotely Sensed Image Time Series*) (<http://spirits.jrc.ec.europa.eu>) et ArcGIS pour le Système d'information géographique (SIG). Les régressions et les tests non paramétriques de corrélation (tests de wilcoxon et des signes) ont été faits avec le logiciel SAS-JMP10.

## **3.4. Résultats et discussion**

### **3.4.1. Résultats**

#### **3.4.1.1 Caractérisation de la productivité**

La moyenne nationale du rendement fourrager varie du simple au double voire même le triple suivant les années (Tableau 3.1). Cette moyenne suit le gradient pluviométrique en s'accroissant du Nord vers le Sud (Tableau 3.2).

Tableau 3.1 : rendement fourrager moyen suivant les années

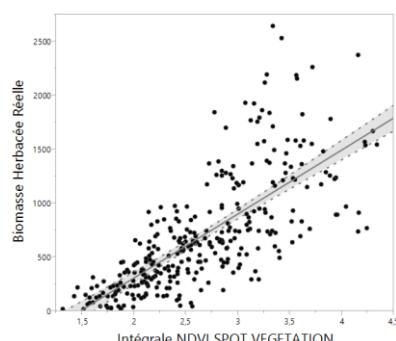
Année	Nombre de sites	Moyenne kg.ha <sup>-1</sup>	Écart Type
2001	21	604,31	349,14
2002	27	382,52	260,47
2003	12	809,18	389,09
2004	26	287,99	191,53
2005	31	713,42	432,80
2006	24	825,99	561,40
2007	23	1119,17	337,16
2008	36	457,30	425,71
2009	16	522,01	235,80
2010	35	1080,86	663,06
2011	40	470,74	272,06
2012	25	1299,49	606,20

Tableau 3.2 : rendement fourrager moyen suivant les zones bioclimatiques

Zone bioclimatique	Nombre d'observations	Moyenne kg.ha-1	Écart Type
saharienne	48	364,11	288,56
Nord sahélique	243	717,80	514,84
sahélienne	23	1203,47	547,04

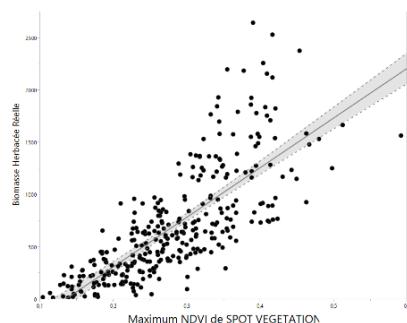
### 3.4.1.2. Analyse à l'échelle globale

Il ressort des résultats de la régression linéaire simple (Figure 3.3 et 3.4) entre la biomasse mesurée au sol (2001 à 2012) et le NDVI (Intégrale et Maximum) de SPOT VEGETATION et eMODIS que : la relation est hautement significative ( $P < 0,0001$ ), les coefficients de détermination  $R^2$  des maximums sont supérieurs à ceux des intégrales. La variance la masse herbacée en  $\text{kg.MS}.\text{ha}^{-1}$  dans la zone pastorale du Niger est alors expliquée à un peu plus de 50 % par le NDVI (maximum ou intégrale) de SPOT VEGETATION et d'EMODIS. La valeur moyenne du RMSE est de  $365 \text{ kg.MS}.\text{ha}^{-1}$  soit une erreur de 52 % par rapport à la moyenne (tableau 3. 3).



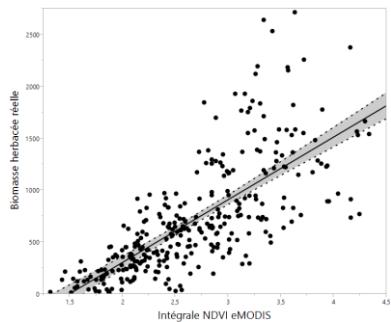
$$\text{Biomasse herbacée observée} = -875,22 + 593,33 * \text{intégrale NDVI SPOT VEGETATION}$$

Figure 3.3 : Ajustement de la biomasse herbacée observée avec l'intégrale NDVI SPOT VEGETATION à l'échelle globale



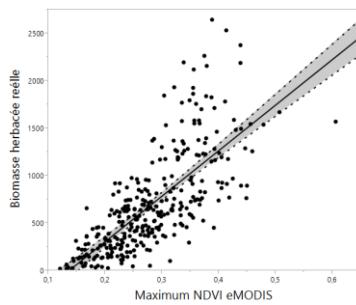
$$\text{Biomasse herbacée observée} = -636,34 + 4749,78 * \text{maximum NDVI SPOT VEGETATION}$$

Figure 3.4 : Ajustement de la biomasse herbacée observée avec le maximum NDVI SPOT VEGETATION à l'échelle globale



$$\text{Biomasse herbacée observée} = -900,36 + 604,28 \times \text{intégrale NDVI eMODIS}$$

Figure 3.5 : Ajustement la biomasse herbacée observée avec l'intégrale NDVI eMODIS à l'échelle globale



$$\text{Biomasse herbacée observée} = -651,57 + 4796,79 \times \text{maximum NDVI_eMODIS}$$

Figure 3.6: Ajustement de biomasse herbacée observée avec le maximum NDVI eMODIS à l'échelle globale

Tableau 3.3 : Comparaison des R<sup>2</sup> des résultats avec SPOT VEGETATION et eMODIS à l'échelle globale

NDVI	R <sup>2</sup>	RMSE
eMODIS intégrale	0,51	374
eMODIS maximum	0,53	368
SPOT VEGETATION intégrale	0,53	367
SPOT VEGETATION maximum	0,57	354

### 3.4.1.3. Analyse par zone bioclimatique

L'examen des résultats de la régression linéaire simple entre la biomasse mesurée au sol (2001 à 2012) et le NDVI (Intégrale et Maximum) de SPOT VEGETATION et EMODIS suivant les zones bioclimatiques (**Figure 3.5 à 3.10**) montre que :

- La relation est hautement significative ( $P < 0,0001$ ),
- Les coefficients de détermination R<sup>2</sup> proches de 0,5 pour les zones Nord sahélienne et saharienne, et par contre le R<sup>2</sup> est inférieur à 0,2 pour la partie sahélienne ;

- La variance de la production de biomasse dans la zone pastorale du Niger est donc expliquée à un peu plus de 50 % par le NDVI (maximum ou intégrale) qu'il soit de SPOT VEGETATION ou d'eMODIS dans les zones Nord sahélienne et saharienne. En revanche elle est faiblement expliquée en zone sahélienne.
- Les valeurs des RMSE sont en moyenne de  $360 \text{ kg MS.ha}^{-1}$  en zone nord sahélienne,  $207 \text{ kg.MS.ha}^{-1}$  en zone saharienne et  $583 \text{ kg MS.ha}^{-1}$  en zone sahélienne, soit des erreurs moyennes respectives de 48 à 60% (tableaux 3.4 ; 3.5 ; 3.6).

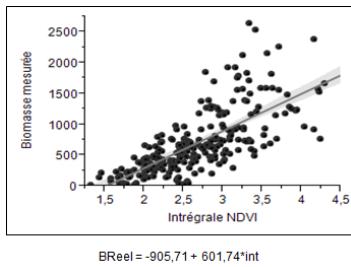


Figure 3.7 : Ajustement avec l'intégrale NDVI SPOT VEGETATION en zone nord-sahélienne

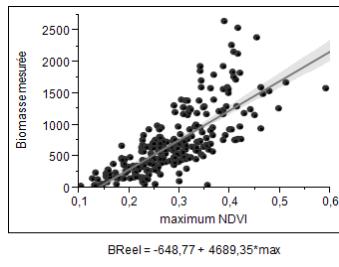


Figure 3.8 : Ajustement avec le maximum NDVI SPOT VEGETATION en zone nord-sahélienne

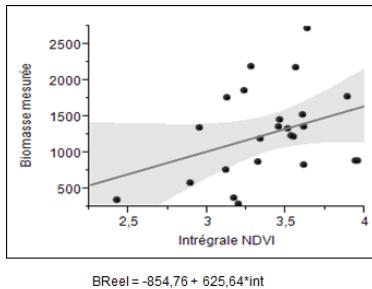


Figure 3.9 : Ajustement avec l'intégrale NDVI SPOT VEGETATION en zone sahélienne

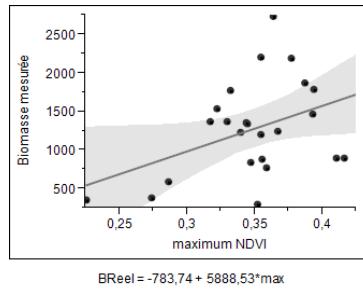


Figure 3.10 : Ajustement avec le maximum NDVI SPOT VEGETATION en zone sahélienne

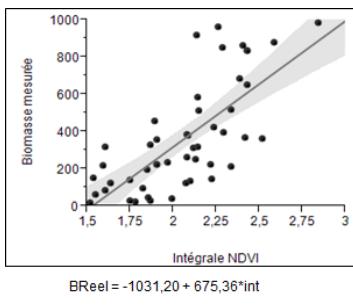


Figure 3.11 : Ajustement avec l'intégrale NDVI SPOT  
VEGETATION en zone saharienne

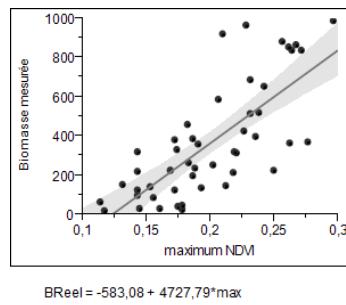


Figure 3.12 : Ajustement avec le maximum NDVI SPOT  
VEGETATION en zone saharienne

Tableau 3.4 : Variation du coefficient de détermination ( $R^2$ ) suivant les zones bioclimatiques

Zone	Nord	Sahélienne	Saharienne
	Sahélienne		
eMODIS intégrale	0,47	0,18	0,45
eMODIS maximum	0,51	0,11	0,43
SPOTVGT intégrale	0,51	0,13	0,53
SPOTVGT maximum	0,56	0,17	0,56

Tableau 3.5 : Variation du RMSE suivant les zones bioclimatiques

Zone	Nord	Sahélienne	Saharienne
	Sahélienne		
eMODIS intégrale	375	573	217
eMODIS maximum	361	594	220
SPOTVGT intégrale	362	591	200
SPOTVGT maximum	344	576	193

Tableau 3.6 : Variation relative du RMSE suivant les zones bioclimatiques

Zone	Nord	Sahélienne	Sahélienne	Saharienne
eMODIS intégrale	52%		48%	60%
eMODIS maximum	50%		49%	60%
SPOTVGT intégrale	50%		49%	55%
SPOTVGT maximum	48%		48%	53%

#### **3.4.1.4. Analyse suivant les années**

Les résultats de la régression linéaire simple entre la biomasse herbacée mesurée au sol au cours de la période allant de 2001 à 2012 et soit l'intégrale, soit le maximum du NDVI de SPOT VEGETATION et d'eMODIS sont illustrés par les tableaux 3.7 à 3.8. La régression montre partout que la relation est hautement significative et pour toutes les années ( $P < 0,0001$ ). Les coefficients de détermination  $R^2$  varient entre 38% et 72 %. Les variations observées sont identiques quel que soit le capteur. La production de biomasse herbacée dans la zone pastorale du Niger est donc expliquée pour certaines années à un peu plus de 70 % par le NDVI (maximum ou intégrale) qu'il soit de SPOT VEGETATION ou d'eMODIS (Tableau 3.9). Les valeurs des RMSE varient de 120 à 455 kg MS. ha<sup>-1</sup>(tableau3.10). Les résultats des tests (Wilcoxon et Signes) montrent qu'il n'y a pas de différence significative entre les RMSE (Tableau 3.11 et 3.14). L'annexe 3 indique les droites d'ajustement des mesures de phytomasse herbacée aérienne en fonction du maximum et de l'intégrale NDVI issus de SPOT-VEGETATION et de MODIS.

Tableau 3.7 : Relation entre l'intégrale de NDVI de SPOT VEGETATION et la biomasse herbacée suivant les années.

Année	Modèle	Nr	R <sup>2</sup>	RMSE
		site	(kg MS/ha)	
2001	BReel = -1253,92 + 758,15*int	21	0,56	384
2002	BReel = -502,31 + 387,64*int	27	0,74	137
2003	BReel = -772,57 + 583,17*int	12	0,62	253
2004	BReel = -571,24 + 435,89259*int	26	0,58	126
2005	BReel = -762,03 + 561,78013*int	31	0,72	233
2006	BReel = -1106,66 + 710,18*int	24	0,63	351
2007	BReel = -343,09 + 446,17*int	23	0,54	233
2008	BReel = -954,22 + 528,74*int	36	0,52	299
2009	BReel = -566,36 + 439,49*int	16	0,59	162
2010	BReel = -1375,12 + 859,19*int	35	0,54	455
2011	BReel = -413,05 + 316,12*int	40	0,65	164
2012	BReel = -1000,97 + 786,70*int	25	0,56	410

BReel: Biomasse totale; int : intégrale NDVI

Tableau 3.8: relation entre le maximum de NDVI de SPOT VEGETATION et la biomasse herbacée suivant les années

Année	Modèle	Nr	R <sup>2</sup>	RMSE
		site	(kg MS/ha)	
2001	BReel = -651,95 + 4835,0476*max	21	0,38	455
2002	BReel = -213,69 + 2405,52*max	27	0,65	160
2003	BReel = -1187,53 + 7636,05*max	12	0,76	198
2004	BReel = -224,94 + 2357,85*max	26	0,6	124
2005	BReel = -502,56 + 4412,1453*max	31	0,63	269
2006	BReel = -663,31 + 4871,8838*max	24	0,59	368
2007	BReel = 16,04+ 2961,89*max	23	0,57	225
2008	BReel = -660,53 + 4278,37*max	36	0,54	288
2009	BReel = -335,25 + 3275,65*max	16	0,52	176
2010	BReel = -968,49 + 6703,91*max	35	0,59	433
2011	BReel = -262,61 + 2676,15*max	40	0,63	168
2012	BReel = -766,32 + 6197,71*max	25	0,6	391

BReel: Biomasse totale; max : maximum NDVI

Tableau 3.9 : Relation entre l'intégrale de NDVI d'eMODIS et la biomasse herbacée suivant les années.

Année	Modèle	Nombre de site	R <sup>2</sup>	RMSE
2001	BReel = -2274,409 + 563,05093*int	21	0,62	357
2002	BReel = -958,7376 + 271,08641*int	27	0,71	144
2003	BReel = -1274,074 + 392,15374*int	12	0,63	247
2004	BReel = -887,96 + 262,05976*int	26	0,54	131
2005	BReel = -1360,005 + 398,9208*int	31	0,73	229
2006	BReel = -1853,817 + 743,39651*int	24	0,62	354
2007	BReel = -685,6192 + 444,54888*int	23	0,56	229
2008	BReel = -1565,205 + 574,01701*int	36	0,53	296
2009	BReel = -1035,214 + 476,31009*int	16	0,57	164
2010	BReel = -2133,852 + 924,50059*int	35	0,52	464
2011	BReel = -769,8143 + 336,08794*int	40	0,64	163
2012	BReel = -1920,488 + 584,59571*int	25	0,59	394

BReel: Biomasse totale; int : intégrale NDVI

Tableau 3.10: relation entre le maximum de NDVI de eMODIS et la biomasse herbacée suivant les années

Année	Modèle	Nombre de site	R <sup>2</sup>	RMSE
2001	BReel = -937,3483 + 5745,7317*max	21	0,45	429
2002	BReel = -292,647 + 2666,006*max	27	0,71	145
2003	BReel = -770,2853 + 5720,8998*max	12	0,68	231
2004	BReel = -309,6009 + 2548,3944*max	26	0,62	120
2005	BReel = -478,6606 + 4132,3514*max	31	0,62	270
2006	BReel = -872,668 + 5573,4358*max	24	0,62	353
2007	BReel = 17,718957 + 3001,7481*max	23	0,65	204
2008	BReel = -663,4033 + 4337,4184*max	36	0,56	284
2009	BReel = -314,2879 + 3358,0384*max	16	0,49	178
2010	BReel = -1033,525 + 7201,3384*max	35	0,58	433
2011	BReel = -290,5028 + 2687,9264*max	40	0,62	169
2012	BReel = -971,6338 + 7380,0853*max	25	0,68	350

BReel: Biomasse totale; int : intégrale NDVI

Tableau 3.11 : Variation annuelle des R<sup>2</sup> issus de la régression entre la biomasse herbacée et l'intégrale ou le maximum NDVI de SPOT VEGETATION et eMODIS

Année	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
eMO int	0,62	0,71	0,63	0,54	0,73	0,62	0,56	0,53	0,57	0,52	0,64	0,59
eMO max	0,45	0,71	0,68	0,62	0,62	0,62	0,65	0,56	0,49	0,58	0,62	0,68
SPOT int	0,56	0,74	0,62	0,58	0,72	0,63	0,54	0,52	0,59	0,54	0,65	0,56
SPOTmax	0,38	0,65	0,76	0,60	0,63	0,59	0,57	0,54	0,52	0,59	0,63	0,60

eMO int : eMODIS Intégrale ; eMO max : eMODIS maximum ; SPOT int : SPOT VEGETATION intégrale ; SPOT max : SPOT VEGETATION maximum

Tableau 3.12 : Variation annuelle des RMSE issues de la régression entre la biomasse herbacée en l'intégrale et maximum NDVI de SPOT VEGETATION et eMODIS

Année	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
eMO int	357	144	247	131	229	354	229	296	164	464	163	394
eMO max	429	145	231	120	270	353	204	284	178	433	169	350
SPOT int	384	137	253	126	233	351	233	299	162	455	164	410
SPOT max	455	160	198	124	269	368	225	288	176	433	168	391

eMO int : eMODIS Intégrale ; eMO max : eMODIS maximum ; SPOT int : SPOT VEGETATION intégrale ; SPOT max : SPOT VEGETATION maximum

Tableau 3.13 : Test de Wilcoxon sur les RMSE issues de la régression entre la biomasse herbacée en l'intégrale et maximum NDVI de SPOT VEGETATION et eMODIS (basé sur les rangs)

	Spot int-spot	Emo max-spot	Emo max-spot	Emo int-spot	Emo int-spot	Emo int-emo
	max	max	int	max	int	max
Statistique de test S	-5,000	-17,000	-5,000	-7,000	-7,500	6,500
Prob > S	0,7334	0,1426	0,7178	0,6221	0,5830	0,6353
Prob>S	0,6333	0,9287	0,6411	0,6890	0,7085	0,3176
Prob <	0,3667	0,0713	0,3589	0,3110	0,2915	0,6824

eMO int : eMODIS Intégrale ; eMO max : eMODIS maximum ; SPOT int : SPOT VEGETATION intégrale ; SPOT max : SPOT VEGETATION maximum

Tableau 3.14 : Test des signes sur les RMSE issues de la régression entre la biomasse herbacée en l'intégrale et maximum NDVI de SPOT VEGETATION et eMODIS

	Spot int-spot	Emo max-spot	Emo max-spot	Emo int-spot	Emo int-spot	Emo int-emo
	max	max	int	max	int	max
Statistique de test M	0,000	-1,500	0,000	0,000	-1,000	1,000
Prob $\geq  M $	1,0000	0,5488	1,0000	1,0000	0,7744	0,7744
Prob $\geq M$	0,5002	0,8867	0,5002	0,5002	0,8062	0,3872
Prob $\leq M$	0,5000	0,2744	0,5000	0,5000	0,3872	0,8062

emoint : eMODIS Intégrale ; Emo max : eMODIS maximum ; spotint : SPOT VEGETATION intégrale ; spotmax : SPOT VEGETATION maximum

### **3.4.2. Discussion partielle**

La production de la carte annuelle du disponible fourrager herbacé est réalisée chaque année à partir d'une équation de régression linéaire obtenue par la méthode des moindres carrés entre l'intégrale du NDVI ou le maximum NDVI et les données de biomasse *in situ*. Cette approche a fait l'objet de plusieurs travaux (Tucker et al., 2005; Maselli et al., 2000; Wylie et al., 1991; Diallo et al., 1991). En revanche, c'est la première fois que l'analyse couvre une longue période (12 ans) offrant la possibilité d'examiner le comportement du modèle dans le temps et l'espace. La zone pastorale du Niger couvre une superficie de 252 207.5 km<sup>2</sup>. Les échantillons utilisés pour les relevés au sol dans le cadre des campagnes d'évaluation de la biomasse varient de 150 à 450 km<sup>2</sup> suivant les années. Les échantillons des premières études faites sur le terrain par Wylie et al. (1991) ont varié entre 25 et 100 km<sup>2</sup> mais ne concernaient que l'Est et le Centre du pays. Elles ne prenaient pas en compte toute la diversité spatiale des écosystèmes pâturés. On note globalement dans cette étude que les coefficients de détermination obtenus pour les années 1986, 1987 et 1988 sont élevés (respectivement 0,68, 0,91 et 0,73) mais varient d'une année à l'autre tout comme ceux trouvés pour les années concernées par notre étude avec une erreur standard de + ou - 200 kg/ha (Wylie et al., 1991). Au Sénégal, une étude similaire menée avec des données NOAA -AVHRR pour les années 1987 et 1988 donnait des coefficients de détermination similaires à ceux du Niger (Diallo et al., 1991). Les coefficients de détermination varient suivant les années dans toutes les études, ils varient aussi suivant les zones. On observe que nos résultats sont légèrement inférieurs à ceux trouvés au niveau des deux études citées précédemment. Cette différence peut s'expliquer en partie par le fait que Wylie et al. (1991) et Diallo et al. (1991) ont utilisé les images NDVI provenant de NOAA AVHRR alors que nous avons utilisé des images issues des capteurs VEGETATION du satellite SPOT et celles de MODIS TERRA lissées en eMODIS. D'autre part, cette différence pourrait s'expliquer par les

conditions de relevé de la végétation notamment la difficulté de respecter rigoureusement les exigences de la méthode et surtout la date des observations qui peut avoir un impact important sur les masses mesurées. Le RMSE moyen est d'environ 360,5 kg/ha, si on rapporte cette erreur à la superficie totale de la zone pastorale (252207.5 km<sup>2</sup>) on trouve une erreur moyenne annuelle de 9092080 tonnes soit environ l'alimentation de 5387899 UBT pour 9 mois (saison sèche). Peut-on raisonnablement fonder la planification des interventions d'un pays avec ce niveau d'erreur ? Si non, il est alors nécessaire d'explorer d'autres pistes d'amélioration du modèle actuel. L'utilisation de l'intégrale ou du maximum NDVI de la période de la croissance de la végétation herbacée NDVI seul pour estimer la phytomasse, nous paraît limitatif, car des études menées ont montré qu'on peut assurer le suivi de la phénologie de la végétation à travers plusieurs paramètres. On peut par exemple explorer l'utilisation des métriques dérivées des séries d'images (Jönsson et Eklundh, 2004). On peut aussi envisager de combiner les indices de satisfaction des besoins en eau pendant les différents stades phénologique de la végétation. Les résultats de la comparaison inter-senseurs de l'étude ont montré que, pour l'estimation de la biomasse, il n'y a pas de différence significative à utiliser les NDVI de SPOT VEGETATION ou d'eMODIS. Ceci est une bonne nouvelle pour les utilisateurs (services techniciens) qui sont dans des pays où l'accès aux données est souvent difficile. L'intégrale ou le maximum NDVI ne présente pas de différence significative en termes de performance pour estimer les productions fourragères dans la zone d'étude, ce qui signifie que l'opérateur peut utiliser l'un ou l'autre source de données sans que cela n'entraîne de grande différence dans les résultats. Ceci va dans le même sens que l'étude de comparaison entre NOAA AVHRR, MODIS et SPOT VEGETATION qui montrait qu'il y a une similarité entre les trois capteurs pour la dynamique de la végétation (Tucker et al., 2005).

### **3.5. Conclusion partielle**

Les résultats ont montré qu'en dépit de corrélations significatives, les performances du modèle fluctuent d'une année à l'autre comme l'atteste les  $R^2$  qui varient suivant les années entre 0,52 et 0,73 mais les corrélations restent très significatives ( $P<.0001$ ). Aussi, il ressort que le coefficient de détermination est plus élevé dans les zones saharienne et nord sahélienne que dans la zone sahélienne typique. Les erreurs quadratiques moyennes (RMSE) sont quand même élevées, elles varient selon les années entre 120 et 460 kg MS/ha. Les tests non paramétriques de comparaison de moyenne (Test de Wilcoxon et des signes) ont montré qu'il n'y a pas de différence significative d'utiliser SPOT VEGETATION ou eMODIS (intégrale ou maximum). Cette possibilité d'intercomparabilité permet aux utilisateurs d'assurer la production de produits sans interruption avec une possibilité d'analyse convergente. Néanmoins, le modèle nécessite des améliorations importantes car l'erreur quadratique moyenne est de 52 %.

## **IV. Amélioration de la modélisation de la biomasse fourragère en zone sahélienne à travers la méthode de régression linéaire multiple**

### **4.1. Introduction**

Le lien entre la biomasse herbacée mesurée au sol et les indices de végétation établi par régression linéaire en région sahélienne a été largement traité (Tucker et al. 1983, Tucker et al. 1985 ; Justice et Hiernaux, 1986 ; Maïdagui et al, 1987 ; Hiernaux et Justice, 1986 ; Wylie et al., 1991 ; Diollo et al., 1991 ; Maselli et al., 1992 ; Groten et al., 1993). Cependant, d'autres études ont montré que la relation n'est pas toujours une fonction linéaire (Santin-Janin *et al.*, 2009 ; Bégué *et al.*, 2011). Selon, Diouf et al., 2014, les fonctions exponentielle et puissance sont plus performantes que les fonctions linéaires. Ailleurs, la méthode de régression linéaire multiple a montré des performances satisfaisantes pour l'estimation des rendements agricoles (Balaghi et al., 2008 ; Kouadio, 2012 ; Kouadio et al., 2014). En plus, des améliorations très importantes ont été obtenues au Sénégal avec la méthode par régression linéaire multiple combinant des métriques dérivées de FAPAR et de GeoWRSI. (Diouf A. et al., en 2015 ; Diouf A. et al., 2016).

Dans les précédentes parties de la thèse, nous avons mis en évidence d'une part, que le modèle BIOMASA du CRA donne des moyennes simulées différentes des moyennes observées ; d'autres part que des modèles comme celui du MEIA présentent des  $R^2$  disparates qui varient fortement d'une année à l'autre. En outre, les RMSE sont souvent très élevées ce qui illustrait l'instabilité de ce modèle limité à une seule variable. Au regard de l'importance de l'estimation de la biomasse fourragère herbacée au Sahel en général et en particulier au Niger, il est nécessaire de trouver un modèle plus fiable permettant aux utilisateurs chargés du suivi et de l'estimation des ressources fourragères, de disposer d'une méthode plus fiable pour estimer la

production fourragère en vue d'aider les décideurs à anticiper les crises pastorales et à satisfaire les besoins du marché.

L'étude s'est dès lors fixée pour objectif global de contribuer à l'amélioration des méthodes d'estimation de la masse aérienne des herbacées. De manière plus spécifique, il s'agira d'utiliser des variables biophysiques et agrométéorologiques pour réaliser un modèle de régression linéaire multiple capable de prédire au mieux le rendement fourrager au Niger.

L'hypothèse de cette étude est de considérer qu'il est possible d'utiliser les métriques dérivées des séries temporelles d'images NDVI décennales de la période de croissance de la végétation grâce au logiciel VAST (Vegetation Analysis in Space and Time) et de les associer aux sorties du logiciel agrométéorologique AgrometShell pour dériver un ensemble de variables explicatives qui vont servir dans un modèle de prévision des rendements fourragers des herbacées. Le modèle issu de la méthode de régression linéaire multiple sera appelé modèle de référence, car il provient de plusieurs étapes de sélections objectives, dont la sélection des variables explicatives et celle du modèle pertinent et parcimonieux.

## **4.2. Matériel et Méthodes**

Les données biophysiques du modèle proviennent du NDVI de SPOT-VEGETATION ; Les données agrométéorologiques d'entrée sont les pluies mesurées au niveau des toutes les stations du pays et celles estimées par satellite RFE2 de FEWSNET, l'ETP (évapotranspiration potentielle) issues du centre européen appelé "European Centre for Medium-Range Weather Forecasts" (ECMWF). Le programme informatique VAST et le logiciel Agrometshell (AMS) ont été utilisés pour générer les variables explicatives. La variable expliquée est le rendement fourrager mesuré sur les sites de contrôle au sol par le MEIA ; les analyses statistiques ont été réalisées avec le logiciel de traitement statistique SAS JMP. Les étapes du traitement statistique

sont subdivisées en sept points : i) l'élimination des variables inutiles ; ii) la sélection des variables les plus pertinentes ; iii) la recherche exhaustive des modèles ( $2^k$  modèles possibles) ; iv) la sélection des meilleurs modèles (RMSE min et parcimonie de paramètres) ; v) la validation croisée, vi) l'analyse des résidus et vii) la présentation des modèles de références suivant les années, écorégions, faciès et à l'échelle globale (Figure 4.1).

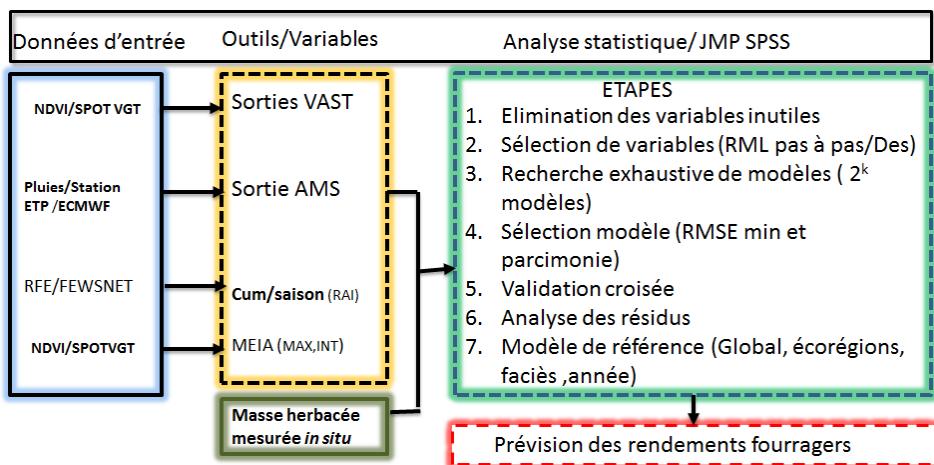


Figure 4.1 : Schéma général de la méthode de régression linéaire multiple conduisant à la prévision des rendements fourragers

#### **4.2.1. Matériel**

Le matériel est l'ensemble des données, logiciels et programmes ayant servi à la réalisation de l'étude. Les données principales utilisées sont :

- les données biophysiques (série d'images SPOT VEGETATION) ayant permis de dériver une partie des variables exogènes (explicatives, indépendantes) du modèle à l'aide de VAST ;
- les données météorologiques et agronomiques au nombre desquelles on peut citer la pluviosité mesurée dans les stations météorologiques, l'évapotranspiration potentielle, la phénologie de la végétation. Elles sont utilisées comme variables d'entrée pour le calcul des sorties du logiciel AMS ;
- le cumul annuel des pluies RFE-FEWSNET ;
- le maximum et l'intégrale du NDVI tels que calculés par le MEIA ;
- enfin, la masse aérienne des herbacées mesurée *in situ* (source MEIA) constituant la variable dépendante.

En ce qui concerne les logiciels et programmes informatiques, on peut citer :

1. les logiciels de traitement, d'analyse et d'affichage d'images satellitaires (VGTEExtract, WINDISP, VAST) ;
2. les logiciels de calcul de variables agrométéorologiques (AGROMETSHELL1.157)
3. les logiciels d'analyse statistique (SAS-JMP et SPSS).

##### **4.2.1.1. Données agrométéorologiques**

###### **Pluies**

Les variables pluviométriques proviennent de la base de données du Centre régional AGRHYMET (CRA). Elles ont été structurées sous un format compatible avec le logiciel AMS (calculs de séries décennales et de la moyenne 1971-2000 par station), ensuite importées dans ledit logiciel. Au total, 199 stations pluviométriques ont été renseignées et utilisées dans la Base de données AgroMetShell (BD AMS) pour le Niger.

## **Évapotranspiration potentielle**

Les données ECMWF utilisées couvrent la période de 1978 à 2012. Elles ont permis de calculer à un pas de temps décadaire. Les données ECMWF ERA INTERIM sont présentées en grille de points espacés de 0,25 degré de latitude et longitude. Pour extraire les ETP (Evapo Transpiration Potentielle) suivant les stations contenant les pluies mesurées, les étapes suivantes ont été suivies : i) construction d'une grille régulière de polygones de 0,25 degré de côté. Dans cette grille chaque polygone est centré autour d'un point ECMWF ; ii) une jointure spatiale pour affecter les attributs des points ECMWF à chaque grille correspondante ; iii) utilisation de la couche polygone issue de l'étape (i) pour réaliser une deuxième jointure spatiale, afin de porter les données ETP sur les stations du Niger.

### **4.2.2.2. Outils**

#### **VGExtract**

VGExtract est un logiciel gratuit, convivial, utilisable en mode batch. Il est développé par VITO pour la décompression, l'extraction suivant une fenêtre spatiale choisie des images SPOT VEGETATION et leur sauvegarde sous un format approprié (ILWIS, ENVI, RST, GeoTiff, RAW, WINDISP).

#### **VAST**

L'acronyme VAST de l'application signifie Vegetation Analysis in Space and Time. Ce programme informatique a été utilisé dans notre travail comme outil pour extraire les paramètres biophysiques dérivés des séries annuelles de NDVI (1998 à 2012). VAST a été développé dans les années 1990 par Felix Lee à l'époque assistant technique de FewsNet au Tchad, pour analyser systématiquement les séries d'images NDVI. Le programme analyse la série annuelle d'images NDVI pour en dériver les paramètres phénologiques suivants (Figure 4.2) :

- PEAK : la décennie à laquelle le NDVI atteint son maximum ;

- SDAT : la décade de début de la saison de végétation ;
- HORZ = PEAK – SDAT ;
- SVAL : la valeur du NDVI à SDAT ;
- PVAL : la valeur du NDVI à PEAK ;
- VERT = PVAL – SVAL ;
- EVAL : le NDVI au temps PEAK + 4 (soit environ la fin de saison) ;
- DROP = PVAL – EVAL ;
- SLOP : la pente de la droite qui rejoint (SDAT, SVAL) à (PEAK, PVAL) ;
- CUMM : la somme des valeurs de NDVI de SDAT à PEAK ;
- SKEW : le rapport entre la somme des 3 valeurs de NDVI suivant PEAK (de PEAK + 1 à PEAK + 3) et la somme des 7 valeurs de PEAK — 3 à PEAK + 3.

Pour des raisons pratiques, les paramètres phénologiques dérivés de VAST seront repris par leurs trois premières lettres (annexe 4.1).

Les images NDVI sont d'abord enregistrées selon une nomenclature particulière : le nom doit comprendre respectivement 2 lettres suivies de 2 chiffres correspondant à l'année, suivis de 2 chiffres pour le mois et d'un chiffre pour la décade terminée par l'extension. Exemple : DV12021.NEG.

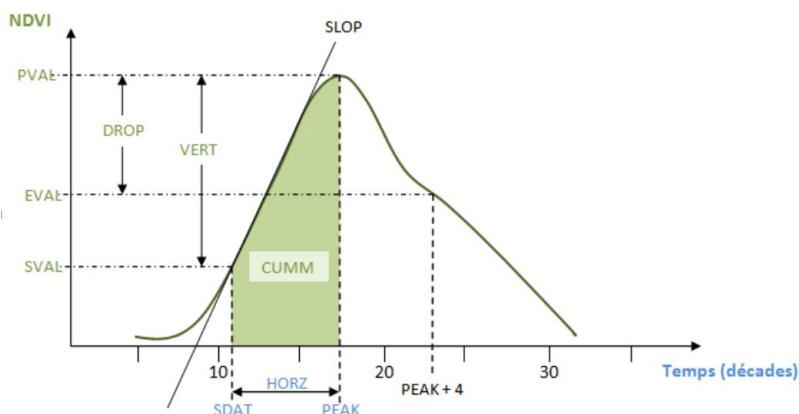


Figure 4.2 : Schéma de métriques calculées sur VAST  
Source : manuel d'utilisation de VAST

## **WINDISP**

C'est un logiciel libre développé par la FAO pour le Global Information and Early Warning System. Il dispose de modules permettant entre autres : l'affichage, l'analyse d'images, des couches vectorielles et des bases de données associées ; de produire des graphes représentant l'évolution d'une série temporelle d'images (NDVI, pluviométrie...) relatifs à une zone d'intérêt ; de superposer des images et des cartes d'unités administratives afin d'en extraire des statistiques ; de calculer aussi des statistiques pour chaque pixel d'une série d'images. Il est principalement reconnu pour les fonctions de visualisation et de traitement des images satellites de NDVI, de pluies estimées par satellite (RFE) et autres comme les sorties de VAST.

## **AGROMETSHELL 1.57 (AMS)**

Le AgroMetShell est un outil développé par la FAO en 2007 pour le suivi des cultures et la prévision des rendements agricoles. Il permet de simuler le bilan hydrique et les risques de déficit de production. Nous avons utilisé AMS dans le cadre de ce travail pour calculer les paramètres agrométéorologiques qui constitueront les variables d'entrée de notre modèle de prévision des rendements fourragers ( $\text{kg.MS.ha}^{-1}$ ). AMS contient une base de données facile à mettre à jour régulièrement afin d'assurer une production régulière des variables. Les modules contenus dans le logiciel, basés sur le calcul du bilan hydrique permettent d'analyser l'impact des facteurs climatiques sur les différentes cultures. En réalité, c'est un programme informatique fondé sur l'hypothèse que les rendements peuvent être expliqués par le contexte agrométéorologique de la zone considérée. Cet outil sert à l'alerte précoce et à la sécurité alimentaire car il permet de procéder à l'évaluation des effets du climat sur les cultures et de prédire les rendements agricoles à travers la modélisation statistique. En résumé, AMS se base sur le modèle CSSWB (Crop Specific Soil Water Balance) qui est un modèle de bilan hydrique permettant d'estimer l'impact des conditions climatiques sur les cultures (Allen *et al.* 1988). CSSWB

calcule le bilan hydrique pour un pas de temps décadaire suivant l'équation :

$$W_t = W_{t-1} + R - ETc - (r + i)$$

Avec :

- $W_t$ , la quantité d'eau stockée dans le sol au moment t
- $W_{t-1}$ , la quantité d'eau stockée dans le sol à la fin de la période précédente t-1
- R, la quantité de pluie cumulée sur la période de temps t (souvent une décennie)
- ETc, l'évapotranspiration potentielle de la culture sur la période t
- r, la perte d'eau due à l'écoulement sur la période t
- i, la perte d'eau liée à la percolation profonde sur la période t

La simulation du bilan hydrique dans AMS fournit plusieurs variables dont les plus importantes sont :

- la quantité d'eau requise pour le cycle complet de la culture (TWR) ;
- l'indice de satisfaction en eau en fin de cycle (Idx, IndxNor, IndxLatest) ;
- le contenu initial en eau dans le sol (SWi)
- l'excès en eau à différents stades phénologiques du cycle de croissance : phase initiale, phase végétative, floraison, maturité et durant tout le cycle (donnée par la somme des autres valeurs) : WEXi, WEXv, WEXf, WEXr, WEXt ;
- le déficit en eau à différents stades phénologiques et déficit en eau total en fin de cycle (WDEFi, WDEFv, WDEFf, WDEFr, WDEFt) ;
- l'évapotranspiration réelle à différents stades phénologiques et sa valeur totale en fin de cycle (ETAi, ETAv, ETAf, ETAr, ETAt) ;
- crossing data Cr1a à Cr4a (calculés avec les données pluviométriques réelles) et Cr1n à Cr4n (Calculés avec les données pluviométriques moyenne sur trente ans) indiquent les décades

auxquelles le rangeland index (RI) croise la ligne de 0.4 \*PET. Ces décades peuvent être associées aux décades de semis des cultures. Le rangeland index (RI) correspond à l'indice de satisfaction en eau développé par la FAO calculé sur une période de 5 décades avec l'évapotranspiration normale prise à son niveau potentiel.

En plus des données dérivées de VAST et d'AgroMetShell, nous avons utilisé le cumul des pluies compté à partir du début de la saison des pluies (première décade de mai à la troisième décade d'octobre) et estimé par satellite (RAI) dérivé de RFE2, le maximum (MAX) et l'intégrale (INT) du NDVI utilisés par le MEIA pour réaliser la régression linéaire simple.

#### **4.2.2. Méthode**

##### **4.2.2.1. Présence minimum de végétation**

La connaissance de la valeur qui représente la présence minimum de végétation est nécessaire pour le calcul de la date de démarrage de la saison de croissance végétative. Pour déterminer ce seuil de présence minimale de la végétation, nous nous sommes basés sur la bibliographie et l'examen des profils NDVI de chaque site. Selon les auteurs, la valeur réelle de NDVI représentant le seuil minimal de présence de la végétation ligneuse et herbacée est 0,1 (Hiernaux, 1984). La formule du calcul de la valeur numérique étant : NDVI SPOT VEGETATION = (DN\* 0,004) - 0,1, cette valeur numérique ou Digital Number (DN) est égale à  $0,2/0,004 = 50$ . Les séries d'images NDVI SPOT VEGETATION des années allant de 2000 à 2012 ont été utilisées pour extraire les paramètres phénologiques de la végétation au niveau des sites de mesure de biomasse *in situ*. Le mois de juin a été retenu comme date minimale du début de la saison de végétation, et celui d'octobre comme date de fin maximum de la saison de végétation, la valeur DN50 comme valeur minimale de présence végétative, 05 comme variation minimale de DN entre deux décades.

La même opération a été faite pour les images eMODIS.  $eMODIS\ NDVI = (DN * 0,01) - 1$ , Si la valeur réelle du NDVI représentant le seuil minimal de présence de la végétation est 0,1, alors la valeur numérique ou Digital Number (DN) de ce seuil est égal à  $1,1 / 0,01 = 110$ . Les variables fixes de détermination des paramètres phénologiques dérivés de chaque série décadaire de NDVI des années allant de 2000 à 2012. La valeur 6 est mise pour le mois de juin comme date du début minimum de la saison de végétation, mois d'octobre (10) comme la fin maximum de la saison de végétation, DN110 comme valeur minimale de présence végétative, 05 comme variation minimale de DN entre deux décades.

#### **4.2.2.2. Décades d'installation (planting dekads).**

Pour déterminer les décades d'installation de la végétation herbacée annuelle, deux possibilités sont disponibles dans le logiciel AMS : la première se base sur un seuil (à fixer) de pluie efficace, suivi d'autres précipitations pendant les deux décades suivantes ; la seconde est basée sur un seuil de pourcentage du besoin total en eau de la végétation. La deuxième option a été utilisée en fixant le seuil à 10%.

#### **4.2.2.3. Longueur de la végétation**

Il y a très peu d'écrits sur le suivi de la phénologie des espèces fourragères naturelles. Le suivi du cycle phénologique de certaines graminées et légumineuses montre que la longueur varie selon les espèces et le régime hydrique du milieu. Les valeurs de cette longueur végétative varient entre 5 et 8 décades selon les espèces (Penning et al., 1982). Une moyenne de 7 décades a été retenue comme durée de la longueur de la végétation fourragère.

#### **4.2.2.4. Analyse statistique**

Avant toute analyse approfondie, il est important de vérifier la sensibilité de notre échantillon en calculant les biais sur les paramètres de l'échantillon de base en procédant à un re-échantillonnage avec

remise. Si notre échantillon initial suit une loi normale, alors il n'y aura pas de grande différence entre ses paramètres et ceux issus du re-échantillonnage. Les biais ainsi calculés correspondent aux erreurs sur les paramètre de l'échantillon. C'est la technique appelée Bootstrap. Elle a été initiée vers la fin des années 1970 comme un autre regard sur la méthode Jackknife (Efron, 1979 et Kouadio, 2012). Dans le cadre de cette étude, nous avons créé 2000 sous échantillons pour calculer les biais sur la moyenne, la variance et l'écart type. La régression est une technique d'analyse statistique qui permet de décider, contrôler, et prévoir. Elle permet de formuler une relation mathématique entre la variable expliquée (endogène, prédictive, dépendante) avec une ou plusieurs variables explicatives (exogènes, prédicteurs, indépendants). Les variables indépendantes dans cette étude ont été dérivées des séries d'images NDVI traitées avec le programme VAST, les données agrométéorologiques dérivées d'AMS, les pluies cumulées, Max NDVI et Int NDVI. Les rendements fourragers mesurés *in situ* représentent la variable dépendante. Les objectifs spécifiques de notre analyse statistique sont : d'ajuster le meilleur modèle pour expliquer le rendement fourrager en fonction des variables pertinentes dérivées de AMS et VAST, RFE, MAX et INT du MEIA ; de prédire les valeurs du rendement fourrager pour des nouvelles valeurs des variables explicatives. La sélection des variables pertinentes a été réalisée avec la méthode pas-à-pas descendante :

### **Étape 1 : Vérification des données**

La vérification des données a été réalisée par l'examen des variances des variables indépendantes en vue de faire un premier choix de celles qui sont les plus fiables. Cette opération a permis d'éliminer toutes les variables explicatives ayant une variance nulle. Ce type de variable représente une constante et n'apporte pas d'information utile dans un modèle de régression.

## **Étape 2 : Sélection des variables**

Il existe plusieurs types de procédure de sélection des variables telle que la méthode pas-à-pas ascendante ou descendante. Les procédures de type pas à pas s'initialisent avec plusieurs variables explicatives. Les différentes variables sont éliminées ou sélectionnées selon des critères de variance. La méthode descendante a été privilégiée tout en nous réservant la possibilité de tester la méthode mixte. Cette approche permet de prendre en compte toutes les variables possibles sans en négliger certaines. Un panel de 34 variables a été utilisé. La sélection des variables explicatives représente une étape importante dans la régression multiple. La procédure de type pas-à-pas disponible sur le logiciel SAS/JMP pour sélectionner les variables a été utilisée.

## **Étape : 3 et 4 : Sélection des modèles**

La sélection des modèles est une problématique très connue en statistique, elle a été largement développée par plusieurs auteurs (Good et Hardin, 2012, Meier et al., 2011, Weiers et al., 2011, Johnson et al., 2010). Plusieurs critères existent pour la sélection du meilleur modèle : le Root Mean Square Error (RMSE); le coefficient de Détermination ( $R^2$  qui augmente avec le nombre de variables); le Coefficient de Détermination ajusté ( $R^2_{aj}$  qui corrige certains défauts de  $R^2$  en tenant compte du nombre de paramètres du modèle); Bayesian Information Criterium (BIC); l'Akaïke Information Criterium (AIC) ; l'Akaïke Information Criterium Corrigé (AICC); coefficient de Mallows (CP) si le modèle est bon le Cp est très proche de P (p étant le nombre de variables) ; etc. En général ces critères ne se contredisent pas, mais peuvent être plus pertinents les uns par rapport aux autres suivant les cas et permettent d'identifier les meilleurs modèles. Un modèle parcimonieux combine un petit nombre de variables explicatives et donne le  $R^2$  ajusté le plus élevé.

Le choix du meilleur modèle s'est appuyé sur le  $R^2$  ajusté le plus élevé, le RMSE minimum et un nombre de variables explicatives faible (pas

plus de 4 variables). La méthode de recherche exhaustive est très efficace pour la sélection des modèles surtout quand on a un nombre limité de variables explicatives. Si nous avons  $k$  variables explicatives, le nombre de modèles possibles est de  $2^k$ . L'étape de sélection des variables a permis de sélectionner un nombre limité de variables explicatives. La démarche retenue réalisée avec SAS/JMP consiste à réaliser les  $2^k$  combinaisons possibles et les classer par ordre décroissant selon le  $R^2$  ajusté et selon le nombre de variables explicatives. Cette démarche permet de sélectionner le meilleur modèle pour chaque nombre de variables explicatives du modèle.

#### **Étape 5 : Validation croisée**

La technique du Leave-one-out cross validation (LOOCV) a été utilisée pour valider les modèles choisis en vue d'identifier le meilleur. Le LOOCV est une technique qui est utilisée quand la taille de l'échantillon utilisée est faible et ne permet pas de créer un groupe de données distinct pour la calibration et pour la validation de taille suffisante. Il permet de choisir le modèle optimal en testant la précision prédictive et/ou l'erreur de généralisation. La démarche consiste à subdiviser l'échantillon de  $n$  observations en  $k$  sous-ensembles égaux, faire la calibration avec  $k-1$  sous-ensembles et valider avec le  $k^{\text{ième}}$  sous-ensemble, répéter la même opération pour tous les  $k$  sous ensemble (Emmert-Streib et Dehmer, 2009). Le LOOCV est un cas particulier de la technique  $k$ -fold où  $k = n$ , c'est une technique très puissante permettant de choisir le modèle le plus intéressant (Cornillon et Matzner-Løber, 2007). Dans le cadre de cette étude, la technique du LOOCV a été utilisée pour la détermination du modèle ayant la meilleure RMSE (minimisée) avec un nombre de variables inférieur ou égal à 4.

#### **Étape 6 : Analyse de résidus**

Dans l'approche de régression, un modèle n'est considéré comme bon que si les résidus obéissent aux hypothèses de normalité qui peuvent

être examinées à travers la droite de Henri ; l'égalité des variances des erreurs (homoscédasticité) qui signifie que toutes les distributions de Y doivent avoir le même écart-type : la variance résiduelle est alors constante sur le domaine étudié ; enfin, quand il n'y a pas de multicolinéarité (Johnson et al., 2010). Dans le cadre de ce travail, les hypothèses sur les résidus ont été vérifiées avec le logiciel SAS/JMP. Il s'agit là de vérifier que les résidus suivent une loi normale ; en pratique, on représente graphiquement les résidus par rapport à la variable réponse prévue pour s'assurer que le graphique ne présente aucune structure particulière et enfin on applique le test de Watson-Durbin qui est très approprié pour analyser l'autocorrélation surtout quand il y a une constante dans la régression (Badi, 2011 ; Weiers et al., 2011).

## **Étape 7 : Prévision**

Le modèle retenu à la fin du processus peut être utilisé pour prédire de nouvelles valeurs responses à l'aide des nouvelles valeurs de variables explicatives. La relation mathématique entre les variables explicatives (métriques issues de AMS et VAST) et la variable expliquée ou dépendante (masse fourragère herbacée) est utilisée pour estimer le rendement fourrager sous forme de produits cartographiques.

### **4.3. Résultats et discussion**

#### **4.3.1. Résultats**

##### **4.3.1.1. Analyse exploratoire des mesures de masse herbacée de 2001 à 2012**

L'analyse exploratoire des mesures de masse herbacée de 2001 à 2012 sur 319 observations montre que la moyenne est d'environ 700 kg MS.ha<sup>-1</sup> avec un écart type de 531 kg MS.ha<sup>-1</sup>. Les résultats de l'analyse bootstrap basée sur 2000 sous échantillons, montrent que cette moyenne varie de 642 kg à 762 kg dans un intervalle de confiance de 95 %, les biais sur la moyenne et l'écart type sont respectivement de 0,96 kg MS.ha<sup>-1</sup> et de -0,85 kg MS.ha<sup>-1</sup> (Tableau 4.1). Il y a au total 34

variables indépendantes provenant des sorties de AMS, de VAST, du cumul saisonnier de pluie RFE de FEWSNET et les deux variables utilisées par le MEIA (INT et MAX).

Tableau 4.1 : Analyse exploratoire des mesures de masse herbacée de 2001 à 2012 par bootstrap (kg MS.ha<sup>-1</sup>)

Termes	Statistique	Erreurs standard	Bootstrap			
			Biais	Erreurs standard	Intervalle de confiance à 95 %	
Moyenne	699,11	30,12	0,96	30,20	642,94	762,26
Masse	Variance	282 097,66	-165,40	28 872,303	228 150,66	343 005,93
herbacée	Écart-type	531,13	-0,85	27,18	477,65	585,67

Unité : kg MS.ha<sup>-1</sup>

#### **4.3.1.2. Ajustement global de la productivité fourragère sur l'ensemble de la zone étudiée**

Il ressort de la mise en œuvre de la procédure de sélection de variables du type pas-à-pas descendante, que les variables retenues sont les suivantes : MAX, DRO, EVA, HOR, PEA, PVA, SLO (annexe 4.1). La méthode de recherche exhaustive a été appliquée à ces 7 variables indépendantes pour établir tous les modèles possibles. Les modèles au nombre de  $2^7$  (128) sont automatiquement classés par ordre décroissant des RMSE et selon le nombre de variables. Les résultats (Tableau 4.2) donnent les quatre meilleurs premiers modèles. Ils sont listés suivant le nombre de variables explicatives selon le critère du RMSE minimum.

Tableau 4.2 : Quatre meilleurs modèles globaux en fonction du nombre de variables

Modèles	n°	R <sup>2</sup> cal	R <sup>2</sup> aj cal	R <sup>2</sup> val	RMSE cal. KG MS.ha <sup>-1</sup>	RMSE val. KG MS.ha <sup>-1</sup>	Dif RMSE
Y= -603,13+4590,81 max	1	0,57	0,57	0,57	354,47	353,42	1,05
Y= -1193,01 + 2822,30 max + 15,51 DRO	2	0,62	0,62	0,61	308,44	310,65	2,21
Y=-388,01+3133,09 max -15,41DRO +17, 62 VER	3	0,66	0,66	0,65	294,16	297,26	3,10
Y= -2190,82+3344,13 MAX -20,46 DRO +20,78 VER+ 74,06 PEA	4	0,69	0,68	0,67	285,22	288,94	3,72

cal : calculé ; aj : ajusté ; Dif : différence ; val : validation

Le meilleur modèle, selon le critère du RMSE est le n° 4. En effet, il donne un RMSE relatif de 40 %. Les quatre variables sélectionnées pour cette équation sont : MAX, DRO, PEA, VER. L'examen des probabilités des coefficients estimés de ce modèle donne des résultats hautement significatifs au seuil de 1 pour 10 000 (Tableau 4.3). Le graphique des valeurs observées en fonction des valeurs prédictes pour l'ajustement global est donné par la Figure 4.3.

Tableau 4.3 : Estimation des coefficients du modèle global à quatre variables

Variables	Estimation	Erreur standard	Rapport t	Prob.> t
Constante	-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*

\*significatif au seuil de 1 pour 10000

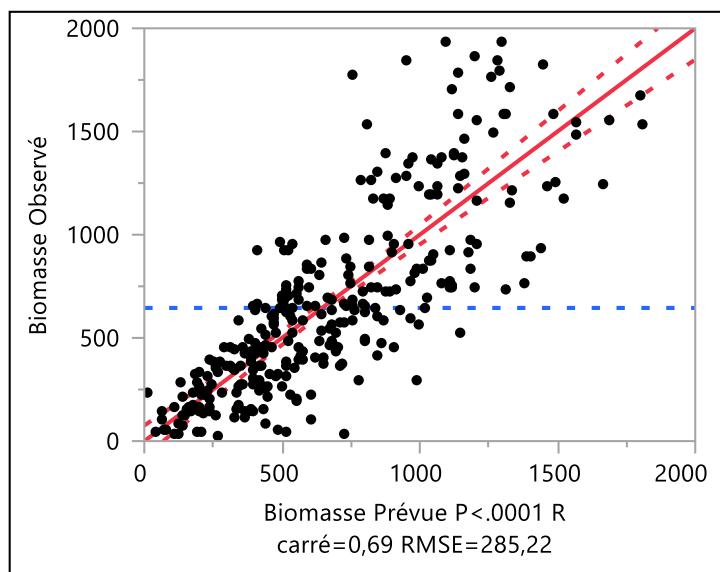


Figure 4.3 : Les valeurs observées en fonction des valeurs prédictes à l'échelle globale

### **Analyse des résidus de l'ajustement global**

L'analyse exploratoire indique un écart type de 283,33 kg. Les résultats de l'analyse bootstrap basée sur 2 000 sous échantillons montrent que la moyenne globale varie de 642 et 762 Kg dans un intervalle de confiance à 95 %, les biais sur la moyenne et l'écart type sont respectivement de 0,89 et -0,35 relèvement faibles (tableau 4.4).

Tableau 4.4 : Analyse exploratoire bootstrap des résidus

Paramètres	Statistiques	Bootstrap			Intervalle de confiance à 95 %	
		Biais	Erreur standard	Inférieur	Supérieur	
Nombre d'observations	319			319	319	
Moyenne	0,00	0,89	16,02	-32,49	33,16	
Écart-type	283,33	-0,35	14,86	255,30	313,02	
Variance	80 273,71	24,69	8440,26	65 176,17	97 979,15	

Unité : kg.MS.ha<sup>-1</sup>

Le diagnostic du modèle global montre une répartition régulière des résidus suivant les lignes (Figures 4.3 a), la distribution suit une loi normale (Figure 4.3 b) attestée par un indice Durbin-Watson DW de 1,83 et un pourcentage d'autocorrélation de 8.

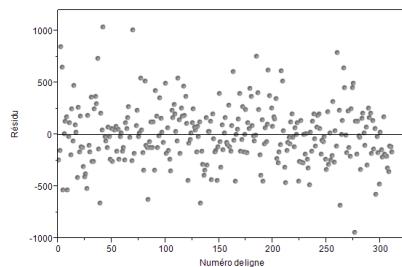


Figure 4.4 a : Graphique des résidus par ligne

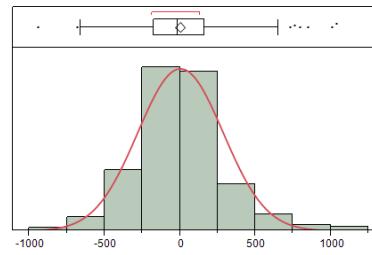


Figure 4.4 b : Distribution des résidus

#### 4.3.1.3. Analyse suivant les écorégions

Le calcul de la répartition spatiale du nombre de sites suivant les écorégions montre que seuls Azaouak, le maga1 et le manga2 ont suffisamment d'observations pour permettre de réaliser des ajustements de la productivité fourragère herbacée en fonction des variables explicatives

##### Les modèles de l'écorégion Azaouak

L'Azaouak compte 150 observations. La moyenne des observations de ces sites est de  $684,69 \text{ kg MS.ha}^{-1}$ . La mise en œuvre de la procédure de sélection de variables de type pas à pas descendante au niveau de cette région écologique a permis d'obtenir les variables pertinentes suivantes : VER, PEA, DRO, MAX (tableau 4.5)

Tableau 4.5 : Les quatre meilleurs modèles de l’Azaouak en fonction du nombre de variables

Modèles	R <sup>2</sup> cal	R <sup>2</sup> aj cal	R <sup>2</sup> val	RMSE cal kg.MS.ha-1	RMSE val .kg.MS.ha-1	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 PEA +13,80 VER	0,72	0,72	0,71	270,71	274,96	3,91
20 PEA +3070,61MAX +43 VER-1192,03	0,74	0,74	0,73	223	226	3
Y = -2800,78 + 2979,32max -23,84 DRO + 98,49 PEA +27,14 VER	0,77	0,76	0,75	250,5	255,25	5,21

cal : calculé ; aj : ajusté ; Dif : différence ; val : validation

## Le meilleur modèle sur l’Azaouak

Le meilleur ajustement sur l’Azaouak du point de vue du minimum RMSE comporte quatre variables que sont le MAX, le DRO, le PEA et le VER. L’estimation des paramètres du modèle montre des probabilités hautement significatives (Tableau 4.6). La représentation des valeurs observées en fonction des valeurs prévues montre un  $R^2$  de 0,76 et un RMSE de 250 kg (Figure 4.4). Il est caractérisé par un RMSE relatif de 36 %.

Tableau 4.6 : Estimation des coefficients du modèle de l’Azawak

Terme	Estimation	Erreur standard	Rapport t	Prob.> t
Constante	-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*

\*significatif au seuil 1 pour 10000

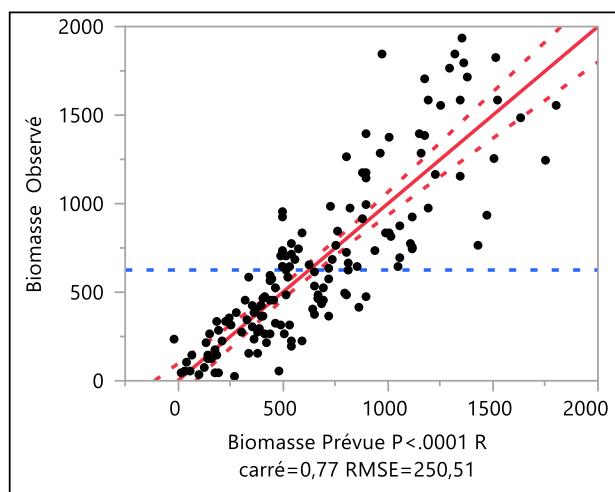


Figure 4.5 : Les valeurs observées en fonction des valeurs prédictes pour le modèle de l’Azaouak

## Analyse des résidus

L'analyse des résidus de l'Azaouak indique un écart type de 247 kg (Tableau 4.7). Le diagnostic du modèle de cette région écologique montre une répartition régulière (Figure 4.5 a) et une distribution des résidus qui suit une loi normale (Figure 4.5 b) comme l'atteste l'un indice DW de 1,59.

Tableau 4.7 : Estimations des coefficients

Type	Coefficient	Estimation	Inférieur à 95 %	Supérieur à 95 %
Position	$\mu$	0	-39,87	39,87
Dispersion	$\sigma$	247	221,96	278,76

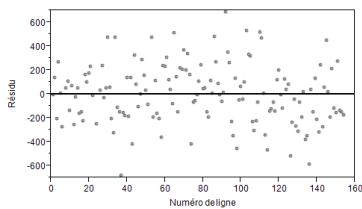


Figure 4.6 a : Graphique des résidus par ligne

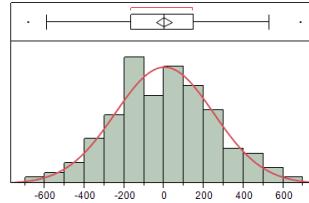


Figure 4.6 b : Distribution normale des résidus

## Le modèle de Manga2

Le Tableau 4.8 montre que le Manga2 compte 48 observations, la moyenne des ces observations est  $865,26 \text{ kg MS.ha}^{-1}$ . La mise en œuvre de la procédure de sélection de variables du type pas à pas descendante au niveau de cette région écologique a permis d'obtenir l'équation suivante qui comporte trois variables pertinentes que sont IndxNor, VER et RAI :  $Y = 13,66 + 22,70 \text{ VER} + 1,76 \text{ RAI} - 1286,37 \text{ IndxNor}$

La représentation des valeurs observées en fonction des valeurs prévues montre un  $R^2$  de 0,73 et un RMSE calculé de  $240 \text{ kg.ha}^{-1}$  (Figure 4.6). Le modèle est caractérisé par un RMSE relatif de 27 % (Tableau 4.8). L'estimation des paramètres du modèle montre des probabilités hautement significatives (Tableau 4.9).

Tableau 4.8 : Résumé de l'ajustement

Paramètres	Valeurs
R <sup>2</sup>	0,72
R <sup>2</sup> ajusté	0,70
R <sup>2</sup> validation	0,66
RMSE calibration	240
RMSE validation	254
Moyenne de la réponse	865
Observations	43

Tableau 4.9 : Estimations des coefficients

Terme	Estimation	Erreur standard	Rapport t	Prob. >  t
Constante	-1286,37	272,35	-4,72	<,0001*
IndxNor	13,66	5,76	2,37	0,0228*
VER	22,70	2,76	8,22	<,0001*
RAI	1,7	0,50	3,51	0,0011*

\*significatif au seuil de 10 pour 1000 ; \*\*\*significatif au seuil de 1 pour 10000

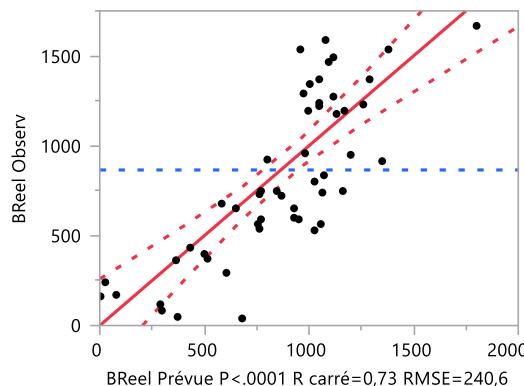


Figure 4.7 : Les valeurs observées en fonction des valeurs prévues pour le modèle du Manga2

## Analyse des résidus du modèle de Manga2

L'analyse des résidus indique un écart type de 275 kg. Le diagnostic du modèle de Manga2 indique une répartition régulière (Figure 4.7 a) et une distribution normale des résidus (Figure 4.7 b), comme l'attestent l'un indice DW de 1,37 et le taux d'autocorrélation de 0,09.

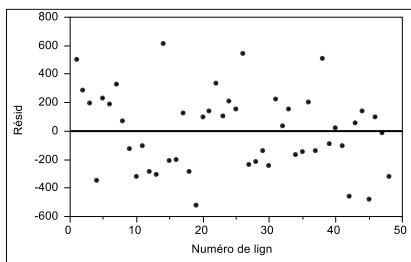


Figure 4.8 a : Graphique du résidu par ligne

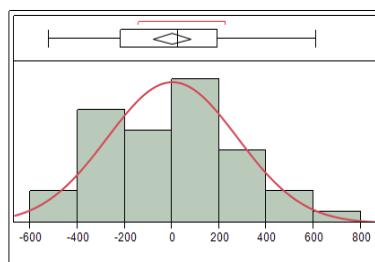


Figure 4.8 b : Distribution des résidus

Tableau 4.10 : Récapitulatif des meilleurs modèles réalisés avec les observations valides enregistrées de 2001 à 2012 par écorégion

Écorégion	Modèle	$R^2$	$R^2$	$R^2$	RMSE	RMSE	Dif
		cal	aj	val	cal kg.ha <sup>-1</sup>	val kg.ha <sup>-1</sup>	RMSE kg.ha <sup>-1</sup>
Azaouak	$Y = -2800,78 + 2979,32 \text{max} -23,84 \text{DRO} + 98,49$ PEA +27,14 VER	0,77	0,76	0,75	250,5	255,25	5,21
Managal	3523,16 MAX -357,65	0,54	0,53	0,50	188	192	16
Managa2	13,66* IndxNor + 22,70 VER + 1,76 RAI-1286,37	0,72	0,70	0,66	240	254	14

Cal : calculé ; aj : ajusté ; Dif : différence ; val : validation ; DW : Indice de Durbin-Watson

#### **4.3.1.4. Analyse par Type de sol FAO**

L’analyse de la répartition spatiale du nombre d’observations valides sur les 12 ans de mesures, suivant les types de sol FAO (1975) indique que seuls Ge5-1a (Gleysols eutriques et Arénosols luviques , grossier, plat à doucement ondulé), Qc1 (Arénosols cambiques), Qc7-1a(Arénosols cambiques et Gleysols eutriques, grossier, plat à doucement ondulé), Ql1-1a (Arénosols luviques, grossier, plat à doucement ondulé ) et Re35-a (Régosols eutriques et Arénosols luviques, plat à doucement ondulé) ont assez d’observations pour permettre de réaliser des ajustements de la productivité fourragère herbacée en fonction des variables explicatives. La mise en œuvre de la procédure de sélection de variables du type pas-à-pas descendante au niveau de ces types de sol a permis de retenir des modèles comprenant de 1 à 3 variables explicatives ; les R<sup>2</sup> des modèles varient de 0,75 à 0,86, ; les variables explicatives diffèrent d’un type de sol à l’autre. Cependant la variable MAX a été pertinente au niveau de 4 types de sol. Les meilleurs modèles sont consignés dans le tableau 4.11. L’annexe 4.2 indique les résultats détaillés des analyses suivant les types de sol.

Tableau 4.11 : Récapitulatif des meilleurs modèles réalisés avec les observations valides enregistrées de 2001 à 2012 par type de sol FAO

Type de sol	Modèle	R <sup>2</sup> cal	R <sup>2</sup> aj	R <sup>2</sup>	RMSE	RMSE	Dif	DW
				val	cal	Val	RMSE	
Ge5-1a	-936,85+ 22,79 EVA	0,85	0,84	0,79	148	161	13	1,82
Qc1	- 758,60 +19,73 VER +2029,28 MAX	0,86	0,85	0,81	190	205	15	2,35
Qc7-1a	-670,44 -10,07 ETAF+ 1,24 RAI +4197,70 MAX	0,75	0,73	0,70	201	209	8	1,83
Ql1-1a	24EVA+101,17PEA+2374,77MAX-4096	0,76	0,76	0,73	207	214	7	1,38
Re35-a	-567 +12,94 VER +2940,82 MAX	0,77	0,76	0,74	152	157	4	1,72

Cal : calculé ; aj : ajusté ; Dif : différence ; val : validation ; B = Cambisols ; G= Gkeysols ; J = Fluvisols ; Q =arenosols ; R =Regosols ; V= vertisols; Y = Yermosols ; L =Luvic ; DW : Indice de Durbin-Watson

#### **4.3.1.5. Analyse de la productivité fourragère herbacée par faciès**

L’analyse de la répartition spatiale du nombre d’observations valides sur les 12 ans de mesures, suivant les faciès indique que : le MA2\_Qc7-1a Nord\_sahélienne ; le MA2\_Qc1 nord-sahélienne ; le MA1\_Ql1-1a nord-sahélienne, Le MA1\_Qc7-1a nord-sahélienne, Le AZ\_Re35-a saharienne ; le AZ\_Re35-a nord-sahélienne ; le AZ Ql1-1a nord-sahélien et le AZ Ge5-1a nord-sahélien ont assez d’observations pour permettre de réaliser une régression linéaire multiple du rendement fourrager (mesures du terrain) en fonction des variables explicatives (métriques dérivées de VAST et AMS) . La mise en œuvre de la procédure de sélection de variables du type pas-à-pas descendante a permis d’obtenir des modèles comprenant 3 à 4 variables explicatives. Les R<sup>2</sup> des modèles varient de 0,93 à 0,72, les variables explicatives différent d’un faciès à l’autre aussi (Tableau 4.12).

Tableau 4.12 : récapitulatif des modèles par faciès

Faciès	N°	Modèles	R <sup>2</sup> cal	R <sup>2</sup> aj cal	R <sup>2</sup> val	RMSE cal.kg. MS.ha <sup>-1</sup>	RMSE val.kg. MS.ha <sup>-1</sup>	D-W
MA2_Qc7-1a Nord_sahélienne	1	2913+1,02 INDXLAT +11,26WDEFF -22,25 ETAF - 185,03 HOR	0,93	0,92	0,86	100	122	2,13
MA2_Qc1 nord- sahélienne	2	-2970,78-25,66 INDXHAR +43,93 INDEXNOR +37,48 ETAI +36,45 EVA	0,82	0,77	0,65	231	282	2
MA1_Ql1-1a nord-sahélienne	3	-4564,35+3580,39 MAX -10,06 DRO + 165,52 PEA + 2,12 RAI	0,86	0,81	0,74	126	141	2,23
MA1_Qc7-1a nord-sahélienne	4	5260,48 — 62,87 DRO +25,37 PVA +49,87 SLO + 176,36 SDA	0,78	0,7	0,42	168	224	2,33
AZ_Re35-a saharienne	5	-3433,08 3154,10 MAX +6,15 TWR -25,90 DRO + 26, 73 VER	0,86	0,82	0,74	121	139	2,26
AZ_Re35-a nord-sahélienne	6	-2304,57 -28,79 DRO +65,18 VER 82,91SDA	0,80	0,77	0,72	233	253	1,83
AZ Ql1-1a nord- sahélien	7	-3218,28 -22, 13 DRO +119,34 PEA +39,06 VER	0,78	0,77	0,75	269	278	2,16
AZ Ge5-1a nord-sahélien	8	-269,31 + 3005,07 MAX – 13, 45 INDXNOR + 2,39 WDEFR +15,83 EVA	0,77	0,72	0,64	279	311	2,15

Cal : calculé ; aj : ajusté ; Dif : différence ; val : validation ; B = Cambisols ; G= Gkeysols ; J = Fluvisols ; Q =arenosols ; R =Regosols ; V= vertisols;  
Y = Yermosols ; L =Luvic ; DW : Indice de Durbin-Watson

#### **4.3.1.6. Modèle d'estimation suivant les années**

La mise en œuvre de la procédure de sélection de variables du type pas-à-pas descendante suivant les années donne les résultats consignés dans le tableau 4.13, qui indique le récapitulatif des paramètres des modèles suivant les années. Le  $R^2$  ajusté varie 0,94 à 0,57 ; le RMSE relatif varie entre 11 % et 58 %. Les variables explicatives diffèrent d'une année à l'autre. L'annexe 4.3 indique les résultats détaillés des analyses suivant les années.

Tableau 4.13 : Récapitulatif des paramètres des Modèles suivant les années

Année	Modèles	R <sup>2</sup>	R <sup>2</sup> aj	R <sup>2</sup>	RMSE cal		RMSE val		DW	
		cal	cal	val	(kg. MS.ha <sup>-1</sup> )	Abs	Relative	Abs	Relative	
2001	72,43 -20,58 INDXNOR + 6,40 ETAT +32,11 EVA -30,99 SVA	0,83	0,79	0,67	156,18	25%		178,89	29%	1,92
2002	-201,30 + 4006,70 MAX - 6,41 DRO -0,55 RAI	0,9	0,88	0,85	83,16	21%		95	24%	1,83
2003	-13,62 + 9475,86 MAX - 50,53 WDEFI - 342,48 HOR	0,96	0,94	0,89	91,73	11%		123,41	15%	1,38
2004	-224,94+ 2357,8581 MAX	0,6	0,58	0,5	126,72	44%		130,5	45%	1,9
2005	-2639,25+ 698,67 INT + 1,91 WDEFI + 106,11 SDA	0,82	0,8	0,77	192,62	27%		205,37	29%	2
2006	-858,74 -686,75 INT + 76,18 CUM + 66,51 EVA - 44,45 SVA	0,87	0,85	0,8	239,53	29%		244	30%	2,41
2007	-1495,72 + 505,03 INT +2,16TWR	0,61	0,57	0,53	223,83	20%		226,29	20%	2,4
2008	-1125,59 + 14,50 PVA	0,59	0,57	0,53	280,26	58%		292,11	60%	1,8
2009	-1551,76+ 3,61 ETAF +31,13 EVA	0,82	0,8	0,75	107,76	22%		119,54	24%	2
2010	-4931,74+ 1057,45 INT +151,27 SDA	0,68	0,66	0,62	623,09	62%		361	36%	1,99
2011	1850,33+ 2329,23 MAX -27,77 WDEFI +6,63 ETAF -89,84 PEA	0,76	0,73	0,65	139,82	29%		154,5	32%	1,92
2012	PVA 720,92 -16,19 INDXLATEST 6,95 WDEFR 19,71	0,78	0,74	0,65	306,42	24%		347,88	27%	2,1

cal : calculé ; aj : ajusté ; val : validation ; DW : Indice de Durbin-Watson

### **4.3.2. Discussion partielle**

Les progrès de la science et des techniques, surtout en statistique, en télédétection et en informatique, ont permis de disposer d'une part de machines de grande capacité de traitement, d'autre part de disposer de séries d'images satellitaires suffisamment longues, permettant de réaliser des études d'estimation des rendements fourragers par régression linéaire multiple avec 34 variables explicatives pour établir un modèle stable et parcimonieux.

Les outils comme VAST, TIMESAT, SPIRITS..., la disponibilité de séries temporelles assez longues d'images d'indice de végétation offrent la possibilité de calculer plusieurs métriques caractérisant le profil de la végétation de la germination à la sénescence. L'utilisation du maximum ou l'intégrale du NDVI de la saison de croissance végétative est insuffisante pour expliquer le niveau de rendement fourrager. Il existe bien d'autres métriques caractéristiques du profil saisonnier de la végétation qui peuvent contribuer à améliorer le modèle d'estimation des rendements fourragers. Ces métriques du profil phénologique de la phase de croissance de la végétation peuvent être associées à d'autres variables agrométéorologiques telles que celles dérivées d'AMS.

L'examen des modèles annuels obtenus par la régression linéaire multiple montre que les variables peuvent provenir de VAST, AMS ou de la méthode de MEIA ou des trois à la fois (tableau 4.47).

Au plan statistique, le modèle global à quatre variables peut être considéré comme bon. En effet, ce modèle est caractérisé par un  $R^2$  ajusté de 0,68, un  $R^2$  de validation de 0,67 et  $3,72 \text{ kg.ha}^{-1}$  de différence RMSE (RMSE calculé et RMSE validation). La qualité de ce modèle est mise en évidence par le diagnostic des résidus (répartition régulière et distribution normale des résidus). Les variables pertinentes du modèle global : MAX (maximum de la végétation tel qu'il est calculé par le MEIA), DRO (petite l'amplitude), PEA (décade du pic de végétation), VER (grande amplitude), sont toutes des métriques

dérivées du NDVI, elles sont liées directement au cycle de la végétation. Il est important de relever que dans le modèle global à une variable, le MAX tel que calculé par MEIA permet d'expliquer déjà 57 % du rendement fourrager, donc les trois autres variables (DRO, VER et PEA) n'apportent en réalité que 10%, mais permettent de réduire l'erreur de 40 kg MS.ha<sup>-1</sup>. La variable du modèle global N°1 du tableau 4.2 est disponible dès que la végétation atteint sa croissance maximale, ce qui signifie que les premiers résultats du modèle global peuvent être obtenus dès la troisième décade du mois d'août ou la première décade du mois de septembre. Après quatre décades, une mise à jour de ces résultats peut être réalisée en utilisant les modèles N°2, 3 et 4 du tableau 4.2. La disponibilité de l'information sur la production fourragère avant la fin du cycle de la végétation permettra aux décideurs de gagner un peu de temps pour anticiper les crises. La régression linéaire multiple réalisée suivant les types de sols de la FAO a permis d'obtenir des modèles caractérisés par des paramètres plus intéressants que ceux issus de l'analyse à l'échelle globale (les R<sup>2</sup> varient de 0,75 à 0,86 ; les RMSE de 148 à 207 kgMS.ha<sup>-1</sup>), la performance des modèles par type de sol est attestée par les R<sup>2</sup> la validation qui varient dans l'intervalle de 0,71 à 0,81.

À l'échelle des faciès, les résultats sont intéressants, car ils indiquent qu'à une échelle plus homogène, il y a une amélioration des R<sup>2</sup> ajustés de 0,70 à 0,90. Aussi, ces résultats renseignent qu'un R<sup>2</sup> élevé ne signifie pas forcément que le modèle est bon. En effet, les faciès ont donné des modèles stables avec des R<sup>2</sup> de validation variant de 0,64 et 0,86, à l'exception du faciès constitué par les arénosols cambiques et Gleysols eutriques, grossier, plat à doucement ondulé de l'écorégion Manga 1 (MA1\_Qc7-1a) nord sahélienne qui a donné un R<sup>2</sup> de validation de 0,46. Malgré, les bons R<sup>2</sup>, il est important de souligner que les résultats des modèles ne peuvent être obtenus, au plus tôt qu'à 4 décades après la date du maximum de végétation (modèles N° 4, 6 et 7 du tableau 4.10), pour les autres modèles, il faut attendre carrément la

fin du cycle de la végétation. Par ailleurs, Il est observé, une variation des variables explicatives selon les faciès, ce qui peut être considéré comme source d'instabilité des modèles.

Les modèles suivant les années ont aussi donné des R<sup>2</sup> ajustés et RMSE qui varient d'une année à l'autre. Cette variation est de 0,57 à 0,94 pour le R<sup>2</sup> ajusté, de 11 % à 44 % pour le RMSE relatif. Ce qui veut dire qu'à ce niveau aussi la stabilité du modèle n'est pas systématique, car les écarts entre les RMSE sont souvent très importants. En plus, une variation des variables explicatives d'une année à l'autre peut être considérée comme une instabilité du modèle. Une analyse des variables explicatives suivant le profil de la campagne peut permettre d'expliquer d'avantage le comportement des modèles.

Selon le Ministère de l'élevage et des industries Animale du Niger, les années 2004, 2008, 2009 et 2011 ont été déficitaires (Annexes 4.4). L'interprétation du tableau4.11 permet d'affirmer, que même en année déficitaire la méthode par régression linéaire multiple peut être performante. En effet, le modèle de 2009 montre un bon R<sup>2</sup> ajusté (0,8) et un RMSE (112 kg), ce qui constitue une amélioration importante par rapport à la régression linéaire simple qui donne un R<sup>2</sup> (0.59) et un RMSE (162 kg).Pour le modèle global à quatre variables (MAX, DRO, PEA, VER), le R<sup>2</sup> est de 0,69 (R<sup>2</sup> ajusté=0,68) et le RMSE est égal à 285,22 kgMS.ha<sup>-1</sup> (RMSE validation de 288,94 kgMS.ha<sup>-1</sup>), contre les valeurs (R<sup>2</sup> = 0,69 et RMSE = 483 kgMS.ha<sup>-1</sup>) trouvés au Sénégal avec un modèle similaire (Diouf et al.2016). La comparaison de ces résultats montre un avantage du modèle réalisé avec les données du Niger par rapport à celui du Sénégal, car le RMSE trouvé au Sénégal est supérieur à celui du Niger d'environ 200 kg, même si pour faire une comparaison plus objective, il faut les RMSE relatifs. Cependant, si on rapporte cette erreur à la superficie totale de la zone pastorale qui est de 350 000 km<sup>2</sup>, on trouve une erreur moyenne annuelle d'environ plus ou moins 10 000 000 tonnes soit environ la consommation d'environ 2 millions d'UBT/9 mois (saison sèche). Peut-on raisonnablement fonder la

planification des interventions d'un pays avec ce niveau d'erreur ? Si non, il est alors nécessaire de continuer d'explorer d'autres pistes d'amélioration du modèle actuel. Les modèles réalisés par faciès et par écorégion donnent des meilleurs résultats, ce qui permet de dire que la stratification est une bonne solution pour améliorer la performance de cette méthode d'estimation des rendements fourragers. Les tableaux 4.5 et 4.10 illustrent les améliorations enregistrées par rapport au modèle global. Ces résultats montrent certes une amélioration par rapport au modèle de régression linéaire simple, mais ils montrent aussi qu'il faut continuer les recherches en explorant d'autres indices tel que le FAPAR dont les fortes corrélations avec la biomasse observée ont été montrées à travers des études récentes de Meroni et al., (2014) ; et de Diouf et al., (2016).

#### **4.4. Conclusion partielle**

Les résultats ont montré que le modèle global est stable. En effet, la différence entre les RMSE calculés ( $282\text{kg MS.ha}^{-1}$ ) et celui de la validation est  $2,72\text{ kg}$ , le  $R^2$  ajusté est de  $0,68$  kg avec des paramètres très significatifs ( $P<, 0001$ ). Au plan statistique, les résultats satisfaisants mettent en évidence une importante amélioration de la modélisation de la production de biomasse fourragère au Niger par rapport au modèle de régression simple, dont le  $R^2$  est de  $0,56$  avec un RMSE de  $367\text{ kg MS. ha}^{-1}$ . La régression linéaire multiple a permis d'améliorer les paramètres du modèle à l'échelle globale. L'étude a montré que la stratification permet d'améliorer la performance des modèles. En effet, les  $R^2$  ajustés sont très élevés à l'échelle des unités spatiales homogènes (faciès, sols et écorégions) par rapport à l'échelle globale. Cependant à l'échelle des années, les modèles sont instables, car les coefficients de détermination ajustés fluctuent ( $R^2_{aj}$ ) de  $0,57$  à  $0,94$ . Le faciès MA1\_Qc7-1a nord-sahélienne indique un  $R^2$  de validation de  $0,42$ , malgré son  $R^2$  ajusté élevé ( $0,70$ ), ce qui signifie qu'un  $R^2$  ajusté élevé n'est pas systématiquement synonyme d'un bon modèle.

## **V. Performance de l'analyse de similarité dans l'estimation des rendements fourrager au Niger**

### **5.1. Introduction**

#### **5.1.1. Contexte**

La zone pastorale du Niger couvre une superficie 35 millions d'hectares (Rhissa, 2010). Le suivi pastoral de cette région est réalisé grâce à un dispositif de collecte de données qui existe depuis plus de vingt ans. La contribution de l'État du Niger à ce dispositif coûte environ 50 millions Fcfa par an. Cette participation ne prend pas en compte les ressources humaines. Ce dispositif étatique reste toujours fragile, du fait des interruptions de travail par le personnel, à l'insécurité résiduelle persistant dans les zones pastorales. En effet, la zone sahélienne traverse une crise sécuritaire (KAS, 2014) empêchant les techniciens et chercheurs de s'y rendre pour faire les observations de terrain. L'apport des satellites d'observations est très intéressant mais reste complémentaire à la composition floristique des parcours, la dynamique des populations végétales, les processus d'érosion éolienne ou hydrique, la fertilité des sols, l'intensité et les modes de pâture... Toutes ces variables ne relèvent pas de la télédétection. Les observations systématiques et mesures au sol doivent se poursuivre, se développer, s'améliorer en tenant compte de leur rôle pour la validation des produits satellites. Ainsi paraît-il impérieux de proposer une alternative viable et durable de suivi et d'estimation de masses herbacée combinant les données d'observation de la terre et celles du terrain. Nous allons tester la méthode de similarité. Elle est, dans son principe, une forme de raisonnement basé sur le Case Based Reasoning ou CBR (Riesbeck et R. C. Schank, 2013; Watson, 1999; Kolodner, 1992.; Kolodner, 1997; Maher et Garza, 1997; Leake, 1996; Aamodt et E. Plaza, 1994.; Allen, 1994; Watson et Marir, 1994) . Le principe est de considérer que deux situations similaires produisent des résultats similaires. Autrement dit, deux profils de végétation similaires

produiront des rendements fourragers similaires. Cette méthode est comparable à celle des analogues utilisée en météorologie pour réaliser, entre autres, la prévision des températures (Lorenz, 1969) ou des précipitations (Fernández-Ferrero et al., 2010; Hamill et Whitaker, 2006 ; Roebber et Bosart, 1998). La méthode de similarité est disponible dans le logiciel SPIRITS qui permet l'identification des années similaires en utilisant trois catégories de mesures de similarité entre deux séries de données : la racine carrée de l'erreur quadratique moyenne (*Root Mean squared error* : RMSE), la déviation absolue à la moyenne (*Mean absolute Deviation* : MAD) et le coefficient de détermination *R square* ( $R^2$ ) tout en indiquant le décalage (*shift*) tolérable en décades. L'objectif du présent chapitre est de tester la méthode de similarité. Il s'agit particulièrement d'utiliser les images NDVI de SPOT VEGETATION et les rendements fourragers du MEIA pour faire l'estimation ; de tester la performance des critères de mesure de similarité comme le RMSE, le MAD et le  $R^2$  ; enfin comparer l'approche par similarité au modèle du MEIA et à celui de la régression linéaire multiple.

L'hypothèse de base de cette approche est de considérer que deux années similaires produisent des rendements équivalents. Aussi, le travail s'articule autour des points suivants :

- La présentation de l'état des connaissances en matière de similarité ;
- La description du matériel et des méthodes ;
- Les résultats et leur discussion ;
- La conclusion ainsi que les perspectives que laisse entrevoir cette étude.

### **5.1.2. Revue de littérature sur la similarité**

La mesure de similarité a initialement été utilisée pour mettre en évidence le degré ou l'importance de la ressemblance ou de la proximité entre deux objets. Plusieurs travaux de recherches en statistique s'appuient sur l'exploration ou l'analyse des données (data mining)

dans des domaines d'application thématique aussi variés que la linguistique, la biologie, l'informatique (Lesot et al., 2009.). Depuis les années 1900 à aujourd'hui, on relève la publication d'un grand nombre de mesures de similarité qui sont appliquées à plusieurs domaines. Plusieurs études ont contribué à faire l'état des connaissances sur les mesures et les coefficients de similarité. Ainsi, concernant les études sur les applications aux données binaires, on note 76 mesures de similarité (Choi et al., 2010), et 22 coefficients de similarité (Cheetham et Hazel, 1969). D'autres études ont conduit à subdiviser les mesures de similarité en deux groupes. Les mesures de similarité que l'auteur a qualifiées de groupe I sont celles qui ne prennent pas en compte le nombre de caractéristiques possédées par aucun des deux objets comparés et celles du groupe II, qui au contraire les prennent en compte (Rifqi, 2010). Il n'y a pas de mesure de similarité universelle s'appliquant à tous les domaines. Mitchell (2010) a subdivisé les mesures de similarité, en mesures globale et locale avec la possibilité de passer d'une échelle à une autre. Malgré, la quantité importante d'indices de similarité étudiés, on constate un faible nombre d'études comparatives de la performance de ces mesures. En biologie cellulaire une étude comparative de 20 coefficients a été faite pour évaluer les performances des différentes conditions d'acquisition des données en matière de formation cellulaire(Yin et Yasuda, 2005). La méthode par similarité est assimilable à la méthode *nearest neighbors* ou méthode des analogues utilisée en météorologie (Berdugo et al., , 2011., Fernández-Ferrero et al., 2010, Bazin et al., 2010, Matulla et al., 2008, Zorita et H. Von Storch, 1999., Tangang et al., 1997, Livezey et al., 1994, Barnett et Preisendorfer, 1978). Elle est surtout utilisée pour la prévision de précipitation en temps réel (Xavier et Goswami, 2007., Barnston A. G. et R. E. Livezey, 1989) soit à très court terme (Berdugo et al., , 2011., Barnett et al., 1978) , soit à moyen terme (Barnston A. G. et al., 1989). Cette méthode est utilisée aussi par l'ACMAD pour les prévisions saisonnières. La limite principale de la méthode par

similarité tout comme celle par analogie réside dans la profondeur historique de la base des données d'archive. Cette profondeur historique ou l'effectif des situations élargissent le champ mais la limite principale reste pour moi la nature des critères de similarité et leur lien fonctionnel avec le processus, ici celui de la production végétale.

Les données agrométéorologiques telles qu'ETP, Water Requirement Satisfaction Index (WRSI), Standardized Precipitation Index (SPI), cumul pluviométrique, sont des indicateurs de cause, utilisés pour le suivi de la végétation et l'estimation des rendements. Ces activités de suivi et prévision des rendements sont réalisées souvent avec des indicateurs d'état comme les indices de végétation (NDVI, LAI, FAPAR, VCI, SNDVI). Pour l'analyse de similarité, entre les indicateurs d'état et les indicateurs des causes, nous avons préféré un indicateur d'état. C'est-à-dire, le NDVI, car ses relations avec la productivité de la végétation ont été largement étudiées.

Il est observé dans les travaux de recherche qu'un bon nombre de mesures de similarité ont déjà fait leur preuve avec le NDVI. C'est le cas de la distance de Mahalanobis qui a été avantageusement utilisée pour quantifier et cartographier la biodiversité (Krishnaswamy et al., 2009). Le carré de l'erreur moyenne (*mean square error* : mse) et la moyenne absolue de l'erreur ( *mean absolute error* : mae) sont considérés comme des mesures de similarité globale avec alignement spatial.  $mse = \sum_k (a_k - b_k)^2 / k$  et  $mae = \sum_k |a_k - b_k| / k$  où  $a_k$  et  $b_k$  sont respectivement les valeurs numériques du  $k$ ième pixel dans les images  $A$  et  $B$ . Elles ne sont utilisées que, quand les images proviennent du même capteur. Le coefficient de corrélation croisée est

$$\rho = \frac{\sum_k a_k b_k}{\sqrt{\sum_k a_k^2 \sum_k b_k^2}}$$

- L'écart absolu moyen (Mean Absolute Deviation : MAD) est considéré comme une mesure d'erreur statistique de prédiction au même titre que l'erreur quadratique moyenne de prédiction (Root Mean Square Prediction Error : RMSPE) (Chang et al., 2011). En

effet, il donne le biais sur l'estimation (Okin et al., 2013). En outre, il permet de comparer la performance prédictive entre deux modèles ou confronter les résultats de simulations à des mesures (Morland et al., 2001, Sano et al., 1998). C'est aussi une mesure de précision d'un modèle (Evrendilek et Gulbeyaz, 2008).

- Le Root Mean squared Error (RMSE) appelé aussi Root Mean Square Déviation (RMSD). La différence individuelle entre une valeur prédictive et une valeur réelle est appelée valeur résiduelle donc le RMSE n'est qu'une agrégation de ces valeurs constituant le pouvoir prédictif du modèle. Il est utilisé aussi pour comparer les résultats de plusieurs approches (Durai et R. Bhadwaj, 2014).
- Le coefficient de détermination  $R^2$  donne le niveau de liaison entre deux variables à travers une relation linéaire. Dans le cadre de cette étude, ces mesures peuvent alors être utilisées pour calculer la similarité entre deux années dans l'hypothèse que deux situations similaires produisent des résultats similaires. En d'autres termes l'hypothèse que deux années similaires donnent des rendements fourragers équivalents a été formulée. Le coefficient de détermination  $R^2$  et le RMSE sont aussi utilisés pour la validation des modèles (Englhart et al., 2011).

## 5.2. Matériel et méthodes

Le logiciel SPIRITS décrit ci-dessous a été utilisé pour analyser la similarité des profils saisonniers de NDVI issues de SPOT VEGETATION de 2001 à 2012, ainsi que les rendements fourragers herbacés annuels collectés par le MEIA sur la même période. Le logiciel SPIRITS permet de détecter les années similaires à l'aide des trois critères : Le RMSE, le MAD et le  $R^2$ .

Des analyses statistiques exploratoires avec bootstrap sont réalisées pour bien caractériser les observations issues de la simulation. Par ailleurs, l'analyse des corrélations paramétriques et non paramétriques permet d'évaluer le niveau de liaison entre les données simulées et les

données observées. Le test de t et le test de Wilcoxon sont ensuite réalisés afin de comparer les moyennes des rendements observés à ceux obtenus par l'analyse de similarité. Les coefficients de corrélation de la similarité sont enfin comparés à ceux du modèle de MEIA et du modèle par régression linéaire Multiple ou modèle de référence (Figure 5.1).

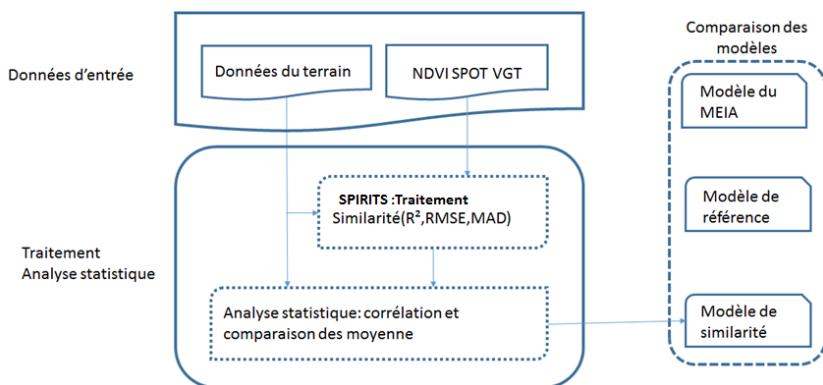


Figure 5.1 : Schéma général de la démarche

### 5.2.1. Généralités sur SPIRITS

Le *Software for the Processing and Interpretation of Remotely Sensed Image Time Series* (SPIRITS) est un logiciel développé par VITO pour le compte du JRC. Dans la perspective d'analyser les séries de données d'Observation de la Terre (OT). La dernière version est publiée en février 2015 sur le site internet du Joint Research Centre JRC<sup>3</sup>. Il contient un large éventail de fonctionnalités permettant d'analyser les séries temporelles d'images satellitaires de basse résolution comme SPOT-VEGETATION, NOAA-AVHRR, METOP-AVHRR, TERRA-MODIS, ENVISAT-MERIS and MSG-SEVIRI. L'algorithme d'analyse de similarité est inclus dans le logiciel SPIRITS (Eerens et Haesen, 2015).

<sup>3</sup> <http://spirits.jrc.ec.europa.eu>

## 5.2.2. Méthodes

La méthode s'articule autour de cinq parties :

- Énoncé du principe de la similarité ;
- Préparation des données ;
- Traitements réalisés avec le logiciel SPIRITS en vue de générer les rendements par similarité ;
- Analyses statistiques exploratoires, tests de corrélation et comparaison des moyennes ;
- Comparaison des trois modèles.

### 2.2.2.1. Principe de la méthode de similarité

Le profil de chaque pixel est réalisé pour la période active de croissance végétative soit dans le cas du sahel, 6 mois allant de mai à octobre (soit 18 décades). Ensuite, une comparaison entre le profil de l'année cible et les profils de l'ensemble des autres années a été réalisée en considérant, soit le  $R^2$  maximum, soit le RMSE minimum ou le MAD minimum, avec un glissement phénologique accepté au maximum de plus ou moins trois décades (Figure 5.2).

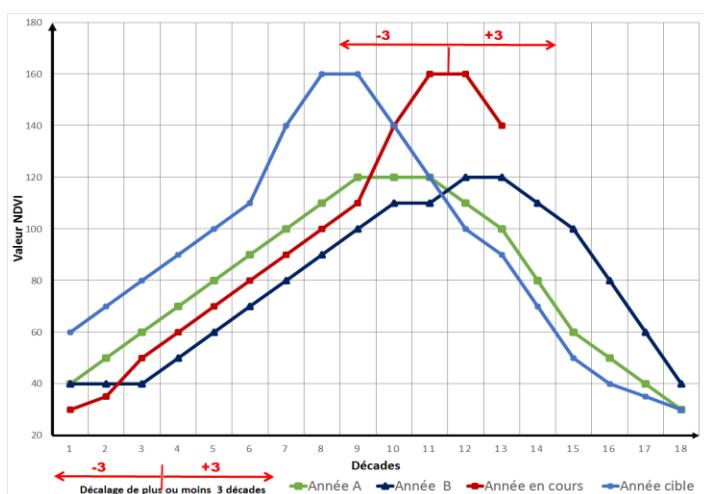


Figure 5.2 : Principe de l'analyse de similarité appliquée à des séries décadiques de NDVI (en code numérique)

### **5.2.2.2. Préparation des données**

La création du masque délimitant les sites constitue une étape préliminaire à l'analyse des données. La seconde étape essentielle consiste à renommer les images NDVI en respectant une nomenclature qui garantit leur identification par le logiciel SPIRITS, et enfin à structurer la table contenant les rendements fourragers mesurés sur le terrain dans un fichier en format texte. Pour faire le traitement avec le module similarité du logiciel SPIRITS sur uniquement les sites de la zone d'étude, il est nécessaire d'utiliser la couche des 68 sites MEIA de relevé in situ de la végétation pour produire le masque. Pour extraire les pixels réellement concernés par les sites, nous avons procédé aussi par les étapes consistant à : générer une grille de carrés de 1 km de coté qui se superpose parfaitement aux pixels des images NDVI de SPOT VEGETATION ; superposer la couche des sites sur celle des grilles pour extraire les neuf polygones contigus de 1 km<sup>2</sup> correspondant à la superficie de chaque site ; faire la fusion des 9 polygones de 1 km<sup>2</sup> pour en faire un seul de 9 km<sup>2</sup> correspondant à la superficie de chaque site ; faire la jointure spatiale pour affecter les attributs de chaque site au polygone correspondant ; donner un code raster à chaque site ; transformer les couches des polygones en raster.

### **5.2.2.3. Traitement des données sur SPIRITS**

Le traitement des données pour la similarité peut être subdivisé en quatre étapes consécutives :

- configuration du projet (Project setting);
- cartographie des années similaires ;
- calcul du rendement fourrager ;
- extraction du tableau contenant les rendements fourragers.

Le module traitement établit une fenêtre permettant de renseigner les différentes données d'entrée : les images multi-annuelles, les informations sur la *temporalité* que sont les périodes couvertes par les

données : la période de croissance végétative qui va de mai à octobre ; les premiers mois et décade ; les derniers mois et décade enfin, le décalage (*shift*) présentant le glissement à prendre en compte dans le calcul. Dans le cadre de cet exercice nous avons pris un décalage maximum de trois décades en considérant qu'un retard ou une avance de trois décades par rapport à la moyenne n'aura pas d'effet sur la production de la masse fourragère au décalage près. L'année similaire est recherchée. Cette démarche permet théoriquement d'augmenter six fois les chances de retrouver une année similaire ; Il faut aussi renseigner la partie procédure qui consiste à donner la mesure de similarité à prendre en considération (RMSE, MAD ou R<sup>2</sup>). Cette technique exige la présence d'au moins 95 % des données dans la série afin de pouvoir démarrer l'analyse des profils.

#### **5.2.2.4. Analyse statistique**

Une analyse exploratoire des observations du terrain et celles générées par la similarité est réalisée pour calculer les moyennes, l'écart type, intervalle de confiance et les biais par bootstrap. Une analyse des corrélations pour calculer les coefficients de Pearson, Kendall et Spearman avec test de signification bilatérale est réalisée pour apprécier le niveau de signification de la relation en les mesures du terrain et les estimations générées par la similarité. La comparaison des moyennes des mesures de terrain et des estimations générées par similarité a été réalisée à l'aide de tests paramétriques et non paramétriques. Il s'agit notamment : du test de t en respectant ses conditions de validité (observations appariées, l'indépendance des observations, l'échantillonnage aléatoire, la distribution normale pour les différences, l'homogénéité de leurs variances), du test de Wilcoxon (Rakotomalala, 2012; Paulson, 2008; Dagnelie, 2013).

#### **Tests de comparaison des moyennes**

Concernant la comparaison des moyennes simulées et les mesures de masse herbacées, nous avons appliqué un test paramétrique (test de t) et un test non paramétrique (test de Wilcoxon). L'objectif de ces tests

est de voir s'il y a une différence significative entre les mesures et les estimations par similarité. S'il n'y a pas de différence significative cela voudrait dire que nous pouvons utiliser les données de la similarité comme proxy des données observées. On fait les deux tests principalement pour deux raisons : la première liée au fait le test de t est plus puissant quand l'échantillon suit une loi normale et que le nombre d'observations est important(Dagnelie, 2013) ; par contre quand la distribution ne suit pas une loi normale ou que le nombre d'observations est faible, il est préférable de privilégier les tests non paramétriques qui sont dans ce cas plus puissants.

### Test de t

Le test de t a été choisi, car les deux séries de rendements fourragers à comparer que sont les observations réelles et celles de la similarité sont en nombre assez important pour appliquer ce test en considérant le théorème central limite. Pour chaque site, il y a la variable issue de la similarité et celle issue des mesures du terrain. La statistique t dans le cas apparié est calculée suivant la formule suivante : $t = \frac{M_d}{SE_d}$  où  $M_d$  est la différence entre les deux moyennes,  $SE_d$  est l'erreur standard de la différence des deux moyennes. Avant ces analyses, les distributions des variables ont été étudiées à travers les analyses des symétries, des aplatissements et des biais (Gilbert, 2004, Dixon et Mood, 1946).

### Test de Wilcoxon

Selon cette méthode, on classe les observations par paires (Rakotomalala, 2008). Ce qui permet d'obtenir un comptage des signes des différences par paires (comme dans le test des Signes) et les rangs de ces différences. Y (+) désigne la somme des rangs des différences positives ; Y (-) désigne la somme des rangs des différences négatives. Le principe est :  $Y (+) + Y (-) = n(n+1)/2$  où n'est le nombre de paires. En moyenne, si les deux échantillons proviennent d'une même population, Y (+) et Y (-) valent tous deux la moitié de cette valeur soit :  $n(n+1)/4$ (Gibbons et S. Chakraborti, 2014; Rakotomalala, 2010)

## **Les corrélations**

Les relations entre les estimations par similarité et les mesures de terrain sont analysées avec différents types de coefficient de corrélation ( $r$  de Pearson,  $\rho$  de Spearman,  $\tau$  de Kendall). La corrélation de Pearson est donnée par le coefficient corrélation  $r$  de Pearson. Elle résume la relation entre 2 variables numériques ainsi que la force de la liaison qui existe entre les variables. La corrélation de Spearman calcule un coefficient de corrélation entre les rangs des valeurs des deux variables, cette corrélation est utilisée lorsque les distributions des variables sont asymétriques (*skewness* en anglais). L'interprétation est identique à celle de la corrélation de Pearson. Pour ce qui est de la corrélation de Kendall, le coefficient de corrélation des rangs de Kendall (Kendall  $\tau$ ) est une mesure de corrélation non paramétrique. Il sert à déterminer la relation qui existe entre deux séries de données (Rakotomalala, 2012; Rakotomalala, 2008).

## **5.3. Résultats et discussion**

### **5.3.1 Résultats**

Les résultats de l'analyse de similarité sont donnés ici, suivant les différents critères de mesure. Ils sont aussi déclinés suivant les différentes échelles spatio-temporelles à savoir l'échelle globale, les écorégions, les zones bioclimatiques et enfin les années. Après les résumés statistiques pour une vue d'ensemble et une caractérisation complète des observations, les résultats des tests de corrélations paramétrique et non paramétrique sont présentés. Les Tableaux de comparaisons des moyennes sont déclinés avant celui de la comparaison des modèles.

#### **5.3.1.1. Analyse globale**

##### **Statistique descriptive et corrélation de la similarité avec $R^2$**

Sur les 319 observations introduites pour la recherche de similarité, nous avons obtenu 153 qui ont trouvé chacune une année similaire. Le Tableau 5.1 montre les résultats de l'analyse statistique descriptive réalisée sur les estimations par similarité avec  $R^2$  et les mesures correspondantes. Les biais et les intervalles de confiance au seuil de 95 % réalisés avec 25 000 sous-échantillons.

Tableau 5.1 : Statistique descriptive de la similarité par utilisation du R<sup>2</sup>

		Statistiques	Erreurs std.	Bootstrap		Intervalle de confiance à 95 %	
				Biais	Erreurs std.	Inférieur	Supérieur
Biomasse	N	153		0,00	0,00		
Similarité (R <sup>2</sup> )	Moyenne	519,59		-0,19	20,48	478,97	559,37
	Ecart type	253,07		-0,95	9,79	234,55	269,47
	Skewness	-0,06	0,20	0,00	0,12	-0,29	0,18
	Kurtosis	-1,09	0,39	0,01	0,11	-1,29	-0,81
Biomasse	N	153		0,00	0,00		
mesurée	Moyenne	584,05		-0,03	36,50	514,85	657,14
	Ecart type	454,95		-3,14	39,21	380,54	522,16
	Skewness	1,47	0,20	-0,04	0,21	1,09	1,74
	Kurtosis	2,71	0,39	-0,15	0,93	1,29	4,05

## **Statistiques descriptives et corrélations des résultats de la similarité avec RMSE**

Sur les 319 observations introduites pour la recherche de similarité, nous avons obtenu 172 ayant enregistré chacune une année similaire. Le tableau 5.2 montre les résultats de l'analyse statistique descriptive réalisée sur les estimations par similarité avec le RMSE et les mesures correspondantes. Les biais et les intervalles de confiance au seuil de 95 % réalisés avec 25 000 sous-échantillons

Tableau 5.2 : Statistiques descriptives de la similarité le RMSE

		Statistiques	Erreur std.	Bootstrap		Intervalle confiance à 95 %	
				Biais	Erreur std.	Inférieur	Supérieur
Biomasse Similarité (RMSE)	N	172		0,00	0,00		
	Moyenne	430,35		-0,51	18,79	393,13	465,46
	Ecart type	249,37		-0,96	10,65	228,85	267,30
	Skewness	0,36	0,19	0,00	0,11	0,15	0,58
	Kurtosis	-0,76	0,37	0,01	0,17	-1,04	-0,36
Biomasse mesurée	N	172		0,00	0,00		
	Moyenne	537,18		-0,04	32,02	475,12	601,12
	Ecart type	415,85		-3,19	38,14	344,35	481,46
	Skewness	1,61	0,19	-0,08	0,29	1,01	1,94
	Kurtosis	3,93	0,37	-0,34	1,33	1,79	5,35

## **Statistiques descriptives et corrélations des résultats de la similarité avec MAD**

Sur les 319 observations introduites pour la recherche de similarité, nous avons obtenu 173 ayant enregistré chacune une année similaire. Le tableau 5.3 montre les résultats de l'analyse statistique descriptive réalisée sur les estimations par similarité avec le MAD et les mesures correspondantes. Les biais et les intervalles de confiance au seuil de 95 % réalisés avec 25 000 sous-échantillons.

Tableau 5.3 : Statistique descriptive de la similarité avec le MAD

			Bootstrap					
			Statistiques	Erreurs std.	Biais	Intervalle de confiance à 95 %		
						Inférieur	Supérieur	
Biomasse	N	173			0,00	173,00	173	
Similarité	Moyenne	439,31			0,06	18,65	402,88	477,49
(MAD)	Ecart type	248,51			-1,13	10,53	225,92	267,81
	Skewness	0,32	0,18		0,00	0,11	0,10	0,54
	Kurtosis	-0,78	0,37		0,01	0,16	-1,06	-0,42
Biomasse	N	173			0,00	173,00	173,00	
mesurée	Moyenne	541,09			-0,83	30,61	483,52	600,32
	Ecart type	396,31			-3,70	32,54	330,98	459,11
	Skewness	1,35	0,18		-0,07	0,27	0,68	1,76
	Kurtosis	2,76	0,37		-0,31	1,14	-0,09	4,58

Le tableau 5.4 consigne les résultats des corrélations de Pearson Kendall et Spearman entre les observations issues de la similarité et les observations mesurées à un seuil de signification de 1 %.

Tableau 5.4 : Corrélations paramétrique et non paramétriques

Similarité	Pearson	Kendall	Spearman
Biomasse Similarité ( $R^2$ )	0,327**	0,262**	0,383**
Biomasse Similarité (RMSE)	0,447**	0,379**	0,521*
Biomasse Similarité (MAD)	0,459**	0,381**	0,540*

\*significatif au seuil de 10 pour 100 ; \*\*significatif au seuil de 5 pour 100 ;

### Tests paramétrique et non paramétrique de comparaison des moyennes à l'échelle globale

La comparaison des moyennes à l'échelle globale montre pour les deux tests qu'il n'y a pas de différence significative entre la moyenne des estimations par la similarité par le  $R^2$  et la moyenne des mesures. Le nombre d'observation s est de 153 avec une différence relative des moyennes de -11 %. Par contre les similarités avec le RMSE et le MAD donnent des différences significatives. Ces deux mesures de similarité sont pratiquement équivalentes en nombre d'observations et en différence relative des moyennes presque égale au double de celle de  $R^2$  (Tableau 5.5).

Tableau 5.5 : Comparaison des moyennes simulées à celles des mesures de masse herbacées à l'échelle globale suivant les tests de t et de Wilcoxon, selon les mesures de similarité

Mesure Sim	Moy mesurée	Moy estimée	Dif absolue	Nb d'obs	Dif Relative %	Test de t	Test wilcoxon
RMSE	537,18	430,35	106,83	172	0,20	0,0003***	0,0075*
R <sup>2</sup>	584,05	519,60	64,45	153	0,11	0,0734	0,7723
MAD	541,10	439,30	101,80	173	0,20	0,0003***	0,0035**

\*significatif au seuil de 1 pour 100 ; \*\*significatif au seuil de 5 pour 1000 ; \*\*\*significatif au seuil de 5 pour 10000

Sim : similarité ; Moy : moyenne ; Dif : Différence ; Nb : nombre ; obs : observations ;

### **5.3.1.2. Tests de t et Wilcoxon selon les années**

La comparaison des moyennes par le test de t et le test de Wilcoxon suivant les années montrent que : chaque année, il y a au moins un test qui indique que la différence des moyennes entre les résultats de la similarité et les données observées n'est pas significative. En ce qui concerne le test paramétrique, l'examen du Tableau 6 montre que pour le R<sup>2</sup>, trois années à savoir 2004, 2007, et 2012 ont donné des différences significatives ; pour le RMSE quatre années que sont 2006, 2007, 2008, 2010 sont significatives ; pour le MAD cinq années à savoir 2006, 2007, 2008, 2010, et 2012. Concernant le test non paramétrique, Il ressort de l'examen de ce même Tableau que pour le R<sup>2</sup>, trois années à savoir 2004, 2005 et 2012 ont donné des différences significatives ; pour le RMSE cinq années que sont 2006, 2007, 2008, 2010, 2012 sont significatives ; pour le MAD quatre années à savoir 2006, 2008, 2010, et 2012 sont significatives. En conclusion, on retrouve chaque année au moins une mesure de similarité donne une différence non significative avec la moyenne observée soit en test de t ou de Wilcoxon. (Tableau 5.6).

Tableau 5.6 : Tests paramétrique et non paramétrique suivant les années

An	Mesure	Moy Similarité	Moy obs	nr obs	Dif Absolute	Dif relative	Test de t	Test Wilcoxon
2001	<b>RMSE</b>	569,9	461,47	17	108,429	0,19	0,2242	0,2842
	<b>R<sup>2</sup></b>	503,79	582	10	78,206	0,15	0,4714	0,4131
	<b>MAD</b>	569,9	491,35	16	78,5467	0,13	0,3405	0,3778
2002	<b>RMSE</b>	340,18	356,21	19	-16,023	-0,04	0,7200	0,5678
	<b>R<sup>2</sup></b>	355,035	464,88	17	-109,85	-0,30	0,1114	0,1415
	<b>MAD</b>	357,365	371,44	17	-14,079	-0,03	0,7591	0,7987
2004	<b>RMSE</b>	260,896	312,86	15	-51,971	-0,19	0,1316	0,1514
	<b>R<sup>2</sup></b>	220,827	382,08	11	-161,26	-0,73	0,0137*	0,0210*
	<b>MAD</b>	280,168	344,5	15	-64,332	-0,22	0,0801	0,1046
2005	<b>RMSE</b>	520,778	497,16	18	23,6109	0,04	0,7931	0,8650
	<b>R<sup>2</sup></b>	613,271	456	17	157,271	0,25	0,1713	0,2462
	<b>MAD</b>	570,373	507	19	63,3735	0,11	0,5028	0,9854
2006	<b>RMSE</b>	700,063	393,71	14	306,349	0,43	0,0429*	0,0353*
	<b>R<sup>2</sup></b>	712,635	576,42	13	136,207	0,19	0,3539	0,8552
	<b>MAD</b>	700,063	391,78	13	308,277	0,44	0,0222*	0,0085*
2007	<b>RMSE</b>	915,376	539	10	376,376	0,41	0,0136*	0,0488*
	<b>R<sup>2</sup></b>	968,589	522,75	7	445,839	0,46	0,0017**	0,0156*
	<b>MAD</b>	934,554	557,55	8	376,999	0,40	0,0140*	0,0547
2008	<b>RMSE</b>	318,963	401,64	17	-82,684	-0,25	0,0224*	0,0348*
	<b>R<sup>2</sup></b>	462,123	551,57	13	-89,448	-0,19	0,3551	0,0906
	<b>MAD</b>	318,963	401,88	17	-82,919	-0,25	0,0235*	0,0267*
2009	<b>RMSE</b>	457,015	355,61	13	101,399	0,22	0,3038	0,3054
	<b>R<sup>2</sup></b>	443,952	465,25	11	-21,298	-0,04	0,8171	1,0000
	<b>MAD</b>	439,878	409,64	14	30,2352	0,06	0,7077	0,6698
2010	<b>RMSE</b>	763,04	502,71	14	260,326	0,34	0,0334*	0,0040***
	<b>R<sup>2</sup></b>	867,291	605,06	15	262,229	0,30	0,1095	0,1167
	<b>MAD</b>	635,492	450,78	14	184,706	0,29	0,0178*	0,0085**
2011	<b>RMSE</b>	440,475	399,27	22	41,2019	0,09	0,3572	0,3021
	<b>R<sup>2</sup></b>	478,425	496,18	15	-17,762	-0,03	0,8025	0,9399
	<b>MAD</b>	448,602	392,52	18	56,0761	0,12	0,2301	0,2253
2012	<b>RMSE</b>	1046,71	561,6	10	485,111	0,46	0,0596	0,0488*
	<b>R<sup>2</sup></b>	989,65	582,72	10	406,923	0,41	0,0442*	0,0420*
	<b>MAD</b>	1048,95	554,83	11	494,115	0,47	0,0214*	0,0161*

\*significatif au seuil de 5 pour 100 ; \*\*significatif au seuil de 1 pour 100 ; \*\*\*significatif au seuil de 5 pour 1000 Sim : similarité ; Moy : moyenne ; Dif : Différence ; Nb : nombre ; obs : observations

### **5.3.1.3. Analyse par écorégion**

Les résultats des tests de t et de Wilcoxon, montrent que le niveau de signification de la différence entre les moyennes varie suivant les écorégions, que ça soit en test paramétrique ou non ainsi que par la mesure de similarité. Concernant le test paramétrique, il ressort de l'examen du Tableau 5.7 que pour la mesure de similarité  $R^2$ , l'Aïr et le Manga2 ont donné des différences significatives ; pour le RMSE, l'Azaouak et le Manga2 ont présenté des différences significatives ; pour le MAD, ce sont l'Azaouak et le Manga2 qui ont présenté des différences significatives. Concernant le test non paramétrique, l'examen du Tableau montre que pour le  $R^2$ , l'Air et Manga2 ont donné des différences significatives ; pour le RMSE seul le Manga2 a donné une différence significative ; pour le MAD les différences sont significatives pour l'Azaouak et le Manga2. En conclusion, pour Ader Doutchi Magia, l'Air, l'Azaouak, le Gourma et le Manga1, on note que les moyennes des résultats obtenus par similarité ne sont pas différentes statistiquement des moyennes observées. Cependant pour le Manga 2 la différence est significative pour tous les tests et pour toutes les mesures de similarité (Tableau 5.7).

Tableau 5.7 : Tests paramétrique et non paramétrique selon les écorégions

	Mesure de sim	Moy observée	Moy sim	Nb obs	Dif absolue	Dif relative	Test de T	Test wilcoxon
ADM	<b>RMSE</b>	422,87	382,71	7	40,16	0,095	0,7232	0,8125
	<b>R<sup>2</sup></b>	354,88	422,80	5	67,91	0,19	0,6162	0,6250
	<b>MAD</b>	422,87	404,29	7	18,59	0,04	0,8842	0,8125
AIR	<b>RMSE</b>	400,45	408,11	9	-7,66	-0,01	0,8274	0,9102
	<b>R<sup>2</sup></b>	460,7	383,62	10	77,08	0,16	0,0379*	0,0488*
	<b>MAD</b>	400,45	408,22	9	-7,77	-0,01	0,8394	0,9102
AZ	<b>RMSE</b>	489,60	411,96	93	77,65	0,16	0,0275*	0,1602
	<b>R<sup>2</sup></b>	575,51	535,13	76	-40,38	-0,07	0,4543	0,9877
	<b>MAD</b>	489,22	416,42	96	72,81	0,15	0,0146*	0,0446*
GR	<b>RMSE</b>	391,65	318,9	10	72,75	0,18	0,5022	0,8457
	<b>R<sup>2</sup></b>	287,48	300,67	6	13,18	0,04	0,8287	0,8438
	<b>MAD</b>	391,65	324	10	67,65	0,17	0,5025	0,8457
MG1	<b>RMSE</b>	562,55	533,91	34	28,64	0,05	0,6235	0,9667
	<b>R<sup>2</sup></b>	547,33	582,58	36	35,25	0,06	0,5354	0,6111
	<b>MAD</b>	563,40	537,11	35	26,29	0,05	0,6610	0,8980
MG2	<b>RMSE</b>	717,99	455,85	14	262,13	0,36	0,0224*	0,0107*
	<b>R<sup>2</sup></b>	681,05	439,13	15	-241,92	-0,35	0,0087**	0,0181*
	<b>MAD</b>	872,57	512	12	360,57	0,41	0,0090**	0,0068**

\*significatif au seuil de 5 pour 100 ; \*\*significatif au seuil de 1 pour 100 ; Sim : similarité ; Moy : moyenne ; Dif : Différence ; Nb : nombre ; obs : observations

#### **5.3.1.4. Analyse par zone bioclimatique**

Le tableau 5.8 consigne les résultats obtenus avec les tests de t et de Wilcoxon selon les zones bioclimatiques. Le niveau de signification de la différence observée entre les moyennes est lié à ces zones et aux critères de mesures de la similarité. Pour le test paramétrique, Il ressort de l'examen de ce Tableau qu'au niveau du  $R^2$ , les différences observées dans les zones sahélienne et saharienne sont significatives. Par contre dans la zone nord-sahélienne elles ne le sont pas ; en ce qui concerne le RMSE et le MAD, les moyennes sont significativement différentes dans les zones nord-sahéliennes et sahéliennes alors qu'elles sont non significatives pour la zone saharienne. Quant au test non paramétrique, l'examen du Tableau montre que pour le  $R^2$ , les moyennes sont significativement différentes dans la zone saharienne ; par contre elles ne le sont pas pour les zones sahélienne et nord sahélienne. S'agissant du RMSE et du MAD, il n'y a pas de différence significative pour ces deux zones. Par contre la différence est significative pour la zone Nord sahélienne.

Tableau 5.8 : Tests paramétrique et non paramétrique par zone bioclimatique

Zones bioclimatiques	Mesure	Moy sim	Moy obs	Nb sim	Dif absolue	Dif relative	Test de T	Test wilcoxon
Nord sahélienne	RMSE	553,24	456,60	133	-96,64	-0,17	0,0022**	0,0415*
	R <sup>2</sup>	593,91	521,13	116	-72,77	-0,12	0,0621	0,2302
	MAD	563,00	468,14	134	-94,86	-0,17	0,0014***	0,0187*
sahélienne	RMSE	1440,42	326,4	5	-1114	-0,77	0,0081**	0,0625
	R <sup>2</sup>	1673,07	548,6	5	-1124,5	-0,67	0,0036***	0,0625
	MAD	1493,22	334,25	4	-1159	-0,77	0,0298*	0,1250
saharienne	RMSE	341,53	342,94	34	1,41	0,00	0,9628	0,9800
	R <sup>2</sup>	378,13	509,46	32	131,33	0,34	0,0105*	0,0189*
	MAD	348,38	340,91	35	-7,47	-0,02	0,8060	0,7122

\*significatif au seuil de 5 pour 100 ; \*\*significatif au seuil de 1 pour 100 ; \*\*\*significatif au seuil de 5 pour 1000

Sim : similarité ; Moy : moyenne ; Dif : Différence ; Nb : nombre ; obs : observations

### **5.3.2. Discussion partielle**

L'analyse statistique descriptive avec bootstrap réalisée sur les données de la similarité utilisant le  $R^2$  montre que l'échantillon issu de la similarité est normalement distribué et que l'asymétrie et l'aplatissement (*skewness, kurtosis*) sont situés dans l'intervalle [-1.96 ; +1.96] autorisant ainsi de procéder à un test paramétrique. Les biais observés étant faibles confortent davantage les résultats obtenus. Par contre, le *kurtosis* des données mesurées s'écarte de cet intervalle ce qui recommande l'utilisation d'un test non paramétrique pour procéder à une comparaison des moyennes. Les corrélations paramétriques et non paramétriques sont toutes significatives au seuil de 0.01, elles sont proches de celles obtenues avec le modèle BIOMASAH, mais inférieures à celles du modèle de MEIA et du Modèle de référence. Sachant qu'à part le  $R^2$ , il y a bien d'autres critères pertinents pour expliquer la performance d'un modèle par rapport à un autre. Il vaut mieux examiner les résultats des tests de comparaison des moyennes avant de se prononcer sur le cas de la présente étude.

Par ailleurs, l'analyse descriptive avec bootstrap des résultats de la similarité par utilisation du RMSE montre que l'asymétrie et l'aplatissement de l'échantillon sont particulièrement bien situés dans un intervalle suggérant de procéder à un test paramétrique, les biais étant tolérables. Les corrélations de Pearson, de Spearman et de Kendall sont toutes significatives comme pour le modèle BIOMASAH.

Quant à l'analyse descriptive sur les résultats de la similarité utilisant le MAD, le *skewness* et le *kurtosis* des observations issues de la similarité sont situés dans l'intervalle autorisant un test paramétrique, les biais sont tolérables. Par contre, le *kurtosis* des données observées est de 2.76 donc assez élevé ce qui suggère un test non paramétrique. Les corrélations de Pearson, de Spearman et de Kendall sont également toutes significatives.

La différence des *Kurtosis* pour les trois mesures de similarité signifie qu'il y a une différence entre les observations constituant les résultats

du traitement selon la mesure de similarité utilisée. L'évidence est donnée par la différence des nombres d'observations obtenues pour R<sup>2</sup>, MAD et RMSE qui sont respectivement 153, 172, 173 ce qui confirme la thèse qui considère le MSE et MAE ont la même performance en analyse de similarité, ils sont sensibles aux observations aberrantes contrairement au coefficient de corrélation (R<sup>2</sup>) qui en est moins sensible(Mitchell, 2010). Ces résultats signifient qu'à l'échelle globale le R<sup>2</sup> est plus sélectif que le MAD et le RMSE. La comparaison des moyennes à l'échelle globale montre que les résultats de la similarité par le R<sup>2</sup> peuvent être utilisés comme proxy des données observées. Ceci est conforté par la différence relative de -11 % entre les moyennes. Par contre, les similarités par le RMSE et le MAD donnent des différences significatives. Ces derniers sont pratiquement équivalents avec une différence relative de 19 % par rapport aux observations mesurées. Par ailleurs la comparaison des moyennes par le test de t et celui de Wilcoxon suivant les années montre qu'il y a au moins chaque année une mesure de similarité qui autorise l'utilisation de la similarité comme proxy aux données observées. Dans 75 % des cas le R<sup>2</sup> permet l'utilisation de ses résultats comme proxy aux données observées. En ce qui concerne le RMSE et MAD, on note que les résultats peuvent être utilisés respectivement dans 66 % et 59 % des cas. Quand on associe les résultats issus des trois mesures de similarité, on trouve que chaque année, on a au moins une possibilité d'utiliser la méthode avec succès. Cependant, il est important de noter que la performance du R<sup>2</sup> est supérieure à celle du RMSE et MAD. L'examen des résultats selon les écorégions montre que le R<sup>2</sup> n'autorise pas l'utilisation de la similarité comme proxy dans l'Aïr et le Manga2. Les mêmes observations sont faites en test paramétrique pour le RMSE et MAD dans l'Azaouak et le Manga2. Concernant le test non paramétrique, il ressort les résultats par le RMSE ne permettent pas l'utilisation des données dans le Manga2 et le MAD dans l'Azaouak et le Manga2. En conclusion, pour l'Ader Doutchi Magia, l'Air, l'Azaouak, le Gourma et

le Manga1, les résultats issus de toutes ces mesures de similarité sont utilisables comme proxy aux données observées. Cependant la méthode n'est pas concluante pour le Manga2. Il faudra alors chercher d'autres mesures de similarité ou d'autres types d'indices pour le Manga2. Selon le test de t pour le  $R^2$ , on ne peut pas utiliser la similarité dans les zones sahélienne et saharienne. Par contre, elle est utilisable dans la zone nord-sahélienne ce qui est une bonne information, car cette zone contient 75 % des sites. Le RMSE et le MAD ne sont pas utilisables pour la zone nord-sahéliennes et sahélienne. Par contre, ils sont utilisables dans la zone saharienne ce qui donne une certaine complémentarité entre les mesures ; le test Wilcoxon montre que pour le  $R^2$  ne peut pas être utilisé dans la zone saharienne, mais utilisable dans les zones Sahélienne et nord-sahélienne. Les mêmes cas sont observés pour le RMSE et le MAD. Au regard de ces premiers résultats, cette approche complexe, avec des variables imparfaitement indépendantes entre elles, on risque de fournir des résultats difficilement interprétables. Les données doivent être exploitées avec un peu de prudence, car le nombre d'observations en moyenne est d'environ 13 pour le  $R^2$ , de 15 pour le MAD et le RMSE ce qui n'est pas une quantité très importante de données pour un tel exercice surtout en considérant l'étendue de la zone pastorale du Niger. Il s'avère alors nécessaire de continuer les recherches. En explorant d'autres indices de végétation et données agrométéorologiques et d'autres mesures de similarité. On peut par exemple explorer le FAPAR, car une étude réalisée au sahel avec les données du Sénégal a montré des corrélations moyennes r de Pearson entre la biomasse mesurée au sol et le cumul FAPAR dans l'ensemble du sahel et sur la partie pastorale qui sont respectivement de 0,78 et 0.75 (Meroni et al., 2014.). Par ailleurs une relation évidente a été trouvée entre le NDVI — GPP et NDVI-FAPAR respectivement 0.72 et 0.79 (Wang et al., 2004). Nous avons également procédé à une comparaison des performances de l'analyse de similarité à celles du modèle du MEIA et de la méthode la référence en nous

basant sur le R<sup>2</sup> et le RMSE, car le RMSE (Chai et Draxler, 2014) et le R<sup>2</sup> (Cornillon et al., 2007) sont utilisés pour qualifier un modèle. L'examen du RMSE et le R<sup>2</sup> montrent que le modèle de référence est le meilleur, il est suivi par le modèle MEIA. Il est néanmoins important de souligner que le R<sup>2</sup> explique seulement comment le modèle ajuste les données observées, mais n'explique pas le pouvoir prédictif d'un modèle (Trendowicz et Jeffery, 2014) et que le RMSE est très sensible aux données aberrantes (Chai et Draxler, 2014).

#### **5.4. Conclusion partielle**

L'étude a été réalisée en utilisant les images NDVI issues de SPOT VEGETATION et les rendements fourragers du MEIA collectés sur 68 sites répartis dans les zones pastorales et agropastorales du Niger. Le logiciel SPIRITS a été utilisé pour faire le calcul de similarité entre profil saisonniers du NDVI de mai à octobre en utilisant le R<sup>2</sup>, le RMSE et MAD comme mesure de la similarité. Les résultats ont permis de tester la performance de ces trois mesures à travers la comparaison des moyennes des résultats issus de la simulation et les rendements fourragers mesurés sur le terrain. À l'échelle globale, les résultats indiquent que le R<sup>2</sup> a été plus performant que le RMSE et MAD qui ont quasiment les mêmes performances. Les résultats de la similarité calculée avec R<sup>2</sup> peuvent être utilisés comme proxy à la phytomasse herbacée mesurée *in situ*, car il n'y a pas de différence significative entre la moyenne simulée et la moyenne mesurée au seuil de 1 %. Par contre, les résultats de la similarité calculée avec le RMSE et le MAD ne sont pas utilisables. Les corrélations paramétriques et non paramétriques sont toutes significatives au seuil de 1 %. Cependant, les R<sup>2</sup> sont faibles, ils varient entre 0,32 et 0,45. Donc, Il est nécessaire de poursuivre les recherches, car il y a beaucoup d'études qui ont montré des très bonnes liaisons entre certains indices comme le FAPAR, l'EVI et le LAI et la masse aérienne.

## **VI. Comparaison des modèles**

L'objectif de ce chapitre est de comparer les résultats des trois méthodes d'estimation de la biomasse herbacée pour déterminer, le modèle qui peut être considéré comme le plus performant. Il s'agit de comparer les résultats du modèle par régression linéaire simple traité dans le chapitre III aux résultats du modèle réalisé par régression linéaire multiple traité dans le chapitre IV et les résultats de la similarité relative au chapitre V.

### **6.1. Méthode de la comparaison de modèles**

Dans la littérature scientifique, on note plusieurs critères pour comparer la performance des modèles. On peut citer entre autres : le BIC (Bayesian Information Criterion) ; le AIC (Akaike Information Criterion) ; Mallow Cp ; le RMSE, le R<sup>2</sup> ; R<sup>2</sup> aj (Hurvich et C.-L. Tsai, 1989, Bolboaca et Jäntschi, 2013, Burnham et al., 2011, Johnson et Omland, 2004, Burnham et Anderson, 2002, Zucchini, 2000, McQuarrie et Tsai, 1998, McQuarrie et al., 1997, Anderson et al., 1994, Bozdogan, 1987, San Martini et Spezzaferri, 1984). Nous avons utilisé le RMSE, le R<sup>2</sup> et le R<sup>2</sup> ajusté pour comparer les résultats obtenus par analyse de similarité, le modèle du MEIA et le Modèle par Régression linéaire Multiple (MRM). Cette comparaison n'a pas concerné le modèle BIOMASAH qui a montré ses limites en donnant une différence significative lorsque du test de comparaison des moyennes qui a été réalisé dans le chapitre II de la thèse (biomasse herbacée simulée avec BIOMASAH par rapport à la masse herbacée mesurée). Cette comparaison des trois modèles permettra non seulement de montrer la contribution de ces travaux de recherche et aussi de donner une synthèse dans un tableau pour permettre d'apprécier les résultats atteints au bout de ces années de recherches en zone sahélienne en général et au Niger en particulier.

### **6.2. Résultats de la comparaison des modèles**

À l'échelle globale, l'examen des RMSE des trois modèles à savoir le modèle du MEIA, le modèle régression linéaire multiple (MRM) et

l'analyse de similarité montre que le Modèle de régression linéaire multiple (MRM) est le meilleur (RMSE en kg.ha<sup>-1</sup>) , le modèle du MEIA et la similarité sont équivalents (Tableau 6.1). En considérant le R<sup>2</sup>, on constate toujours que le MRM confirme sa supériorité (R<sup>2</sup>=0.69), il est suivi par contre par le modèle du MEIA.

Tableau 6.1 : Comparaison du R<sup>2</sup> et du RMSE des trois modèles à l'échelle globale

Modèle	MEIA		MRM		SIM	
	Mesure	R <sup>2</sup>	RMSE	R <sup>2</sup>	RMSE	R <sup>2</sup>
Valeur	0,56	367	(kg.ha <sup>-1</sup> )	0,69	285	(kg.ha <sup>-1</sup> )

Il ressort de la comparaison des trois modèles suivant les années que les R<sup>2</sup> et les RMSE montrent toujours la prédominance du modèle MRM sur les deux autres. Selon le R<sup>2</sup>, le modèle MEIA est meilleur que celui de la similarité. Par contre, en considérant le RMSE, ces deux modèles sont équivalents (Tableau 6.2).

Au niveau de l'analyse de similarité, le R<sup>2</sup> est beaucoup plus variable et faible d'une année à l'autre que pour les deux autres méthodes. Ce constat peut s'expliquer par le fait que l'analyse de la similarité est très dépendante de la profondeur historique de la base de données. En effet, dans la base de données, sur les 12 ans d'observations, les sites présentent des données manquantes, ce qui réduit les chances d'avoir des années similaires.

Cela peut constituer une raison suffisante pour exclure la méthode de similarité en attendant d'avoir une série plus longue. Cependant, l'analyse par similarité présente un potentiel car, elle a donné le meilleur résultat pour l'année 2008 (R<sup>2</sup>=0,71) contre R<sup>2</sup> =0,52 et R<sup>2</sup>= 0,58 respectivement pour la méthode du MEIA et la MRM.

Tableau 6.2 : Comparaison du R<sup>2</sup> et du RMSE des trois modèles selon les années

Année	MEIA		MRM			SIM	
	R <sup>2</sup>	RMSE	R <sup>2</sup>	R <sup>2</sup> aj	RMSE	R <sup>2</sup>	RMSE
2001	0,56	384	0,83	0,79	178	0,2	323
2002	0,74	137	0,9	0,88	95	0,56	191
2003	0,62	253	0,96	0,94	100	###	###
2004	0,58	126	0,59	0,57	124	0,59	118
2005	0,72	233	0,81	0,79	194	0,07	401
2006	0,63	351	0,87	0,84	219	0,52	415
2007	0,54	233	0,61	0,57	100	0,12	219
2008	0,52	299	0,58	0,57	283	0,71	136
2009	0,59	162	0,82	0,79	112	0,02	274
2010	0,54	455	0,82	0,79	348	0,45	412
2011	0,65	164	0,76	0,73	138	0,34	185
2012	0,56	410	0,77	0,74	308	0,003	552

## VII. Discussion générale

Nous avons d'abord testé le modèle BIOMASAH qui a été développé par le CRA. C'est un outil qui continue d'être utilisé par certains pays par manque de modèles alternatifs fiables, leur permettant d'évaluer la biomasse fourragère en vue d'aider aux actions de planification en faveur des éleveurs. C'est un modèle basé sur le bilan hydrique qui n'a jamais fait l'objet d'une validation sur une période de 12 ans. La validation de ce modèle qui a été réalisée à travers des tests paramétriques et non paramétriques de comparaison de moyennes a montré qu'il y a des différences importantes entre la moyenne de la biomasse herbacée mesurée et celle de la biomasse herbacée simulée

avec BIOMASAH. À l'échelle globale, Les tests ont montré que la moyenne du modèle fait quasiment le double de la moyenne mesurée (983,17 contre 591,17), mais cette grande différence ne provient pas seulement de l'imperfection du modèle. En effet, l'examen de la Figure 1.11 qui compare le cadre théorique établi par Breman H. (1982), et la réalité montrée par les données du MEIA, indique que les données observées sont en général sous-estimées. Au niveau des isohyètes 250 mm et 350 mm, la sous-estimation oscille respectivement dans les intervalles  $100\text{-}250 \text{ kg.MS.ha}^{-1}$  et  $350 \text{ à } 600 \text{ kg MS.ha}^{-1}$ . Cette tendance à la sous-estimation du rendement fourrager peut s'expliquer par les retards de mesure (maximum de la végétation comme période optimale) et par l'impact éventuel de la pâture des animaux avant ces mesures étant donné que les zones utilisées comme références dans le modèle n'étaient pas préalablement mises en défens. À cette faiblesse du système de mesure des données au sol, il faut ajouter les limites du modèle. En effet, BIOMASAH ne prend pas en compte la distribution temporelle des pluies. Le sahel, à l'instar des autres régions du monde fait face aux effets des changements climatiques, qui se manifestent par une forte fréquence des événements extrêmes tels que les fortes pluies, les inondations et la sécheresse. Il y a aussi, la variabilité intra saisonnière des pluies qui influence la croissance de la végétation. En effet, une étude menée par Cissé et al. (2016), au Sénégal, dans le bassin du Ferlo a mis en évidence les relations entre la pluie, l'humidité du sol, et la croissance de la végétation. Les résultats de cette étude ont mis en évidence la sensibilité de la croissance de la végétation aux précipitations suivant les types de sols. Aussi, l'étude a montré que la végétation met environ deux à trois semaines pour réagir aux anomalies de précipitation. Une séquence sèche d'au moins 7 jours peut déjà affecter significativement la croissance végétative. Alors, du fait de la mauvaise distribution temporelle des pluies, une année peut enregistrer un cumul de pluie excédentaire par rapport à la moyenne pluviométrique et donner une production de biomasse fourragère

inférieure à la moyenne de la masse fourragère de la période de référence.

Le Niger (MEIA) et le Sénégal (CSE) sont les deux pays qui utilisent chaque année un modèle de régression simple combinant à travers la méthode des moindres carrés, les mesures masse fourragère herbacée *in situ* et le NDVI pour évaluer la biomasse fourragère en fin de saison. Le modèle du MEIA tire son origine dans les travaux menés au Sénégal et au Niger (Diallo et al., 1991, Wylie et al., 1991). Ce sont des méthodes anciennes d'une quarantaine d'années qui sont améliorables grâce aux progrès scientifiques et technologiques. Le test de la performance de ce modèle démontre que ce modèle n'est pas stable. En effet, il donne en moyenne sur la période de 2001 à 2012 et à l'échelle globale un  $R^2$  de 0,56 avec un RMSE de l'ordre de 367 Kg.ha<sup>-1</sup>. Mais la même approche prise année par année donne des résultats assez disparates. Ce qui montre que la relation entre le NDVI et la masse fourragère herbacée n'est pas toujours une fonction linéaire. En effet, Santin-Janin et al., 2009 et Bégué et al., 2011 ont montré que la relation est aussi non linéaire. Aussi, une étude récente réalisée au Sénégal par Diouf et al. 2014, à l'aide des données collectées par le Centre de Suivi Écologique (CSE), a montré que les fonctions exponentielle et puissance sont plus performantes que les fonctions linéaires. En réalité cette relation est beaucoup plus complexe car, elle peut être influencée par l'éclairement au moment du passage du satellite, la réflectance respective de chaque espèce fourragère, la réflectance spécifique du sol, les conditions atmosphériques (poussières, vapeur d'eau...), la durée réelle de la saison de végétation, etc...

La connaissance de l'information fiable sur la disponibilité de la biomasse fourragère est d'une importance capitale pour la gestion des ressources naturelles. En effet, les communautés pastorales, aux niveaux local, national et régional, attendent chaque année la publication de cette information pour planifier leurs actions. Donc, il n'est pas raisonnable, de baser la prise de décision sur un modèle qui

fait une erreur relative de 52%. Il était alors nécessaire d'avoir une méthode plus fiable qui minimise considérablement ces erreurs d'estimation.

La régression linéaire multiple a permis de réaliser des modèles liant des variables agrométéorologiques dérivées d'AMS et de variables biophysiques dérivées de VAST. La disponibilité d'une série de mesures au sol et d'images NDVI permet de tester la régression linéaire multiple qui a fait ses preuves ailleurs. En effet, cette méthode de régression multiple a montré des performances satisfaisantes pour l'estimation de rendements agricoles (Balaghi et al. 2008 ; Kouadio, 2012 ; Kouadio et al., 2014). Les résultats très intéressants que nous avons obtenus avec les métriques dérivées de AGROMET SHELL et VAST confirment les conclusions des travaux de Diouf et al., en 2015 ; Diouf et al. 2016 qui utilisent d'une part, les mesures de biomasse *in situ* et les paramètres phénologiques dérivés de la série de FAPAR provenant de SPOT VEGETATION ; d'autre part, les mêmes métriques associées à d'autres paramètres agrométéorologiques dérivés de GeoWRSI. Une amélioration importante a été trouvée par rapport à la régression linéaire simple, aussi bien au Sénégal qu'au Niger.

La mesure de la similarité est semblable à la méthode des analogues utilisée par les météorologistes pour faire la prévision du temps (Berdugo et al., 2011., Fernández-Ferrero et al., 2010, Bazin et al., 2010, Matulla et al., 2008, Zorita et al., 1999., Tangang et al., 1997, Livezey et al., 1994, Barnett et al., 1978). Cette méthode est limitée par la profondeur historique de l'archive. Les tests de comparaison de moyennes des résultats de la similarité avec les données de terrain ne sont pas significatifs au seuil de 1%, c'est-à-dire que les moyennes du terrain ne sont pas statistiquement différentes de celles issues de la similarité. Les corrélations paramétriques et non paramétriques sont aussi significatives au même seuil (1%). La comparaison de ces méthodes montre que le Modèle de Régression linéaire Multiple (MRM) est le meilleur car il est plus stable. En effet, sur la base des

RMSE et des R<sup>2</sup> le modèle par régression linéaire multiple s'est montré plus performant que le modèle du MEIA et l'analyse de la similarité, aussi bien à l'échelle globale que suivant les années (tableaux 6.1 et 6.2). Cette conclusion montre clairement que le MRM constitue l'outil alternatif pour l'estimation de la production fourragère, en vue de l'établissement d'un bilan fourrager prévisionnel au Niger. Nous pouvons ainsi affirmer que l'étude a contribué à l'amélioration des méthodes d'estimation des rendements fourragers au Niger. Aussi, elle a permis de : faire le point sur les méthodes utilisées pour évaluer la ressource fourragère annuelle ; comparer et confronter les résultats des méthodes (MEIA, MRM, et similarité) ; et enfin de proposer des pistes pour l'amélioration. Cependant, il est nécessaire de poursuivre la recherche avec d'autres indices de végétation et données agrométéorologiques qui ont montré leur pertinence ailleurs.

### **Exploitation opérationnelle des résultats**

Dans un contexte de crise sécuritaire et de défis climatiques, les résultats de ces travaux de thèse contribueront à améliorer la qualité des informations produites annuellement sur la production fourragère pour aider à la prise de décision en vue d'anticiper les crises pastorales. Un Système d'Alerte Précoce Pastorale (SAPP) efficace et efficient nécessite de faire une évaluation fiable et à temps de la production fourragère. La fiabilité du bilan fourrager prévisionnel dépend entre autres de la qualité des résultats de l'évaluation de la production fourragère et de la fiabilité des chiffres sur l'effectif du cheptel. Les variables explicatives du modèle global sont toutes dérivées du NDVI : la valeur maximale du NDVI (MAX) ; la valeur PVAL – EVAL (DRO), la décade au maximum de la végétation (PEA), et l'amplitude (VER). Ce sont des variables simples à produire avec le programme VAST, ce qui facilitera l'adoption du modèle par le Ministère. L'utilisation pratique du Modèle de régression linéaire Multiple (MRM) n'est pas très différente de celle de la Régression Linéaire simple. La seule

différence est que l'équation n'est plus linéaire à une seule variable explicative , mais plutôt linéaire à plusieurs variables. La carte de la masse fourragère peut être réalisée avec cette équation à 4 variables explicatives, un mois après la PEAK (décade du maximum NDVI). Pour améliorer l'estimation de la production fourragère par unité administrative, en plus de la stratification par faciès, il est possible d'envisager l'utilisation de la carte d'utilisation / d'occupation des sols produite par l'USGS et AGRHYMET comme masque pour estimer la production des zones uniquement réservées à l'exploitation pastorale. La méthode de régression linéaire multiple appliquée aux rendements agricoles ayant donné des résultats intéressants ailleurs, peut être utilisée pour calculer les rendements en grain des principales cultures en vue d'en déduire les résidus de récoltes qui jouent un rôle important dans l'alimentation du bétail.

Une extension « plugin » sur le logiciel libre Quantum GIS (QGIS) peut être envisagée. Cette extension peut aider à rendre plus opérationnelle l'exploitation des résultats de cette thèse et ceux des études antérieures obtenus ailleurs dans le domaine agricole pour réaliser la prévision des rendements agricoles. La carte de l'utilisation/occupation des sols (LU/LC) ou toute autre carte de meilleure qualité est d'une grande utilité pour servir de masque de calcul. Le MRM sera appliqué pour l'estimation des rendements fourragers dans les zones pastorales. Pour prendre en compte la production en résidus de récoltes, il serait intéressant d'envisager l'utilisation de la régression linéaire multiple qui a montré ailleurs sa performance dans les travaux de Balaghi et al. (2008), Kouadio (2012), Kouadio et al. (2014) en matière de prévision des rendements agricoles, en vue d'en déduire la production en résidus de récoltes. Le résultat sera une carte contenant la production fourragère des zones pastorales et celle des résidus de récoltes des zones cultivées.

### **VIII. Conclusion générale et perspectives**

En somme, cette étude contribue à l'amélioration des méthodes d'estimation des rendements fourragers au sahel en général et au Niger en particulier. En effet, la comparaison des moyennes réalisée entre les données de biomasse herbacées mesurées sur le terrain et les estimations à l'aide du modèle BIOMASAH montre une différence significative. Le modèle BIOMASAH surestime la masse herbacée, cela est attesté par les tests paramétriques (test de t) et non paramétriques (de Wilcoxon et des signes), même si les conditions de mesure des données observées présentaient de limites qui favorisent cette conclusion. La corrélation de Pearson donne un coefficient qui ne dépasse guère 0,15, mais, tout en étant hautement significative, telle qu'attestée par les corrélations  $\rho$  de Spearman,  $\tau$  de Kendal, et D de Hoeffding). Néanmoins, il est nécessaire de poursuivre les recherches pour améliorer ce modèle notamment il faudrait pouvoir prendre en compte les séquences sèches, le bilan hydrique qui doit être à un pas de temps journalier, à la limite décadaire, les pressions de pâture et aussi les ligneux. La disponibilité de séries assez longues d'images d'indice de végétation et d'autres données agrométéorologiques ouvre de nouvelles perspectives de prise en compte de l'évolution décadaire de la végétation et de la phénologie des pâturages.

Les résultats de l'analyse de la performance du modèle MEIA indiquent une fluctuation des paramètres d'une à année à l'autre. Ce qui est attesté par la variation du  $R^2$  entre 0,52 et 0,73 suivant les années. Par contre, les corrélations restent très significatives ( $P < 0,001$ ). Aussi, il ressort que le coefficient de détermination est plus élevé dans les zones sahariennes et nord-sahélienne que dans la zone sahélienne typique. Les RMSE annuelles sur l'ensemble de la zone pastorale varient entre 120 et 460  $\text{KgMS.ha}^{-1}$ . Les tests non paramétriques de comparaison de moyenne (Test de Wilcoxon et des signes) montrent qu'il n'y a pas de différence significative à utiliser SPOT VEGETATION ou eMODIS (intégrale ou maximum). Ce qui permet aux utilisateurs de continuer

avec les images eMODIS compte tenu de la fin de vie du satellite SPOT VEGETATION. Néanmoins, le modèle nécessite des améliorations, car l'erreur quadratique moyenne sur la production est énorme.

Il est proposé un modèle de régression linéaire multiple qui utilise des métriques des profils saisonniers de NDVI de SPOT VEGETATION et les rendements fourragers herbacés collectés par le MEIA sur 68 sites répartis dans les zones pastorales et agropastorales du Niger. Les résultats montrent que le modèle global est stable, par exemple la variation entre les RMSE calculés ( $282\text{kg.ha}^{-1}$ ) et celui de la validation est 2,72 kg, le  $R^2$  ajusté de 0,68 avec des paramètres très significatifs ( $P < 0,001$ ). Ce qui constitue une amélioration importante par rapport au modèle de régression linéaire simple utilisé par le MEIA dont le  $R^2$  est de 0,56 avec un RMSE de  $367 \text{ kg MS.ha}^{-1}$ .

Les modèles réalisés par Régression linéaire Multiple suivant les types de sols FAO indiquent des paramètres très intéressants comme l'attestent les  $R^2$ : Ge5-1a ( $R^2=0,85$ ) ; Qc1 ( $R^2=0,86$ ) ; Qc7-1a ( $R^2=0,75$ ) ; Ql1-1a (0,75) ; Re35-a ( $R^2=0,76$ ). Les paramètres des Modèles par faciès sont plus intéressants, car les  $R^2$  varient entre 0,77 et 0,93. Ces résultats permettent d'affirmer que la stratification permet d'améliorer la modélisation des productions fourragères. La qualité des paramètres des modèles augmente avec le niveau d'homogénéité des strates sur lesquelles l'analyse est faite. Les  $R^2$  ajustés sont très élevés dans les faciès, ce qui veut dire qu'il serait plus intéressant pour un projet qui travaille à ces échelles d'utiliser les modèles sur les faciès correspondants. Cependant, il est toujours important de bien analyser les modèles (validation croisée, analyse des résidus) avant de tirer toute conclusion sur sa qualité, car un  $R^2$  ajusté élevé ne signifie pas forcément que le modèle est stable. En effet, le faciès MA1\_Qc7-1a nord-sahélienne a donné un bon  $R^2$  ajusté de 0,70, mais un faible  $R^2$  de validation de 0,42.

Selon les années les modèles de régression linéaire multiple fluctuent comme l'attestent les variations des  $R^2$  ajustés de 0,94 et 0,57. Ces

résultats sont nettement plus intéressants que ceux de la régression linéaire simple, mais les modèles sont instables, car les variables extraites de l'analyse statistique fluctuent d'une année à l'autre. D'où la nécessité de poursuivre les recherches en explorant les indices comme le LAI, l'EVI et le FAPAR et aussi le calage des métriques de NDVI sur la phénologie, la prise en compte des ligneux et celle de la pâture pour voir s'il est possible d'améliorer la prévision.

En ce qui concerne l'approche par la mesure de la similarité des profils saisonniers de NDVI, les résultats permettent de tester les performances du R<sup>2</sup>, du RMSE et du MAD à travers la comparaison des moyennes de la simulation et les rendements fourragers observés . L'analyse de résultats à l'échelle globale montre que le R<sup>2</sup> a été plus performant que le RMSE et MAD qui ont quasiment les mêmes performances. À cette même échelle, il n'y a pas de différence significative entre les moyennes de l'estimation par similarité R<sup>2</sup> et les mesures au seuil de 1 % ce qui permet de dire qu'il est possible d'utiliser ces résultats comme approximation des masses herbacées mesurées. Pour ce qui est des corrélations paramétriques et non paramétriques, elles sont toutes significatives au seuil de 1 %. L'analyse des résultats selon les années et les écorégions montrent que ces mesures de similarité sont complémentaires.

La comparaison des résultats des trois modèles à savoir : le modèle du MEIA, le Modèle par Régression linéaire multiple (MRM) et la similarité a montré que le MRM est le meilleur, il apporte une nette amélioration par rapport à la méthode appliquée par le Ministère de l'Élevage et des Industries Animales du Niger. Le MRM peut alors être considéré comme le modèle de référence. Cependant, il est nécessaire de poursuivre les recherches sur d'autres types de données biophysiques et agrométéorologiques, car il y a eu beaucoup d'études qui ont montré des très bonnes liaisons entre certains indices comme FAPAR, EVI, le LAI et la biomasse observée. Aussi, il est nécessaire d'entreprendre des recherches pour examiner les possibilités

d'application du MRM dans l'ensemble du sahel. Toutefois, il est important de souligner que les strates représentent seulement la diversité dans la zone pastorale du Niger et que le manque de données terrain dans les autres zones sahéliennes peut constituer une contrainte à l'application de la méthode.

## **IIX. Références Bibliographiques**

- Aamodt A. et Plaza E. (1994).** Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI communications*, 7(1) : 39-59.
- AGRHYMET (2001).** *Les aptitudes agricoles et pastorales des sols dans les pays du CILSS*. Centre Régional AGRHYMET (CRA), Niamey, NIGER, 173 pages.
- AGRHYMET (2002).** *Atlas agro climatique du Niger*. n° 240, Centre Régional AGRHYMET (CRA), Niamey, NIGER, 136 pages.
- AGRHYMET (1992).** *Les applications de la télédétection au suivi et à la gestion des ressources pastorales au Sahel*. Centre Régional AGRHYMET (CRA), NIAMEY, NIGER, 15 pages.
- Aguiar L. A. A. (2009).** *Impact de la variabilité climatique récente sur les écosystèmes des Niayes du Sénégal entre 1950 et 2004*. Thèse de doctorat: Université du Québec à Montréal, 208 pages.
- Allen B. P. (1994).** Case-based reasoning: Business applications. *Communications of the ACM*, 37(3): 40-42.
- Allen R. G., H. Hurlbert L. S., Pereira M., R. D. and Smith M. (1998)**). "Crop evapotranspiration - guidelines for computing crop water requirements." Food and Agricultural Organization of the United Nations (FAO) Irrigation and Drain. Paper No. 56, Rome. Pp 300
- AMMA (2002). La Mousson Ouest Africaine et ses composantes, Livre blanc, 69 pages
- Anderson D. R., Burnham K. P. et White G. C. (1994).** AIC model selection in overdispersed capture-recapture data. *Ecology*, 75(6) : 1780-1793.
- Anderson D. R., Dennis J. Sweeney, Thomas A. Williams (2011).** Essentials of Statistics for Business and Economics, 6<sup>e</sup> éd, South-Western: Cengage Learning, 725 pages.
- Andrea D. v., Patrizio V. et Djaby B. (2002).** *Les crises alimentaires et les systèmes de prévision au Sahel*. Centre Régional AGRHYMET (CRA), Niamey (NIGER), 8 pages.
- Aubréville A. (1949).** *Contribution à la paléohistoire des forêts de l'Afrique tropicale*. Paris: Société d'Éditions géographiques, maritimes et coloniales, 98 pages .
- Badi H. B. (2011).** Econometrics, Fifth Edition Springer, Verlag Berlin Heidelberg, 415 pages.

- Barnett T. P. et Preisendorfer R. W. (1978).** Multifield analog prediction of short-term climate fluctuations using a climate state vector. *Journal of the Atmospheric Sciences*, 35(10): 1771-1787.
- Balaghi R, Tychon B, Eerens H, Jlibene M (2008).** Empirical regression models using NDVI, rainfall and temperature data for the early prediction of wheat grain yields in Morocco. *International Journal of Applied Earth Observation and Geoinformation* 10: 438–452
- Barnston A. G. et Livezey R. E. (1989).** An Operational Multifield Analog/Anti-Analog Prediction System for United States Seasonal Temperatures. Part II : Spring, Summer, Fall and Intermediate 3-Month Period Experiments. *Journal of Climate*, 2(6) : 513-541.
- Bazin E., K. J. Dawson et Beaumont M. A. (2010).** Likelihood-free inference of population structure and local adaptation in a Bayesian hierarchical model. *Genetics*, 185(2): 587-602.
- Beck P. S. A., Jönsson P., Høgda K. A., Karlsen S. R., Eklundh L. et Skidmore A. K. (2007).** A ground-validated NDVI dataset for monitoring vegetation dynamics and mapping phenology in Fennoscandia and the Kola peninsula. *International Journal of Remote Sensing*, 28(19): 4311-4330.
- Efron B.** (1979). Bootstrap Methods: Another Look at the Jackknife , *Annals of Statistics*, vol. 7, no 1, p. 1-26
- Bégué A., Vintro 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.
- Berdugo V. G., Chaussin C., Dubus L., Hebrail G. et Leboucher V. (2011).** Analog method for collaborative very-short-term forecasting of power generation from photovoltaic systems. Next Generation Data Mining Summit: Ubiquitous Knowledge Discovery for Energy Management in Smart Grids and Intelligent Machine-to-Machine (M2M). *Telematics*, 5 pages.
- Bolboaca S. D. et Jäntschi L. (2013).** Quantitative Structure-Activity Relationships: Linear Regression Modelling and Validation Strategies by Example. *Biomath*, 2(1).
- Boudet G. (1977).** Contribution au contrôle continu des pâturages tropicaux en Afrique occidentale. *Revue d'Elevage et de Médecine Vétérinaire des Pays Tropicaux*, 30(4): 387-406.

- Bozdogan H. (1987).** Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, 52(3): 345-370.
- Bozzini C. et Maselli F. (2002).** Analysis of Multitemporal NDVI Data for Crop Yield Forecasting in the Sahel. *GeoCarto International*, 17(1): 53-60.
- Brase C. H. et Brase C. P. (2012).** *Brase Understandable Statistics : Concepts and Methods*. Australia , Brazil, Japan, Korea, Mexico, Singapore, Spain, United Kingdom, United States: Richard Stratton. 844 page.
- Breman H. (1982).** La production actuelle des pâturages naturels en relation avec des facteurs naturels et des facteurs anthropiques. In : Penning de Vries F. W. T. et M. A. Djitéye. *La productivité des pâturages sahéliens*. Wageningen : Centre for Agriculture Publishing and documentation, 284-384.
- Breman H. et de Ridder N. (1991).** *Manuel sur les pâturages des pays sahéliens* : Paris : ACCT : Karthala ; Wageningen : CTA, 481 pages.
- Burnham K. P. et Anderson D. R. (2002).** *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. New York : Springer, 512 pages.
- Burnham K. P., Anderson D. R. et Huyvaert K. P. (2011).** AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology*, 65(1): 23-35.
- Burke S. (2001).** Missing values, outliers, robust statistics & non-parametric methods. *LC-GC Europe Online Supplement, Statistics and Data Analysis*, 2(0), 19-24.
- Butt B., Turner M. D., Singh A. et Brottem L. (2011).** Use of MODIS NDVI to evaluate changing latitudinal gradients of rangeland phenology in Sudano-Sahelian West Africa. *Remote Sensing of Environment*, 115(12): 3367-3376.
- Chai T. et Draxler R. R. (2014).** Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geoscientific Model Development Discussions*, 7(3):1247-1250.
- Chang C.-L., Franses P. H. et McAleer M. (2011).** How accurate are government forecasts of economic fundamentals? The case of Taiwan. *International Journal of Forecasting*, 27(4): 1066-1075.

- Cheetham A. H. et Hazel J. E. (1969).** Binary (presence-absence) similarity coefficients. *Journal of Paleontology* : 1130-1136.
- Cissé S. , Eymard L., Ottlé C., Ndione J., Gaye A., Pinsard F. (2016).** Rainfall Intra-Seasonal Variability and Vegetation Growth in the Ferlo Basin (Senegal). *Remote Sens.* 8(1), 66.
- Choi Y.-S., Lindzen R. S., Ho C.-H. et Kim J. (2010).** Space observations of cold-cloud phase change. *Proceedings of the National Academy of Sciences*, 107(25): 11211-11216.
- Cornillon P. A. et Matzner-Løber E. (2007).** *Régression. Théorie et applications. Collection Statistique et Probabilités Appliquées*. Paris: Springer, 314 pages.
- Dagnelie P. (2013).** *Statistique théorique et appliquée: 1. Statistique descriptive et base de l'inférence statistique*, 3<sup>e</sup> éd, Bruxelles: De Boeck. 518 pages.
- Diallo O., Diouf A., Hanan N. P., Ndiane A. et Prevost Y. (1991).** AVHRR monitoring of savanna primary production in Senegal, West Africa: 1987-1988. *International Journal of Remote Sensing*, 12(6):1259-1279.
- Diouf A. A., Djaby B., Diop M. B., Wele A. Ndione J.-A., Tychon B. (2014).** Fonction d'ajustement pour l'estimation de la production fourragère herbacée des parcours naturels du Sénégal à partir du NDVI S10 de SPOT-VEGETATION. XXVII Colloque de l'Association Internationale de Climatologie, Dijon (France).
- Diouf A., Brandt M., Verger A., Jarroudi M., Djaby B., Fensholt R., . . . Tychon, B. (2015).** Fodder Biomass Monitoring in Sahelian Rangelands Using Phenological Metrics from FAPAR Time Series. *Remote Sensing*, 7(7), 9122.
- Diouf A., Hiernaux P., Brandt M., Faye G., Djaby B., Diop M., Tychon B. (2016).** Do Agrometeorological Data Improve Optical Satellite-Based Estimations of the Herbaceous Yield in Sahelian Semi-Arid Ecosystems? *Remote Sensing*, 8(8), 668.
- Dixon W. J. et Mood A. M. (1946).** The statistical sign test. . *Journal of the American Statistical Association*, 41(236): 557-566.
- DNEP (2013).** *Déclaration de N'djamena pour l'Elevage Pastoral: une contribution durable au développement et à la sécurité des espaces saharo-sahéliens*. colloque régional conférence ministérielle, 7 pages.
- DNP (2013).** *Déclaration de Nouakchott pour le Pastoralisme: Mobilisons ensemble un effort ambitieux pour un pastoralisme sans frontière*. conférence des chefs d'Etats et de gouvernements. 4 pages.

- Durai V. R. et Bhradwaj R. (2014).** Evaluation of statistical bias correction methods for numerical weather prediction model forecasts of maximum and minimum temperatures. *Natural Hazards*, 73(3) : 1229-1254.
- Eerens H. et Haesen D. (2015).** *SPIRITS Manual: Software for the Processing and Interpretation of Remotely sensed Image Time Series. User's manual Version 1.2.1*, Mol:VITO,339 pages.
- Emmert-Streib F. et Dehmer M. (2009).** *Information Theory and Statistical Learning*. New York ; [London] : Springer, 443 pages.
- Englhart S., Keuck V. et Siegert F. (2011).** Aboveground biomass retrieval in tropical forests — The potential of combined X- and L-band SAR data use. *Remote Sensing of Environment*, 115(5): 1260-1271.
- Evrendilek F. et Gulbeyaz O. (2008).** Deriving vegetation dynamics of natural terrestrial ecosystems from MODIS NDVI/EVI data over Turkey. *Sensors*, 8(9): 5270-5302.
- FAO (1981).** *Manuel d'inventaire forestier*. n° 27, Organisation des Nations Unies pour l'Alimentation et l'Agriculture(FAO), Rome, Italie, 194 pages.
- FAO et Unesco (1975).** *Carte mondiale des sols 1 : 5 000 000*, Volume I, Légende, Unesco, Paris, 68 pages.
- Fernández-Ferrero A., Sáenz J. et Ibarra-Berastegi G. (2010).** Comparison of the Performance of Different Analog-Based Bayesian Probabilistic Precipitation Forecasts over Bilbao, Spain. *Monthly Weather Review* 138(8): 3107-3119.
- Fontana D. C., Potgieter A. B. et Apan A. (2007).** Assessing the relationship between shire winter crop yield and seasonal variability of the MODIS NDVI and EVI images. *Applied GIS*, 3(7): 1-16.
- Gaston A., Dieye K., Bergès J. C. et Citeau J. (1987).** Evaluation de la productivité primaire herbacée annuelle au Sahel sénégalais: saison des pluies 1986. *Veille Climatique Satellitaire* (17): 49-54.
- Gibbons J. D. et Chakraborti S. (2014).** *Nonparametric statistical inference International Encyclopedia of Statistical Science*. Berlin Heidelberg: Springer, 977-979.
- Gilbert C. (2004).** *Statistique non paramétrique élémentaire*. cours de M2 ESA, Université d'Orléans, 52 pages.
- Good P. I. et Hardin J. W. (2012).** *Common Errors in Statistics (and How to Avoid Them)*. 4<sup>e</sup> éd, Hoboken, New Jersey: John Wiley et Sons, 335 pages.

- Groten S. M. E. (1993).** NDVI-crop monitoring and early yield assessment of Burkina Faso. *Remote sensing*, 14(8): 1495-1515.
- Grubbs F. E. (1969).** Procedures for Detecting Outlying Observations in Samples. *Technometrics*, Vol. 11 (1): 1-21.
- Hamill T. M. et Whitaker J. S. (2006).** Probabilistic Quantitative Precipitation Forecasts Based on Reforecast Analogs: Theory and Application. *Monthly Weather Review* 134(11): 3209-3229.
- Hiernaux P. (1984).** *Distribution des pluies et production herbacée au Sahel: une méthode empirique pour caractériser la distribution des précipitations journalières et ses effets sur la production herbacée. Premiers résultats acquis dans le Sahel malien*, CIPEA, Bamako, 46 pages.
- 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.
- Hiernaux P. et Le Houérou H. N. (2006).** Les parcours du Sahel. *Science et changements planétaires/Sécheresse*, 17(1): 51-71.
- Hmimina G., E. Dufrêne, Pontailler J. Y., Delpierre N., Aubinet M., Caquet B., de Grandcourt A., Burban B., Flechard C., Granier A., Gross P., Heinesch B., Longdoz B., Moureaux C., Ourcival J. M., Rambal S., Saint André L. et Soudani K. (2013).** Evaluation of the potential of MODIS satellite data to predict vegetation phenology in different biomes: An investigation using ground-based NDVI measurements. *Remote Sensing of Environment*, 132(0): 145-158.
- Holben B., D. Kimes et Fraser R. S. ( 1986).** Directional reflectance response in AVHRR red and near-IR bands for three cover types and varying atmospheric conditions. *Remote Sensing of Environment*, 19(3): 213-236.
- Holben B. N. et Fraser R. S. (1984).** Red and Near infrared response to off-nadir viewing. *International .Journal of Remote Sensing*, 5: 145-160.
- Huete A., K. Didan, Miura T., Rodriguez E. P., Gao X. et Ferreira L. G. (2002).** Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1): 195-213.
- Hulme M. (1992).** *Confronting climate change: Risks, implications and responses*. Cambridge: Cambridge University Press, 352 pages.
- Hurvich C. M. et Tsai C.-L. (1989).** Regression and time series model selection in small samples. *Biometrika*, 76(2) : 297-307.

- IEMVT (1987).** Élevage et potentialités pastorales sahéliennes. Synthèses cartographiques, Niger. CTA, Wageningen, IEMVT, Maisons-Alfort: IEMVT-CIRAD, CTA.,32 pages.
- Johnson J. B. et Omland K. S. (2004).** Model selection in ecology and evolution. *Trends in ecology & evolution* 19(2) : 101-108.
- Johnson R. A. et Bhattacharyya G. K. (2010).** *Statistics- Principles and Methods*. Hoboken, NJ: John Wiley & Sons, 706 pages.
- Jönsson P. etEklundh L. (2004).** TIMESAT—a program for analyzing time-series of satellite sensor data. *Computers & Geosciences*, 30(8): 833-845.
- Justice C. O. et Hiernaux P. H. Y. (1986).** Monitoring the grasslands of the Sahel using NOAA AVHRR data: Niger 1983. *International Journal of Remote Sensing*, 7(11): 1475-1497.
- Justice C. O., Holben B. N. et Gwynne M. D. (1986).** Monitoring East African vegetation using AVHRR data. *International Journal of Remote Sensing*, 7(11): 1453-1474.
- Justice C. O., Markham B. L., Townshend J. R. G. et Kennard R. L. (1989).** Spatial degradation of satellite data. *International Journal of Remote Sensing*, 10(9): 1539-1561.
- Justice C. O., Townshend J. R. G., Holben B. N. et Tucker C. J. (1985).** Analysis of the phenology of global vegetation using meteorological satellite data. *International Journal of Remote Sensing*, 6(8): 1271-1318.
- Kouadio L. (2012).** Prévisions des rendements du blé d'hiver à échelle régionale par modélisation de la courbe de chute de l'indice foliaire. Thèse de doctorat: Université de Liège, Belgique,187 pages.
- Kouadio L., Newlands N., Davidson A., Zhang Y., et Chipanshi A. (2014).** Assessing the Performance of MODIS NDVI and EVI for Seasonal Crop Yield Forecasting at the Ecodistrict Scale. *Remote Sensing*, 6(10), 10193.
- KAS (2014).** *La sécurité au Sahel après la crise du Mali Quels enjeux et défis pour les pays régionaux et internationaux[en ligne]*. Séminaire international organisé le 28 mars 2014 à Rabat, Maroc, 8 pages. Disponible sur le site <http://www.kas.de/marokko/fr/publications/>
- Kastens J. H., Kastens T. L., Kastens D. L. A., Price K. P., Martinko E. A. et Lee R.-Y (2005).** Image masking for crop yield forecasting using AVHRR NDVI time series imagery. *Remote Sensing of Environment*, 99 (3): 341-356.

- Kolodner J. L. (1997).** Educational implications of analogy: A view from case-based reasoning. *American psychologist* 52(1).
- Kolodner J. L. ( 1992.).** An introduction to case-based reasoning. *Artificial Intelligence Review*, 6(1) : 3-34.
- Krishnaswamy J., Bawa K. S., Ganeshia K. N.et Kiran M. C. (2009).** Quantifying and mapping biodiversity and ecosystem services: Utility of a multi-season NDVI based Mahalanobis distance surrogate. *Remote Sensing of Environment*, 113(4): 857-867.
- Le Houérou H. N. (1980).** The role of browse in the Sahelian and Sudanian zones, in Browse in Africa, the current state of knowledge, H. N. Le Houero (ed), ILCA, Addis Ababa, Ethiopia, 83-100.
- Leake D. B. (1996).** *Case-Based Reasoning: Experiences, lessons and future directions*. Menlo Park, Calif. : AAAI Press ; Cambridge, Mass. : MIT Press, 525 pages.
- Lejeune M. (2010).** *Statistique: La théorie et ses applications*. 2<sup>e</sup> éd, Paris, Berlin, Heidelberg, New York, Hong Kong, Londres, Milan, Tokyo: Springer, 448 pages.
- Lesot M.-J., Rifqi M.et Benhadda H. (2009).** Similarity measures for binary and numerical data: a survey. *International Journal of Knowledge Engineering and Soft Data Paradigms*, 1(1): 63-84.
- Livezey R. E., Barnston A. G., Gruza G. V.et Ran'kova E. Y. (1994).** Comparative skill of two analog seasonal temperature prediction systems: Objective selection of predictors. *Journal of climate*, 7(4) : 608-615.
- Lobell D. B. et Asner G. P. (2004).** Cropland distributions from temporal unmixing of MODIS data. *Remote Sensing of Environment*, 93(3): 412-422.
- Lorenz E. N. (1969).** Atmospheric predictability as revealed by naturally occurring analogues. . *Journal of the Atmospheric sciences*, 26(4) : 636-646.
- Lona I. (2014).** Risques climatiques et pratiques culturales du mil et du sorgho au Niger. thèse de doctorat ,Université Abdou Moumouni de Niamey, 248 pages.
- Lüdeke M. K. B., Ramage P. H.et Kohlmaier G. H. (1996).** The use of satellite NDVI data for the validation of global vegetation phenology models: application to the Frankfurt Biosphere Model. *Ecological Modelling*, 91(1): 255-270.
- Maïdagui B., Denda I., Wylie B., Harrington J. (1987)** *Pasture production in the central and eastern pastoral zones of Niger*. Technical

report, Integrated Livestock Production Project, Government of Niger/Tufts University/USAID, 32 pages.

**Maher M. L. et Garza A. G. d. S. (1997).** Case-based reasoning in design. *IEEE Intelligent Systems*, 12(2) : 34-41.

**Maselli F., Conese C., Petkov L. et Gilabert M. A. (1992).** Environmental monitoring and crop forecasting in the Sahel through the use of NOAA NDVI data. A case study: Niger 1986–89. *International Journal of Remote Sensing*, 14(18): 3471-3487.

**Maselli F., Romanelli S., Bottai L. et Maracchi G. (2000).** Processing of GAC NDVI data for yield forecasting in the Sahelian region. *International Journal of Remote Sensing*, 21(18): 3509-3523.

**Matulla C., Zhang X., Wang X. L., Wang J., Zorita E., Wagner S. et Storch H. v. (2008).** Influence of similarity measures on the performance of the analog method for downscaling daily precipitation. *Climate Dynamics*, 30(2-3): 133-144.

**McQuarrie A., Shumway R. et Tsai C.-L. (1997).** The model selection criterion AICu. *Statistics & probability letters*, 34(3): 285-292.

**McQuarrie A. D. R. et Tsai C.-L. (1998).** *Regression and time series model selection*. Singapore: World Scientific, 480 pages.

**MEIA (2012).** Rapport d'activité de l'année 2012. Ministère de l'elevage et des industries Animale, Niamey, NIGER, 82 pages.

**Meier K., Brudney J. et Bohte J. (2011).** *Applied statistics for public and nonprofit administration*. Belmont, Calif. : Wadsworth ; Andover : Cengage Learning.

**Meroni M., Rembold F., Verstraete M. M., Gommes R., Schucknecht A. et Beye G. (2014).** Investigating the relationship between the inter-annual variability of satellite-derived vegetation phenology and a proxy of biomass production in the Sahel. *Remote Sensing*, 6(6) : 5868-5884.

**Mitchell H. B. (2010).** *Image fusion: theories, techniques and applications*. Berlin: Springer, 242 pages.

**Mkhabela M. S., Bullock P., Raj S., Wang S. et Yang Y. (2011).** Crop yield forecasting on the Canadian Prairies using MODIS NDVI data. *Agricultural and Forest Meteorology*, 151(3): 385-393.

**Morland J. C., Grimes D. I. F. et Hewison T. J. (2001).** Satellite observations of the microwave emissivity of a semi-arid land surface. *Remote Sensing of Environment*, 77(2): 149-164.

**Ndiaye C., Guèye E. F., Ngom S., Traoré E. H. et Akpo L. E. (2015).** Impact de la pluviométrie sur la production de biomasse

- herbagère en milieu sahélien : un cas en Région de Matam, Nord du Sénégal. *Livestock Research for Rural Development*, 27 (55).
- Nicholson S. E. (1981).** The historical climatology of Africa. *Climate and history*, 249-270.
- Nori M. et Davies J. (2007).** *Vent de changement: changement climatique, adaptation et pastoralisme : Initiative mondiale pour un pastoralisme durable*, Rapport. UICN, 26 pages.
- Novella N. S. et Thiaw W. M. (2012).** African Rainfall Climatology Version 2 for Famine Early Warning Systems. *Journal of Applied Meteorology and Climatology*, 52(3):588-606.
- Okin G. S., Clarke K. D. et Lewis M. M. (2013).** Comparison of methods for estimation of absolute vegetation and soil fractional cover using MODIS normalized BRDF-adjusted reflectance data. *Remote Sensing of Environment*, 130 (0): 266-279.
- OMD1 (2011).** *Cadre d'accélération des OMD* : Réduire l'extrême pauvreté et la faim Sécurité alimentaire et nutritionnelle au Niger. Objectif du Millénaire pour un Développement Durable (OMD). Niamey, NIGER, 74 pages.
- Ozer P., Hountondji Y., Niang A. J., Karimoune S., Laminou Manzo O., et Salmon M. (2010).** Désertification au Sahel: historique et perspectives. *Bulletin de la Société Géographique de Liège*, 54, 69-84.
- Paulson D. S. (2008).** *Biostatistics and Microbiology*: A Survival Manual. New York, NY: Springer, 216 pages.
- Penning D. V., F. W. T. et Djiteye M. A. (1982).** *La productivité des pâturages sahéliens: une étude des sols, des végétations et de l'exploitation de cette ressource naturelle*. Wageningen : Centre for Agricultural Pub. and Documentation, 525 pages.
- Potdar M. B., Manjunath K. R. et Purohit N. L. (1999).** Multi-season atmospheric normalization of NOAA AVHRR derived NDVI for crop yield modeling. *Geocarto International*, 14(4): 52-57.
- Rakotomalala R. (2010).** *Comparaison de populations*, Cours de statistique. Université Lumière, Lyon 2, Lyon , France, 201 pages.
- Rakotomalala R. ( 2008).** *Tests de normalité*, Cours de statistique. Université Lumière Lyon 2, 59 pages.
- Rakotomalala R. ( 2012).** *Analyse de corrélation*. Cours de statistique. Université Lumière Lyon 2, 89 pages.
- Rasmussen M. S. (1998).** Developing simple, operational, consistent NDVI-vegetation models by applying environmental and climatic

information. Part II: Crop yield assessment. *International Journal of Remote Sensing*, 19(1): 119-139.

**RGAC. (2008).** *Recensement générale de l'agriculture et du cheptel*, Rapport préparé pour le Gouvernement du Niger par l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO). Rome, 47 pages.

**Rhissa Z. (2010).** *Revue du secteur de l'élevage au Niger*, rapport d'étude. Niamey, NIGER , 115 pages.

**Riesbeck C. K. et Schank R. C. (2013).** *Inside case-based reasoning*. New York : Psychology Press, 448 pages.

**Rifqi M. (2010).** *Mesures de similarité, raisonnement et modélisation de l'utilisateur*. Thèse de doctorat: Université Pierre et Marie Curie, 97 pages.

**Roebber P. J. et Bosart L. F. (1998).** The Sensitivity of Precipitation to Circulation Details. Part I: An Analysis of Regional Analogs. . *Monthly Weather Review*, 126(2): 437-455.

**Rouse J. W., Haas R. H., Deering D. W., Schell J. A. et Harlan J. C. (1973).** *Monitoring vegetation systems in the Great Plains with ERTS*. In Proceedings of the Third ERTS Symposium; NASA SP-351; NASA: Washington, DC, USA; pp. 309–317.

**Rousson V. (2013).** *Statistique appliquée aux sciences de la vie*. Paris, Berlin, Heidelberg, New York, Hong Kong, Londres, Milan, Tokyo: Springer, 327 pages.

**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.

**San Martini A. et Spezzaferri F. (1984).** A predictive model selection criterion. *Journal of the Royal Statistical Society. Series B (Methodological)* : 296-303.

**Sano E. E., Moran M. S., Huete A. R. et Miura T. (1998).** C- and Multiangle Ku-Band Synthetic Aperture Radar Data for Bare Soil Moisture Estimation in Agricultural Areas. *Remote Sensing of Environment*, 64(1): 77-90.

**Sivakumar M. V. K., Maidoukia A. et Stern R. D. (1993).** *Agroclimatology of west africa: Niger*. Patancheru, Andhra Pradesh, India : International Crops Research Institute for the Semi-Arid Tropics ; Niamey, Niger : Direction de la météorologie nationale du Niger.

- Soudani K., Hmimina G., Delpierre N., Pontailler J. Y., Aubinet M., Bonal D., Caquet B., de Grandcourt A., Burban B., Flechard C., Guyon D., Granier A., Gross P., Heinesh B., Longdoz B., Loustau D., Moureaux C., Ourcival J. M., Rambal S., Saint André L. et Dufrêne E. (2012).** Ground-based Network of NDVI measurements for tracking temporal dynamics of canopy structure and vegetation phenology in different biomes. *Remote Sensing of Environment*, 123(0): 234-245.
- Steinfeld H., Gerber P., Wassenaar T., Castel V., Rosales M. et De Haan C. (2009).** *L'ombre portée de l'élevage: impacts environnementaux et options pour leur atténuation*. Rome : Organisation des nations unies pour l'alimentation et l'agriculture (FAO), 464 pages.
- Sullivan M. III (2013).** Statistics Informed Decisions Using Data. 4<sup>e</sup> ed, UK: Pearson, 968 pages.
- Swets D. L., Reed B. C. et Rowland J. D. (1999).** *A weighted leastsquares approach to temporal smoothing of NDVI*. In Proceedings of the 1999 ASPRS Annual Conference: From Image to Information, Portland, Oregon, May 17–21.
- Swinnen E., Hoolst R. V., Eerens H. et Toté C. (2015).** Dray Matter Productivity (DMP)-PROBA-V, Quality assessment Report. Belgium: VITO, 77 pages.
- Tangang F. T., Hsieh W. W. et Tang B. (1997).** Forecasting the equatorial Pacific sea surface temperatures by neural network models. *Climate Dynamics*, 13(2):135-147.
- Toure I., Ickowicz A., Ancey V., Gerber P., Wane A., Saley M., Garba I. et Atte I. (2011).** *SIPSA: un système d'information et d'alerte précoce pour accompagner les prises de décision dans le domaine du pastoralisme au Sahel*. Colloque : la Politique sectorielle du pastoralisme au Tchad. N'Djaména : Ministère de l'élevage et des ressources animales, p. 85-92.
- Touré I., Ickowicz A., Wane A., Garba I., Gerber P., Atte I., Cesaro J. D., Diop A. T., Djibo S., Ham F. (2012).** *Atlas des évolutions des systèmes pastoraux au Sahel: 1970-2012*. Rome: Food and Agriculture Organization of the United Nations, 32 pages.
- Trendowicz A. et Jeffery R. (2014).** Software Project Effort Estimation. Cham : Springer, 470 pages.
- Tucker C. J., Holben B. N., Elgin J. H. et McMurtrey J. E. (1981).** Remote sensing of total dry-matter accumulation in winter wheat. *Remote Sensing of Environment*, 11(0):171-189.

- Tucker C.J., Vanpraet C., Boerwinkel E., Gaston A. ( 1983).** Satellite remote sensing of total dry matter production in the senegalese sahel. *Remote Sens. Environ.* , 13, 461–474.
- Tucker C. J., Justice C. O. et Prince S. D. (1986).** Monitoring the grasslands of the Sahel 1984-1985. *International Journal of Remote Sensing*, 7(11):1571-1581.
- Tucker C. J., Pinzon J. E., Brown M. E., Slayback D. A., Pak E. W., Mahoney R., Vermote E. F. et Saleous N. (2005).** An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. *International Journal of Remote Sensing*, 26(20): 4485-4498.
- Tucker C. J., Vanpraet C. L., Sharman M. J. et Van Ittersum G. (1985).** Satellite remote sensing of total herbaceous biomass production in the Senegalese Sahel: 1980–1984. *Remote Sensing of Environment*, 17(3): 233-249.
- UA (2013).** Union africaine *cadre stratégique pour le pastoralisme en Afrique: Sécuriser, Protéger et améliorer les vies, les moyens de subsistance et les droits des communautés pastorales*, Addis Abeba ,Ethiopie, 56 pages.
- UNESCO, PNUD et FAO (1981).** Écosystèmes pâturels tropicaux. Un rapport sur l'état des connaissances préparé par l'UNESCO, le PNUD et la FAO. Paris: unesco, 675 pages.
- Wagenseil H. et Samimi C. (2006).** Assessing spatio-temporal variations in plant phenology using Fourier analysis on NDVI time series: results from a dry savannah environment in Namibia. *International Journal of Remote Sensing*, 27(16): 3455-3471.
- Wang Q., Tenhunen J., Dinh N. Q., Reichstein M., Vesala T. et Keronen P. (2004).** Similarities in ground- and satellite-based NDVI time series and their relationship to physiological activity of a Scots pine forest in Finland. *Remote Sensing of Environment*, 93(1–2): 225-237.
- Watson I. ( 1999).** Case-based reasoning is a methodology not a technology. *Knowledge-based systems*, 12(5) : 303-308.
- Watson I. et Marir F. (1994).** Case-based reasoning: A review. *The knowledge engineering review*, 9 (04) : 327-354.
- Weiers R. M. et Heinz H. J. (2011).** *Introduction to Business Statistics*. USA: Joe Sabatin, 892 pages.

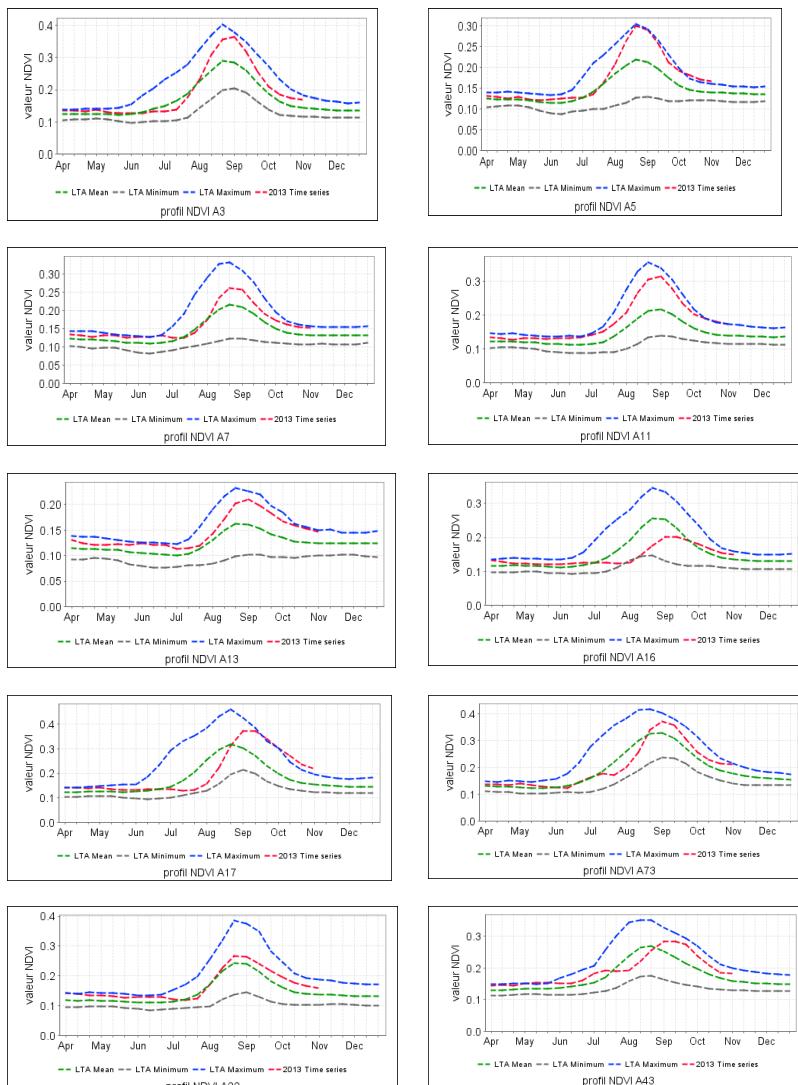
- Wilding G. E. et Mudholkar G. S. (2008).** Empirical approximations for Hoeffding's test of bivariate independence using two Weibull extensions. *Statistical Methodology*, 5(2):160-170.
- Wittemyer G., Rasmussen H. B. et Douglas-Hamilton I. (2007).** Breeding phenology in relation to NDVI variability in free-ranging African elephant. *Ecography*, 30(1): 42-50.
- Wylie B. K., Harrington J. A., Prince S. D. et Denda I. (1991).** Satellite and ground-based pasture production assessment in Niger: 1986-1988. *International Journal of Remote Sensing*, 12(6): 1281-1300.
- Xavier P. K. et Goswami B. N. ( 2007).** An analog method for real-time forecasting of summer monsoon subseasonal variability. *Monthly Weather Review* 135(12): 4149-4160.
- Yin Y. et Yasuda K. (2005).** Similarity coefficient methods applied to the cell formation problem: a comparative investigation. *GroupTechnology/Cellular Manufacturing*, 48(3): 471-489.
- Zorita E. et Von Storch H. (1999).** The analog method as a simple statistical downscaling technique: comparison with more complicated methods. *Journal of climate* 12(8): 2474-2489.
- Zucchini W. (2000).** An introduction to model selection. *Journal of mathematical psychology*, 44(1) : 41-61.

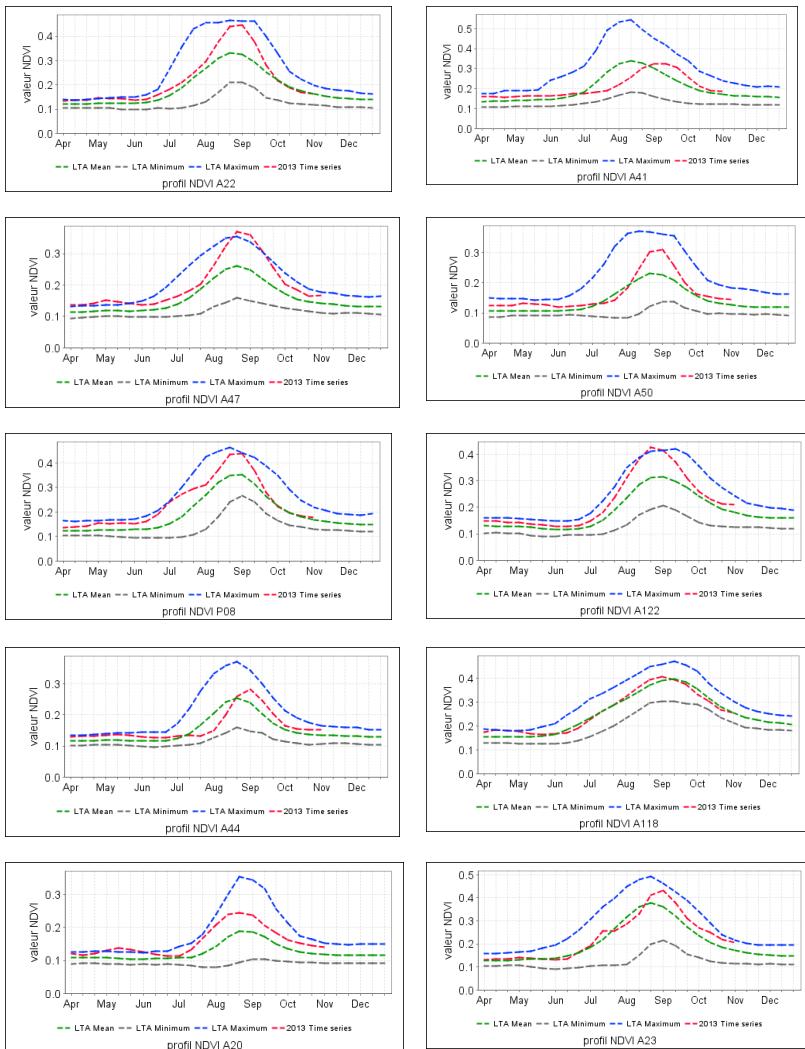


## Annexes

### Annexes 1

#### Annexe 1.1 : Exemples de profil du NDVI sur les sites de mesures de biomasse herbacée aérienne du MEIA





LTA Mean :moyenne de la série ; LTA minimum : minimum de la série ;

LTA maximum : maximum de la série;  
2013 time series: profil ndvi de la saison 2013

## Annexe 1.2 : localisation des sites du MEIA

Sites	Nom localité	Longitude	Latitude
A64	Aborak	6,813069	15,847222
A26	Tchintaborak	7,383594	15,765573
A77	Rigia Malam Bio	7,606410	15,725534
A15	Amataltalen	7,134800	16,149400
A19	Ablama	7,803400	16,132400
A33	Tamalolo	6,813297	16,443647
A20	Toumbelaga	7,822515	15,827207
A13	RN 25 Pk 55	7,539740	16,793316
A11	Ingall	7,056415	16,613082
A80	Tiguidan Tagait	7,593611	17,234117
A23	Fako	6,671119	14,962770
P08	Bermo	7,172500	15,036100
A66	Gadabeji Nord	7,252837	15,064663
A65	Gadabeji	6,744779	15,115499
A22	Amoulasse	7,045334	15,101667
A24	Abouhaya	6,527738	15,375873
A16	Iksman	6,378361	15,593056
A17	Ibeceten2	5,862818	15,403571
A7	Targa	5,705034	16,491389
A3	Tchintab	5,723350	15,725534
A5	Tchintassalaten	5,772836	16,081608
A30	Ibeceten 1	5,838932	15,325043
A62	Obserata	4,554423	15,697799
A32	Tofaminir	6,615618	16,013827
A61	Amilal	5,154180	15,861098
A72	Doutchi	3,747800	13,884200
A118	Kara	2,810832	12,710832
A40	Desert de Tall	12,731721	14,149390
A41	Sayam	12,545295	13,793316
A42	Bouti	11,424800	13,951300

Sites	Nom localité	Longitude	Latitude
A43	Birnia	11,291120	13,945825
N82		13,979700	13,813300
n83		13,227800	14,616100
n84		13,160278	14,454444
N85		13,117222	14,356667
n86		13,103100	14,517800
n87		13,227800	14,616100
n88		13,160278	14,454444
N89		12,873056	14,090556
N93		13,103100	14,517800
n95		12,961667	14,608611
n96		12,736944	14,066944
n97		12,873056	14,090556
n98		12,766100	14,172800
A110		12,772200	13,953600
A111		13,107778	13,731667
A112		13,128611	13,619722
A113		11,689600	13,848900
A114		11,231200	13,944600
A115		11,035800	13,887100
A117		10,914400	13,715000
A44	Yougoum	11,443639	14,579195
A45	Aljannari	11,087797	14,731686
A46	Aborak	10,443619	14,948889
A47	Téjira	10,070847	14,861098
A48	Bathé1	7,743333	14,894988
A49	N Tabanot	7,776389	15,318844
A50	Jeptoji	8,331344	15,318844
A51	Tenbia	9,221074	15,430000
A52	Bathé	7,938328	14,996661
A67	Didiga	3,991913	15,542222
A68	Ekrafane	3,671151	15,403571
A69	banibangou	2,831667	15,166335

Sites	Nom localité	Longitude	Latitude
A70	Inékére	2,676389	15,200226
A73	Toukounous	3,336417	14,606916
A105	Tiguit	0,630000	14,674700
A106	Alkongui	0,623900	14,731700
A121	Tin béré béré	0,970200	14,826000
A122	Boni	1,175700	14,745000

## Annexes 2

### Annexe 2.1 Estimation des valeurs de f

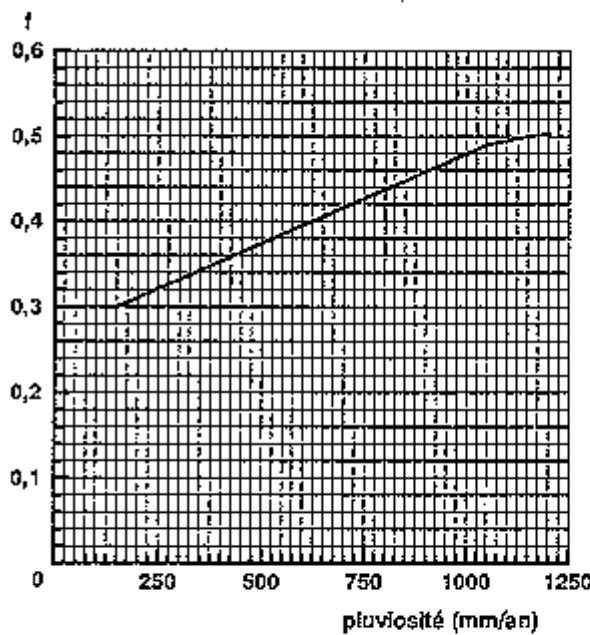


Figure II.4

Valeurs de la fraction moyenne de  $N_b$  perdue annuellement ( $f$ ) par rapport aux isohyètes.

Selon Penning de Vries & Djiteye, 1982 : p. 305.

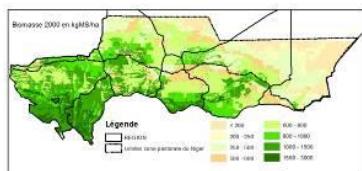


Figure 2 : Biomasse herbacée 2000

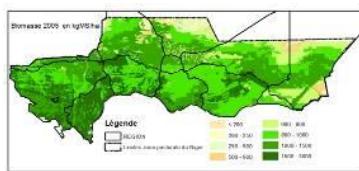


Figure7: Biomasse herbacée 2005

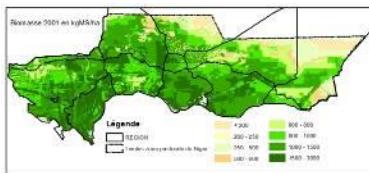


Figure 3 : Biomasse herbacée 2001

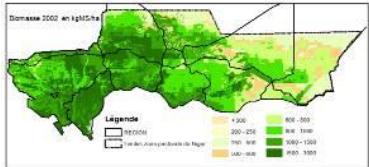


Figure 4 : Biomasse herbacée 2002

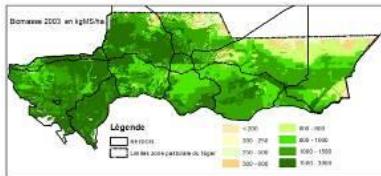


Figure 5 : Biomasse herbacée 2003

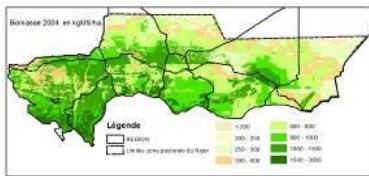


Figure 6 : Biomasse herbacée 2004

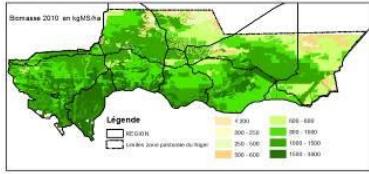


Figure 12 : Biomasse herbacée  
2010

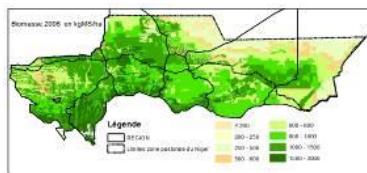


Figure8 : Biomasse herbacée 2006

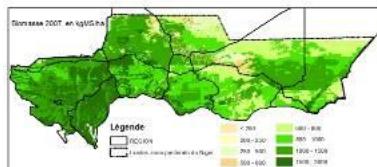


Figure9 : Biomasse herbacée 2007

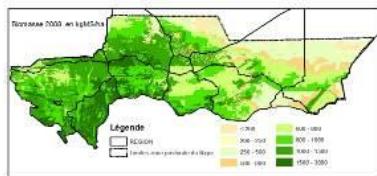


Figure10 : Biomasse herbacée 2008

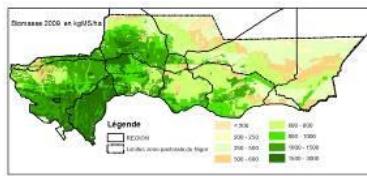


Figure11 : Biomasse herbacée 2009

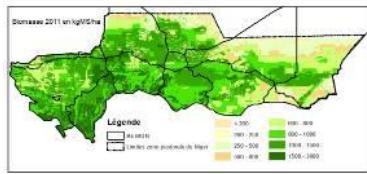


Figure13 : Biomasse herbacée 2011

## **Annexe 2.1 Biomasse herbacée suivant les années**

## **Annexe 2.2 Corrélation paramétrique sur l'ensemble des données**

Variable	variable	Corrél.	nbr obs	Limite de confiance inférieure 95 %	Limite de confiance supérieure 95 %	P-value
bioReel	bioherb	0,19	319	329,57	591,17	0,0001*

### Annexe 3 : ajustement de la masse herbacée aérienne en fonction du maximum et de l'intégrale NDVI suivant les années.

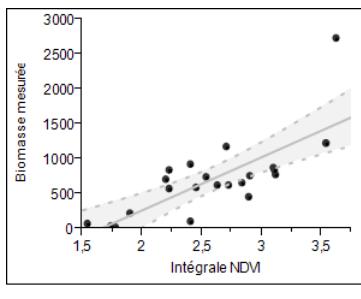


Figure 3.11 : Ajustement avec l'intégrale NDVI SPOT  
VEGETATION pour l'année 2001

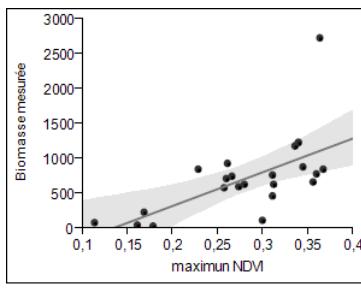


Figure 3.12 : Ajustement avec le maximum NDVI SPOT  
VEGETATION pour l'année 2001

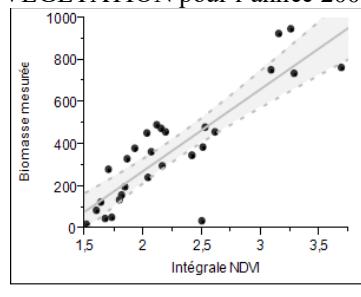


Figure 3.13 : Ajustement avec l'intégrale NDVI SPOT  
VEGETATION pour l'année 2002

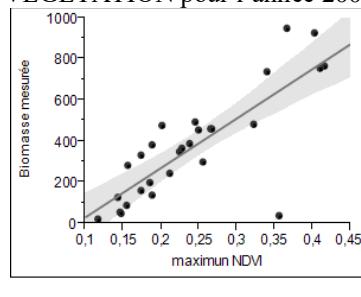


Figure 3.14 : Ajustement avec le maximum NDVI SPOT  
VEGETATION pour l'année 2002

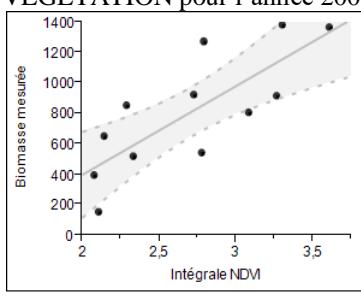


Figure 3.15 : Ajustement avec l'intégrale NDVI SPOT  
VEGETATION pour l'année 2003

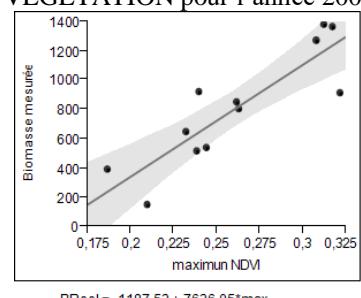
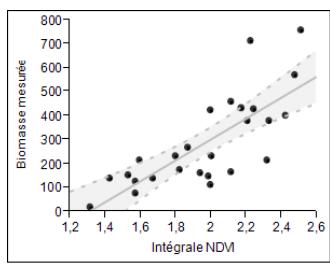
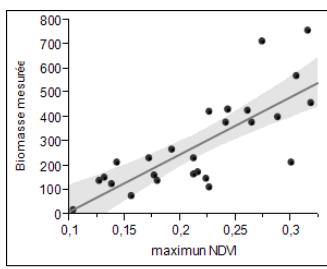


Figure 3.16 : Ajustement avec le maximum NDVI SPOT  
VEGETATION pour l'année 2003



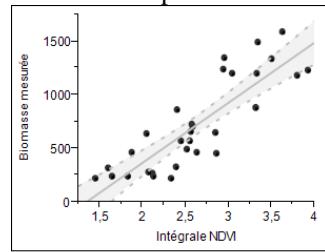
$$BR_{\text{Reel}} = -571,2492 + 435,89259 \cdot \text{Int}$$

Figure 3.17 : Ajustement avec l'intégrale NDVI SPOT  
VEGETATION pour l'année 2004



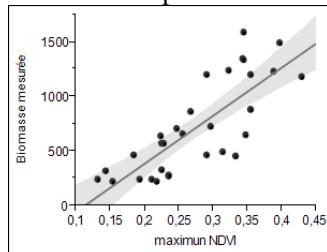
$$BR_{\text{Reel}} = -224,94 + 2357,85 \cdot \text{max}$$

Figure 3.18 : Ajustement avec le maximum NDVI SPOT  
VEGETATION pour l'année 2004



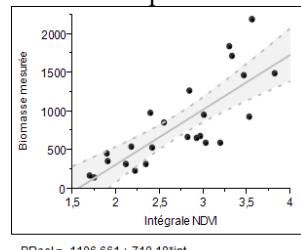
$$BR_{\text{Reel}} = -762,0385 + 561,78013 \cdot \text{int}$$

Figure 3.19 : Ajustement avec l'intégrale NDVI SPOT  
VEGETATION pour l'année 2005



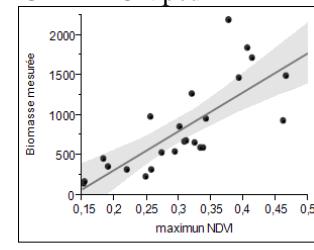
$$BR_{\text{Reel}} = -502,5637 + 4412,1453 \cdot \text{max}$$

Figure 3.20 : Ajustement avec le maximum NDVI SPOT  
VEGETATION pour l'année 2005



$$BR_{\text{Reel}} = -1106,661 + 710,18 \cdot \text{int}$$

Figure 3.21 : Ajustement avec l'intégrale NDVI SPOT  
VEGETATION pour l'année 2006



$$BR_{\text{Reel}} = -663,3123 + 4871,8838 \cdot \text{max}$$

Figure 3.22 : Ajustement avec le maximum NDVI SPOT  
VEGETATION pour l'année 2006

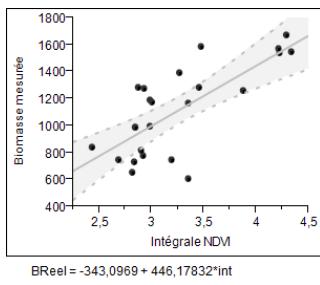


Figure 3.23 : Ajustement avec l'intégrale NDVI SPOT  
VEGETATION pour l'année 2007

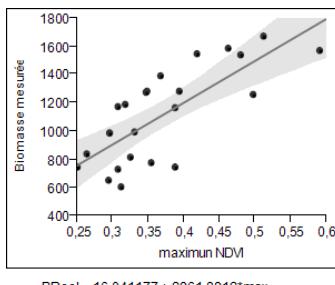


Figure 3.24 : Ajustement avec le maximum NDVI SPOT  
VEGETATION pour l'année 2007

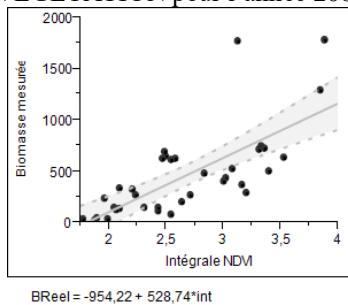


Figure 3.25 : Ajustement avec l'intégrale NDVI SPOT  
VEGETATION pour l'année 2008

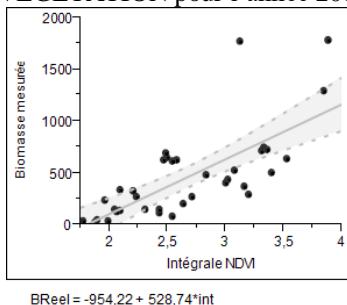


Figure 3.26 : Ajustement avec le maximum NDVI SPOT  
VEGETATION pour l'année 2008

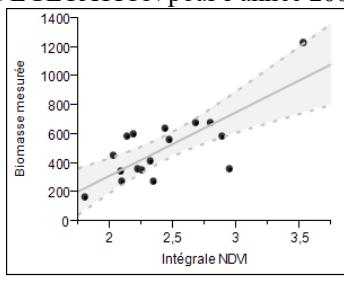


Figure 3.27 : Ajustement avec l'intégrale NDVI SPOT  
VEGETATION pour l'année 2009

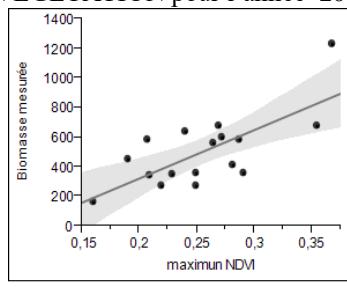
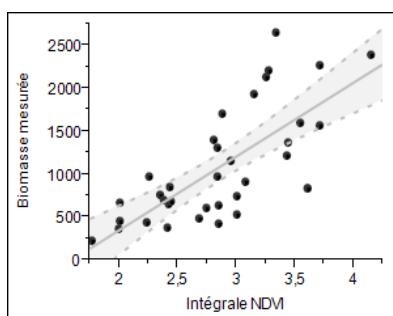
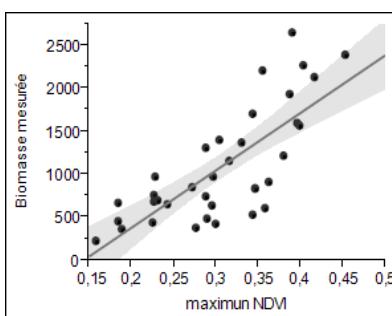


Figure 3.28 : Ajustement avec le maximum NDVI SPOT  
VEGETATION pour l'année 2009



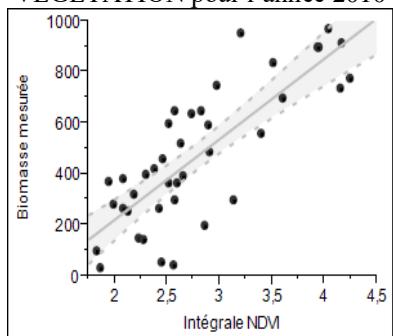
$$B_{\text{Reel}} = -1375,12 + 859,19 \cdot \text{int}$$

Figure 3.29 : Ajustement avec l'intégrale NDVI SPOT  
VEGETATION pour l'année 2010



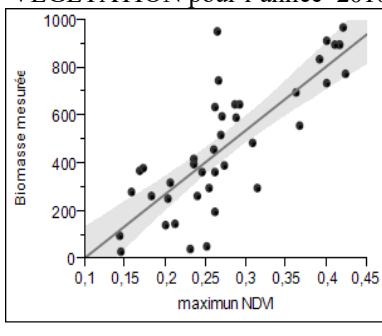
$$B_{\text{Reel}} = -968,49 + 6703,91 \cdot \text{max}$$

Figure 3.30 : Ajustement avec le maximum NDVI SPOT  
VEGETATION pour l'année 2010



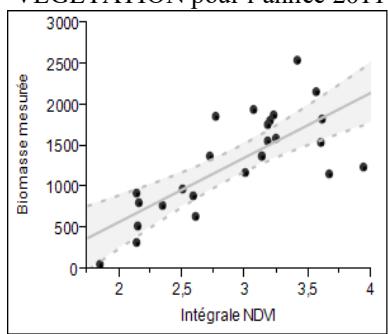
$$B_{\text{Reel}} = -413,05 + 316,12 \cdot \text{int}$$

Figure 3.31 : Ajustement avec l'intégrale NDVI SPOT  
VEGETATION pour l'année 2011



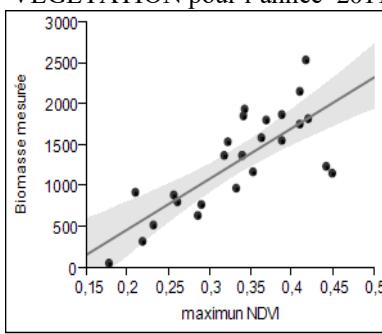
$$B_{\text{Reel}} = -262,61 + 2676,15 \cdot \text{max}$$

Figure 3.32 : Ajustement avec le maximum NDVI SPOT  
VEGETATION pour l'année 2011



$$B_{\text{Reel}} = -1000,97 + 786,70 \cdot \text{int}$$

Figure 3.33 : Ajustement avec l'intégrale NDVI SPOT  
VEGETATION pour l'année 2012



$$B_{\text{Reel}} = -766,32 + 6197,71 \cdot \text{max}$$

Figure 3.34 : Ajustement avec le maximum NDVI SPOT  
VEGETATION pour l'année 2012

## Annexe 4

### Annexe 4.1 Les variables indépendantes

Variables indépendantes	signification
PEA	PEAK : la décade à laquelle le NDVI atteint son maximum ;
SDA	SDAT : la décade de début de la saison de végétation ;
HOR	HORZ = PEAK – SDAT ;
SVA	SVAL : la valeur du NDVI à SDAT ;
PVA	PVAL : la valeur du NDVI à PEAK ;
VER	VERT = PVAL – SVAL ;
EVA	EVAL : le NDVI au temps PEAK + 4 (soit environ la fin de saison) ;
DRO	DROP = PVAL – EVAL ;
SLO	SLOP : la pente de la droite qui rejoint (SDAT, SVAL) à (PEAK, PVAL) ;
CUM	CUMM : la somme des valeurs de NDVI de SDAT à PEAK ;
SKEW	SKEW : le rapport entre la somme des 3 valeurs de NDVI suivant PEAK (de PEAK + 1 à PEAK + 3) et la somme des 7 valeurs de PEAK — 3 à PEAK + 3
MAX	maximum NDVI calculé par le MEIA
INT	intégrale NDVI calculé par le MEIA
RAI	cumul annuel de pluies (RFE)
TWR	La quantité d'eau requise pour le cycle complet de la culture (TWR) ;
Indx, IndxNor, IndxLatest	L'indice de satisfaction en eau en fin de cycle (Indx, IndxNor, IndxLatest) ;
SWi	Le contenu initial en eau dans le sol (SWi)
WEXi, WEXv, WEXf, WEXr, WEXt	L'excès en eau à différents stades phénologiques du cycle de croissance : phase initiale, phase végétative, floraison, maturité et durant tout le cycle (donnée par la somme des autres valeurs) : WEXi, WEXv, WEXf, WEXr, WEXt ;
WDEFi, WDEFv, WDEFf, WDEFr, WDEFt	Le déficit en eau à différents stades phénologiques et déficit en eau total en fin de cycle (WDEFi, WDEFv, WDEFf, WDEFr, WDEFt) ;
ETAi, ETAv, ETAf, ETAr, ETAt	L'évapotranspiration réelle à différents stades phénologiques et sa valeur totale en fin de cycle (ETAi, ETAv, ETAf, ETAr, ETAt) ;

## **Annexes 4.2. Modèle de régression linéaire multiple selon les types de sol**

### **Le modèle de Ge5-1a**

Le Tableau 4.13 montre que le sol Ge5-1a comprend 17 observations, la moyenne des ces observations est  $773 \text{ kg MS.ha}^{-1}$ . La mise en œuvre de la procédure de sélection de variables du type pas à pas descendante au niveau de ce type de sol a permis d'obtenir l'équation suivante :  $Y = -936,85 + 22,79 \text{ EVA}$ , ; l'expression de la prévision la variable pertinente (EVA).

La représentation des valeurs observées en fonction des valeurs prévues montre un  $R^2$  de 0,85 et un RMSE calculé de  $148 \text{ kg.ha}^{-1}$  (Figure 4.8). Le modèle est caractérisé par un RMSE relatif de 22 %. L'estimation des paramètres du modèle montre des probabilités hautement significatives (Tableau 4.14).

Tableau 4.13 : Résumé de l'ajustement

Paramètres	Valeurs
$R^2$	0,85
$R^2$ ajusté	0,84
$R^2$ validation	0,79
RMSE calculé	148
RMSE validation	161
Moyenne de la réponse	773
Nombre d'observations	17

Tableau 4.14 : Estimations des coefficients

Terme	Estimation	Erreur standard	t ratio	Prob. >  t
Constante	-936,8512	190,27	-4,92	0,0002*
EVA	22,793592	2,49	9,15	<,0001*

\*significatif au seuil de 10 pour 1000 ;

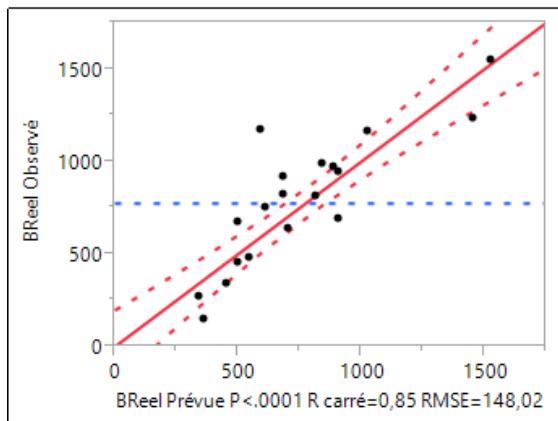


Figure 4.8 : Les valeurs observées en fonction des valeurs prévues pour le modèle du Ge5-1a

### Analyse des résidus du modèle de Ge5-1a

L’analyse des résidus indique un écart type de 153 kg. Le diagnostic du modèle de Ge5-1a indique une répartition régulière (Figure 4.9 a) et une distribution normale des résidus (Figure 4.9 b), cependant l’un indice DW de 2,87 et le taux d’autocorrélation de -0,16.

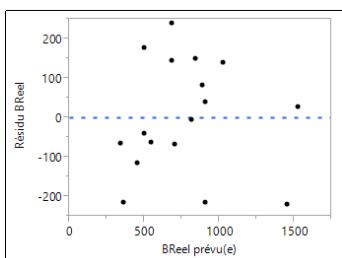


Figure 4.9 a : Graphique du résidu par ligne

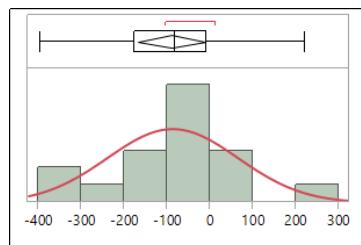


Figure 4.9 b : Distribution des résidus

### Le modèle de Qc7-1a (Arenosol cambique)

Le Tableau 4.15 montre que le sol **Qc7-1a** compte 44 observations, la moyenne des ces observations est  $773 \text{ kg MS.ha}^{-1}$ . La mise en œuvre de la procédure de sélection de variables du type pas à pas descendante au

niveau de cette région écologique a permis d’obtenir l’équation suivante

$$Y = -670,44 -10,07 \text{ ETAF} + 1,24 \text{ RAI} + 4197,70 \text{ MAX}$$

L’expression de la prévision contient 3 variables pertinentes (ETAF, RAI et MAX). La représentation des valeurs observées en fonction des valeurs prévues montre un  $R^2$  de 0,75 et un RMSE calculé de 201 kg.ha<sup>-1</sup> (Figure 4.9). Le modèle est caractérisé par un RMSE relatif de 22 %. L’estimation des paramètres du modèle montre des probabilités hautement significatives (Tableau 4.16).

Tableau 4.15 : Résumé de l’ajustement

Paramètres	Valeurs
$R^2$	0,75
$R^2$ ajusté	0,73
$R^2$ validation	0,70
RMSE calculé	201
RMSE validation	209
Moyenne de la réponse	643
Nombre d’observations	44

Tableau 4.16 : Estimations des coefficients

Terme	Estimation	Erreur standard	t ratio	Prob. >  t
Constante	-670,4453	137,8459	-4,86	<,0001*
ETAF	-10,07803	2,639674	-3,82	0,0005*
RAI	1,2440066	0,389485	3,19	0,0027*
MAX	4197,7037	557,6312	7,53	<,0001*

\*significatif au seuil de 10 pour 1000 ;

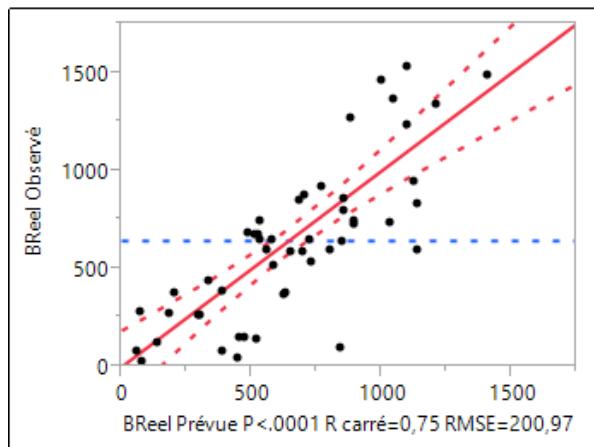


Figure 4.9 : Les valeurs observées en fonction des valeurs prévues pour le modèle du Qc7-1a

#### Analyse des résidus du modèle de Qc7-1a -

L’analyse des résidus indique un écart type de 193 kg. Le diagnostic du modèle de **Qc7-1a** indique une répartition régulière (Figure 4.10 a) et une distribution normale des résidus (Figure 4.10 b), comme l’attestent l’un indice DW de 1,83 et le taux d’autocorrélation de 0,05.

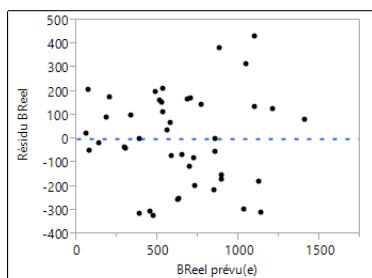


Figure 4.10 a : Graphique du résidu par ligne

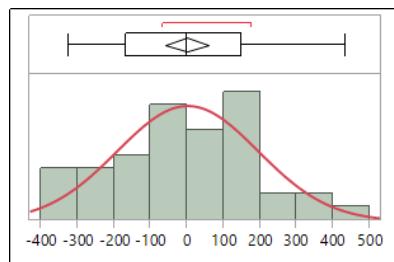


Figure 4.10 b : Distribution des résidus

## Le modèle de QI1-1a

Le Tableau 4.17 montre que l'unité de sol **QI1-1a** contient 85 observations, la moyenne des ces observations est  $650 \text{ kg MS.ha}^{-1}$ . La mise en œuvre de la procédure de sélection de variables du type pas à pas descendante au niveau de cette région écologique a permis d'obtenir l'équation suivante **Y= -4205,69 +24,34 EVA+ 104,11PEA +2438,14 MAX**

L'expression de la prévision contient 3 variables pertinentes (EVA, PVA et MAX). La représentation des valeurs observées en fonction des valeurs prévues montre un  $R^2$  de 0,85 et un RMSE calculé de  $207 \text{ kg.ha}^{-1}$  (Figure 4.11). Le modèle est caractérisé par un RMSE relatif de 31 %. L'estimation des paramètres du modèle montre des probabilités hautement significatives (Tableau 4.18).

Tableau 4.17 : Résumé de l'ajustement

Paramètres	Valeurs
$R^2$	0,76
$R^2$ ajusté	0,76
$R^2$ validation	0,74
RMSE calculé	207
RMSE validation	214
Moyenne de la réponse	650
Nombre d'observations	85

Tableau 4.18 : Estimations des coefficients

Terme	Estimation	Erreur standard	t ratio	Prob. >  t
Constante	-4096,214	641,8367	-6,38	<,0001*
EVA	24,00407	4,176959	5,75	<,0001*
PEA	101,17962	22,31335	4,53	<,0001*
MAX	2374,7752	530,3896	4,48	<,0001*

\*significatif au seuil de 10 pour 1000 ;

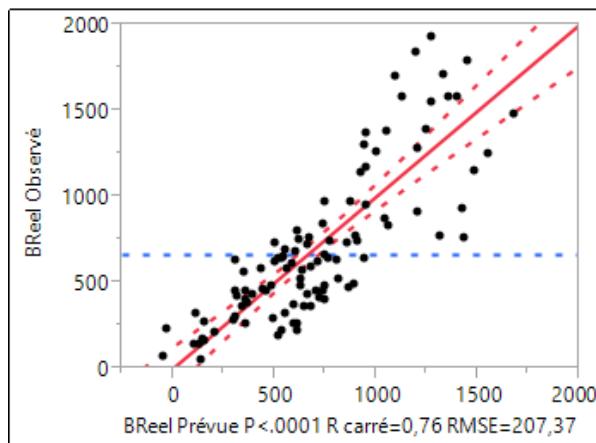


Figure 4.11 : Les valeurs observées en fonction des valeurs prévues pour le modèle du QI1-1a

### Analyse des résidus du modèle de QI1-1a

L'analyse des résidus indique un écart type de 203 kg. Le diagnostic du modèle de **QI1-1a** indique une répartition régulière (Figure 4.12 a) et une distribution normale des résidus (Figure 4.12 b), comme l'attestent l'un indice DW de 1,38 mais le taux d'autocorrélation de 0,29 est un peu élevé.

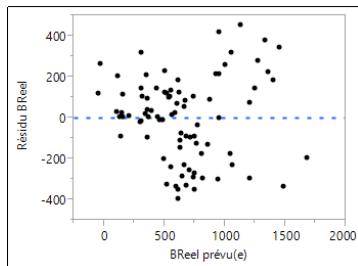


Figure 4.12 a : Graphique du résidu par ligne

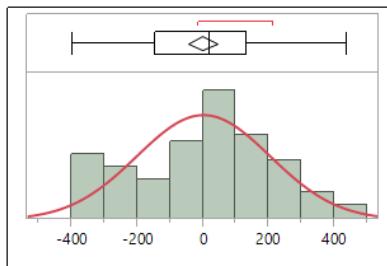


Figure 4.12 b : Distribution des résidus

### Le modèle de Re35-a

Le Tableau 4.19 montre que le sol **Re35-a** compte 39 observations, la moyenne des ces observations est  $391 \text{ kg MS.ha}^{-1}$ . La mise en œuvre de la procédure de sélection de variables du type pas à pas descendante au niveau de cette région écologique a permis d'obtenir l'équation suivante :  **$Y= -567 +12,94 \text{ VER} +2940,82 \text{ MAX}$** , l'**expression de la prévision comprend 2** variables pertinentes (VER et MAX) ;

La représentation des valeurs observées en fonction des valeurs prévues montre un  $R^2$  de 0,77 et un RMSE calculé de  $152 \text{ kg.ha}^{-1}$ . Le modèle est caractérisé par un RMSE relatif de 38 %. L'estimation des paramètres du modèle montre des probabilités hautement significatives (Tableau 4.20).

Tableau 4.19 : Résumé de l'ajustement

	Paramètres	Valeurs
R <sup>2</sup>		0,77
R <sup>2</sup> ajusté		0,76
R <sup>2</sup> validation		0,74
RMSE calculé		152
RMSE validation		157
Moyenne de la réponse		391
Nombre d'observations		39

Tableau 4.20 : Estimations des coefficients

Terme	Estimation	Erreur standard	t ratio	Prob. >  t
Constante	-567,01	108,94	-5,20	<,0001*
VER	12,94	5,49	2,36	0,0240*
MAX	2940,82	986,34	2,98	0,0051*

\*significatif au seuil de 10 pour 1000 ;

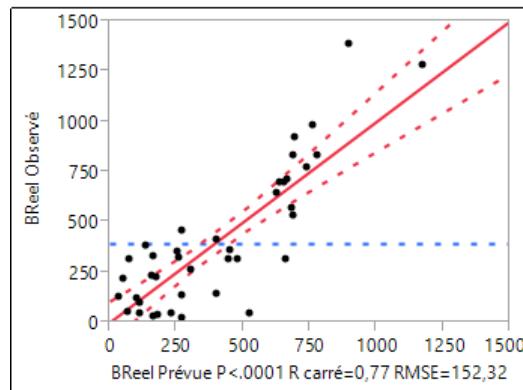


Figure 4.13 : Les valeurs observées en fonction des valeurs prévues pour le modèle du Re35-a

## Analyse des résidus du modèle de Re35-a

L'analyse des résidus indique un écart type de 148 kg. Le diagnostic du modèle de **Re35-a** indique une répartition régulière (Figure 4.14 a) et une distribution normale des résidus (Figure 4.14 b), comme l'attestent l'un indice DW de 1,72 et le taux d'autocorrélation de 0,11.

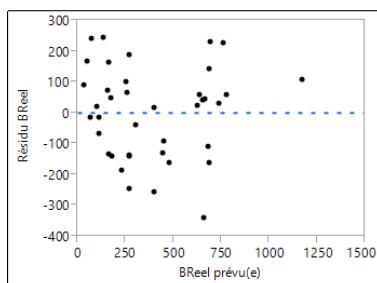


Figure 4.14 a : Graphique du résidu par ligne

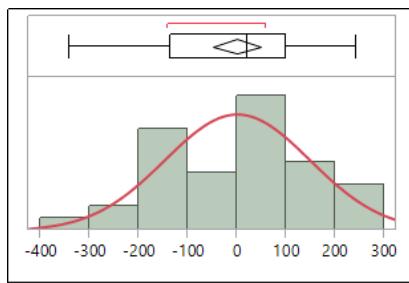


Figure 4.14 b : Distribution des résidus

## Le modèle de Qc1

Le Tableau 4.21 montre que le sol **Qc1** compte 22 observations, la moyenne des ces observations est 728 kg MS.ha<sup>-1</sup>. La mise en œuvre de la procédure de sélection de variables du type pas à pas descendante au niveau de cette région écologique a permis d'obtenir l'équation suivante **Y= - 758,60 +19,73 VER +2029,28 MAX**, l'**expression de la prévision** comporte 2 variables pertinentes (VER et MAX).

La représentation des valeurs observées en fonction des valeurs prévues montre un R<sup>2</sup> de 0,86 et un RMSE calculé de 190 kg.ha<sup>-1</sup> (Figure 4.15). Le modèle est caractérisé par un RMSE relatif de 26 %. L'estimation des paramètres du modèle montre des probabilités hautement significatives (Tableau 4.22).

Tableau 4.21 : Résumé de l'ajustement

	Paramètres	Valeurs
R <sup>2</sup>		0,86
R <sup>2</sup> ajusté		0,85
R <sup>2</sup> validation		0,81
RMSE calculé		190
RMSE validation		205
Moyenne de la réponse		728
Nombre d'observations		22

Tableau 4.22 : Estimations des coefficients

Terme	Estimation	Erreur standard	t ratio	Prob. >  t
Constante	-758,60	156,65	-4,84	0,0001*
VER	19,73	5,10	3,87	0,0010*
MAX	2029,28	982,51	2,07	0,0528

\*significatif au seuil de 10 pour 1000 ;

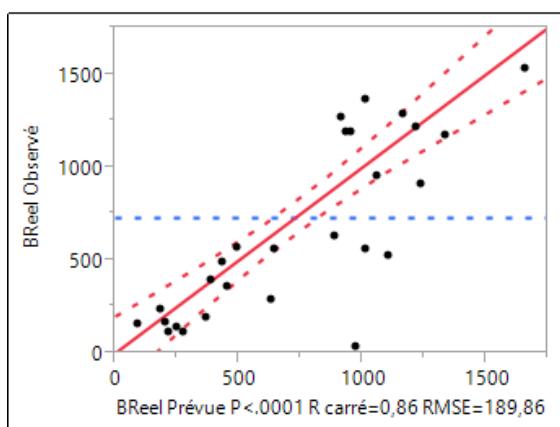


Figure 4.15 : Les valeurs observées en fonction des valeurs prévues pour le modèle du Qc1

## Analyse des résidus du modèle de Qc1

L'analyse des résidus indique un écart type de 180 kg. Le diagnostic du modèle de **Qc1** indique une répartition régulière (Figure 4.16 a) et une distribution normale des résidus (Figure 4.16 b), mais l'un indice DW de 2,35 et le taux d'autocorrélation un peu élevé (-0,27).

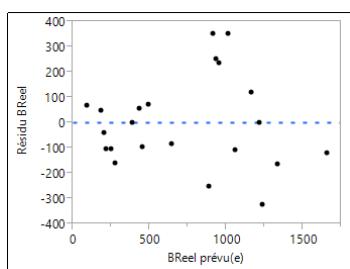


Figure 4.16 a : Graphique du résidu par ligne

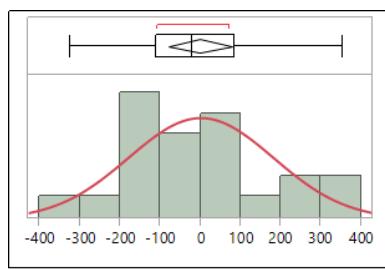


Figure 4.16 b : Distribution des résidus

## Annexe 4.3 : Modèle de régression linéaire multiple suivant les années

### Analyse de l'année 2001

La mise en œuvre de la procédure de sélection de variables du type pas à pas, selon sa variante descendante pour l'année 2001 a permis d'obtenir l'équation suivante ayant quatre variables pertinentes que sont : INDXNOR, ETAT, EVA, SVA.

$$Y = 72,43 -20,58 \text{ INDXNOR} + 6,40 \text{ ETAT} + 32,11 \text{ EVA} - 30,99 \text{ SVA}$$

L'estimation des paramètres du modèle liés à ces variables montre des probabilités significatives (Tableau 4.24). La corrélation des valeurs observées en fonction des valeurs prévues donne un R<sup>2</sup> de 0,83 et un RMSE de 158 kg.(Figure 4.17 ). Il est caractérisé par un RMSE relatif de 25 % (Tableau 4.25).

Tableau 4.24 : Estimation des coefficients d'ajustement

Variables	Estimation	Erreur standard	Rapport t	Prob.> t
Constante	72,43	592,16	0,12	0,9044
INDXNOR	-20,58	4,75	-4,33	0,0007**
ETAT	6,40	1,19	5,34	0,0001***
EVA	32,11	7,64	4,20	0,0009**
SVA	-30,99	14,12	-2,19	0,0456*

\*significatif au seuil de 5 pourcent ; \*\*significatif au seuil de 10 pour 10000 ;

\*\*\*significatif au seuil de 1 pour 10000

Tableau 4.25 : Résumé de l'ajustement

Paramètres	Valeurs
R <sup>2</sup>	0,83
R <sup>2</sup> ajusté	0,79
R <sup>2</sup> validation	0,67
RMSE Calibration	25 %
RMSE Validation	178,89
Moyenne de la réponse	624,71
Observations	19

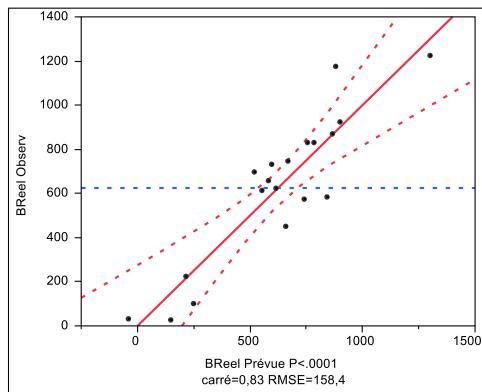


Figure 4.17 : Les valeurs observées en fonction des valeurs prévues pour le modèle de l'année 2001

## Analyse des résidus

Il ressort de l'analyse, que les résidus de l'année 2001 donnent un écart type de 140 kg (Tableau 4.26). Le diagnostic du modèle de cette année donne une répartition régulière (Figure 4.17 a), une distribution normale des résidus (Figure 4.17 b) comme l'atteste l'indice DW de 1,92.

Tableau 4.26 : Estimations des coefficients

Type	Coefficient	Estimation	Inférieur à 95 %	Supérieur à 95 %
Position	$\mu$	-0	-67	67
Dispersion	$\sigma$	140	106	207

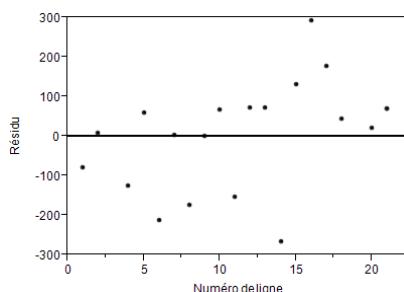


Figure 4.18 a : Graphique du résidu par ligne

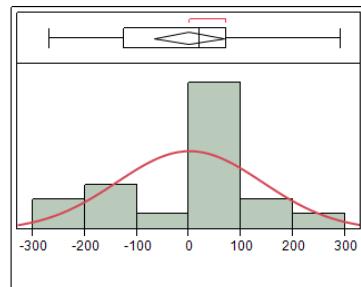


Figure 4.18 b : Distribution des résidus

## Analyse de l'année 2002

Pour l'année 2002, nous avons 27 observations avec une moyenne de 382,51 KgMS.ha<sup>-1</sup>. La mise en œuvre de la procédure de sélection de variables du type pas à pas à travers sa variante descendante sur les données de l'année 2002 a permis d'obtenir l'équation suivante ayant trois seules variables pertinentes que sont MAX, DRO et RAI.

$$Y = -201,3049 + 4006,7013 \text{ MAX} - 6,418996 \text{ DRO} - 0,557239 \text{ RAI}$$

L'estimation des paramètres du modèle liés à cette variable et la constante montre des probabilités significatives (Tableau 4.27). La corrélation des valeurs observées en fonction des valeurs prévues

montre un R<sup>2</sup> de 0,90 et un RMSE de 95 kg (Figure 4.19). Il est caractérisé par un RMSE relatif de 21 % (Tableau 4.28).

Tableau 4.27 : Estimation des coefficients du modèle

Terme	Estimation	Erreur standard	Rapport t	Prob.> t
Constante	-201,30	61,83	-3,26	0,0036**
max	4006,70	410,29	9,77	<, 0001***
DRO	-6,41	1,97	-3,25	0,0037**
rai	-0,55	0,24	-2,29	0,0320*

\*significatif au seuil de 5 pourcent ; \*\*significatif au seuil de 5 pour 1000 ; \*\*\*significatif au seuil de 1 pour 10000

Tableau 4.28 : Résumé de l'ajustement

Paramètres	Valeurs
R <sup>2</sup>	0,90
R <sup>2</sup> ajusté	0,88
Racine de l'erreur quadratique moyenne	21 %
RMSE validation	95 kg
Moyenne de la réponse	396 kg
Observations (ou sommes pondérées)	27

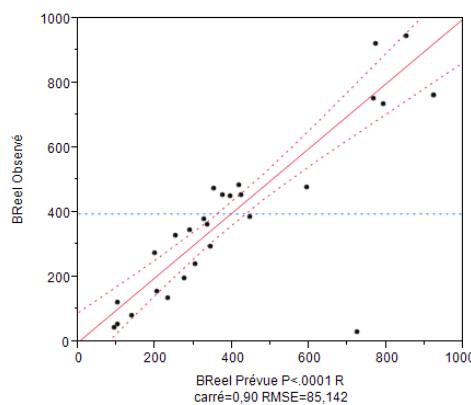


Figure 4.19 : Les valeurs observées en fonction des valeurs prévues pour l'année 2002

## Analyse des résidus

Il ressort de l'analyse, que les résidus de l'année 2002 indiquent un écart type de 145 kg (Tableau 4.29). Le diagnostic du modèle de cette année donne une répartition régulière (Figure 4.20 a) et une distribution normale des résidus (Figure 4.20 b).

Tableau 4.29 : Estimations des coefficients

Type	Coefficient	Estimation	Inférieur à 95 %	Supérieur à 95 %
Position	$\mu$	0	-42	42
Dispersion	$\sigma$	106	83,3	145

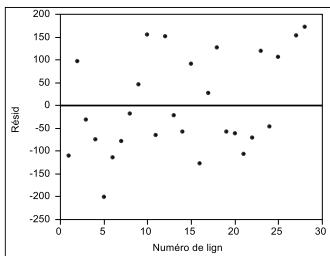


Figure 4.20 a : Graphique du résidu par ligne

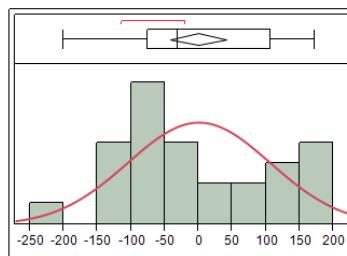


Figure 4.20 b : Distribution des résidus

## Analyse de l'année 2003

Pour cette année, nous ne disposons que de 11 observations. La mise en œuvre de la procédure de sélection de variables du type pas à pas, à travers sa variante descendante a permis d'obtenir l'équation suivante ayant trois variables pertinentes que sont MAX, WDEFI, HOR.  $\mathbf{Y = -13,62 + 9475,86 MAX - 50,53 WDEFI - 342,48 HOR}$

L'estimation des paramètres du modèle liés à ces variables et la constante montre des probabilités significatives (Tableau 4.30). La corrélation des valeurs observées en fonction des valeurs prévues montre un  $R^2$  de 0,96 et un RMSE de 100 kg (Figure 4.21). Il est caractérisé par un RMSE relatif de 11 % (Tableau 4.31).

Tableau 4.30 : Estimations des coefficients

Variables	Estimation	Erreur standard	Rapport t	Prob.> t
Constante	-13,62	302,50	-0,05	0,9653
Max	9475,86	1128,12	8,40	<, 0001**
WDEFI	-50,53	9,82	-5,15	0,0013*
HOR	-342,48	88,89	-3,85	0,0063*

\*significatif au seuil de 5 pour 1000 ; \*\*significatif au seuil de 1 pour 10000

Tableau 4.31 : Résumé de l'ajustement

	Paramètres	Valeurs
R <sup>2</sup>		0,96
R <sup>2</sup> ajusté		0,94
R <sup>2</sup> de la validation		0,89
Racine de l'erreur quadratique moyenne		11, %
RMSE validation		123,41 kg
Moyenne de la réponse		833,97
Observations (ou sommes pondérées)		11

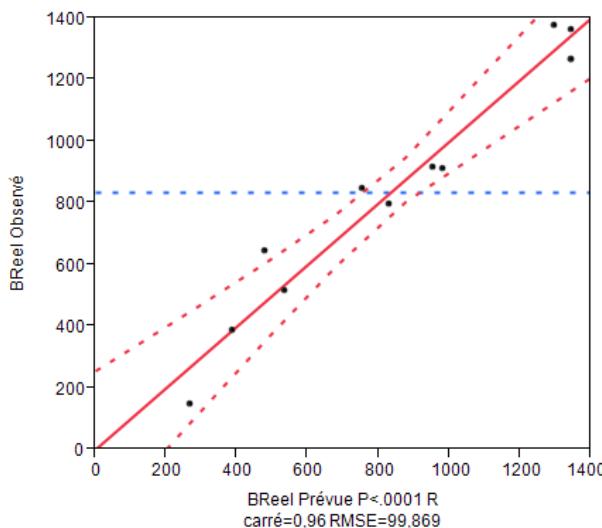


Figure 4.21 : Les valeurs observées en fonction des valeurs prévues pour l'année 2003

## Analyse des résidus

Il ressort de l'analyse que les résidus de l'année 2003 indiquent écart type de 84 kg (Tableau 4.32). Le diagnostic du modèle de cette année donne une répartition régulière (Figure 4.22 a) et une distribution normale des résidus (Figure 4.22 b) attestée par un indice DW de 1,38.

Tableau 4.32 : Estimations des coefficients

Type	Coefficient	Estimation	Inférieur à 95 %	Supérieur à 95 %
Position	$\mu$	0	-56	56
Dispersion	$\sigma$	84	58	147

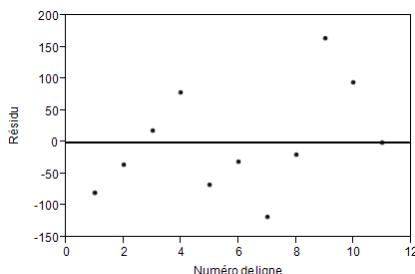


Figure 4.22 a : Graphique du résidu par ligne

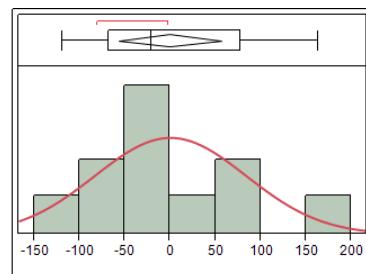


Figure 4.22 b : Distribution des résidus

## Analyse de l'année 2004

Les observations de l'année 2004 sont au nombre de 26 avec une moyenne de  $288 \text{ kg.MS.ha}^{-1}$ . La mise en œuvre de la procédure de sélection de variables du type pas à pas à travers sa variante descendante a permis d'obtenir l'équation suivante ayant une seule variable pertinente qu'est MAX.  **$Y = -224,94 + 2357,8581 \text{ MAX}$**

L'estimation des paramètres du modèle liés à cette variable et la constante montre des probabilités significatives (Tableau 4.33). La corrélation des valeurs observées en fonction des valeurs prévues donne un  $R^2$  de 0,60 et un RMSE de 124 kg. (Figure 4.23). Il est caractérisé par une erreur (RMSE) relative de 44 % (Tableau 4.34).

Tableau 4.33 : Estimations des coefficients

Variables	Estimation	Erreur standard	Rapport t	Prob.> t
Constante	-224,94	89,70	-2,51	<,01*
Max	2357,85	396,82	5,94	<, 0001*

\*significatif au seuil de 1 pour cent ; \* significatif au seuil de 1 pour 10000

Tableau 4.34 : Résumé de l'ajustement

Paramètres	Valeurs
R <sup>2</sup>	0,60
R <sup>2</sup> ajusté	0,58
R <sup>2</sup> Validation	0,50
Racine de l'erreur quadratique moyenne	44 %
RMSE Validation	130,50
Moyenne de la réponse	288
Observations (ou sommes pondérées)	26

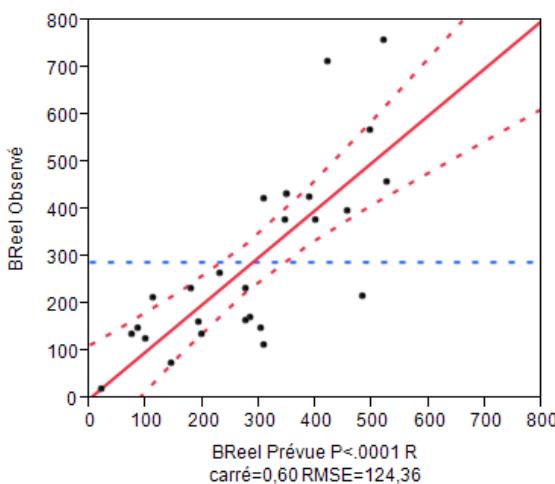


Figure 4.23 : Les valeurs observées en fonction des valeurs prévues pour l'année 2004

## Analyse des résidus

Il ressort de l'analyse que les résidus de l'année 2004 indiquent un écart type de 122 kg (Tableau 4.35). Le diagnostic du modèle de cette année montre une répartition régulière avec trois points excentrés (Figure 4.24 a) et une distribution normale des résidus (Figure 4.24 b).

Tableau 4.35 : Estimations des coefficients

Type	Coefficient	Estimation	Inférieur à 95 %	Supérieur à 95 %
Position	$\mu$	-0	-49	49
Dispersion	$\sigma$	122	96	168

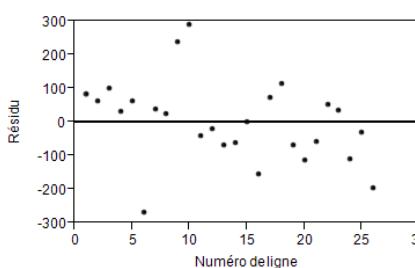


Figure 4.24 a : Graphique du résidu par ligne

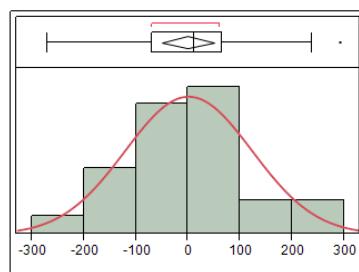


Figure 4.24 b : Distribution des résidus

## Analyse de l'année 2005

Les observations de l'année 2005 sont au nombre de 31 avec une moyenne de 713 kg MS/ha. La mise en œuvre de la méthode de sélection de variables du type pas-à-pas à travers la procédure descendante a permis d'obtenir l'équation suivante, avec trois variables pertinentes que sont INT, WDEFI et SAD.  $Y = -2639,25 + 698,67 \text{ INT} + 1,91 \text{ WDEFI} + 106,11 \text{ SDA}$ .

L'estimation des paramètres du modèle liés à ces variables et la constante montre des probabilités significatives (Tableau 4.36). la corrélation des valeurs observées en fonction des valeurs prévues donne

un R<sup>2</sup> de 0,82 et un RMSE de 194 kg.(Figure 4.25). Il est caractérisé par une erreur (RMSE) relative de 27 % (Tableau 4.37).

Tableau 4.36 : Estimations des coefficients

Variables	Estimation	Erreurs standard	Rapport t	Prob.> t
Constante	-2639,26	968,23	-2,73	0,01**
int	698,68	97,49	7,17	<, 0001****
WDEFR	1,91	0,96	1,98	0,05*
sda	106,12	35,99	2,95	0,006***

\*significatif au seuil de 5 pour 100 ; \*\*significatif au seuil de 10 pour 1000 ; \*\*\*significatif au seuil de 10 pour 1000 ; \*\*\*\*significatif au seuil de 1 pour 10000

Tableau 4.37 : Résumé de l'ajustement

Paramètres	valeurs
R <sup>2</sup>	0,82
R <sup>2</sup> ajusté	0,80
R <sup>2</sup> validation	0,77
Racine de l'erreur quadratique moyenne	27 %
RMSE validation	205,37
Moyenne de la réponse	713,42
Nombre d'observations	31

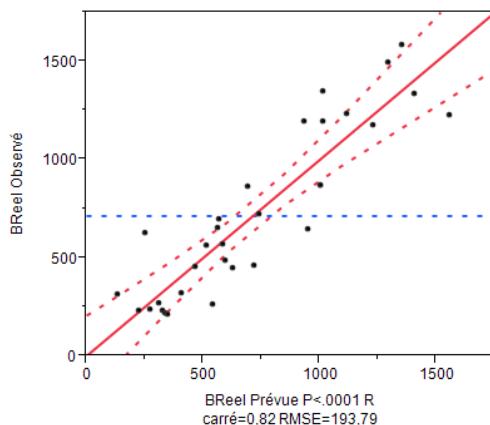


Figure 4.25 : Les valeurs observées en fonction des valeurs prévues pour l'année 2005

## Analyse des résidus

Il ressort de l'analyse que les résidus de l'année 2005 indiquent un écart type de 184 kg (Tableau 4.38). Le diagnostic du modèle de cette année montre une répartition régulière (Figure 4.26 a) et une distribution normale des résidus (Figure 4.26 b) comme l'attestent l'indice DW de 2 un taux d'autocorrélation de -0,009.

Tableau 4.38 : Estimations des coefficients

Type	Coefficient	Estimation	Inférieur à 95 %	Supérieur à 95 %
Position	$\mu$	0	-67	67
Dispersion	$\sigma$	184	147	246

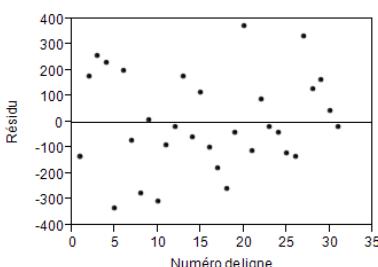


Figure 4.26 a : Graphique du résidu par ligne

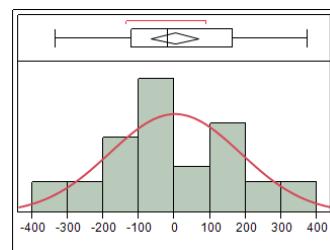


Figure 4.26 b : Distribution des résidus

## Analyse de l'année 2006

Pour cette année, nous disposons de 24 observations ayant une moyenne de 826 kg.MS.ha<sup>-1</sup>. La mise en œuvre de la procédure de sélection de variables du type pas à pas par la méthode descendante a permis d'obtenir l'équation suivante ayant quatre variables pertinentes que sont INT, CUM, EVA, SVA.  $Y = -858,74 -686,75 \text{ INT} + 76,18 \text{ CUM} + 66,51 \text{ EVA} - 44,45 \text{ SVA}$

L'estimation des paramètres du modèle liés à ces variables et la constante donne des probabilités significatives (Tableau 4.39). la corrélation des valeurs observées en fonction des valeurs prévues montre un R<sup>2</sup> de 0,87 et un RMSE de 220 kg.(Figure 4.27). Il est caractérisé par une erreur (RMSE) relative de 29 % (Tableau 4.40).

Tableau 4.39 : Estimations des coefficients

Terme	Estimation	Erreurs Standard	Rapport t	Prob.> t
Constante	-858,75	760,78	-1,13	0,27
INT	-686,76	327,37	-2,10	0,04*
CUM	76,19	19,14	3,98	0,0008***
EVA	66,51	17,21	3,87	0,001**
SVA	-44,46	10,38	-4,28	0,0004****

\*significatif au seuil de 5 pour 100 ; \*\*significatif au seuil de 1 pour 1000 ;

\*\*\*significatif au seuil de 10 pour 10000 ; \*\*\*\*significatif au seuil de 5 pour 10000

Tableau 4.40 : Résumé de l'ajustement

Paramètres	Valeurs
R <sup>2</sup>	0,87
R <sup>2</sup> ajusté	0,85
R <sup>2</sup> validation	0,80
Racine de l'erreur quadratique moyenne	29 %
RMSE validation	244
Moyenne de la réponse	825,98
Observations (ou sommes pondérées)	24

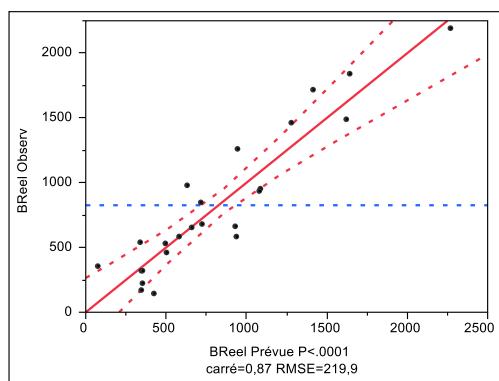


Figure 4.27 : Les valeurs observées en fonction des valeurs prévues pour l'année 2006

### Analyse des résidus

Il ressort de l'analyse, que les résidus de l'année 2006 donnent un écart type de 200 kg (Tableau 4.41). Le diagnostic du modèle de cette année

montre une répartition régulière (Figure 4.28 a) et une distribution normale des résidus (Figure 4.28 b), un indice DW de 2,43.

Tableau 4.41 : Estimations des coefficients

Type	Coefficient	Estimation	Inférieur à 95 %	Supérieur à 95 %
Position	$\mu$	0	-84	84,43
Dispersion	$\sigma$	200	155	280,48

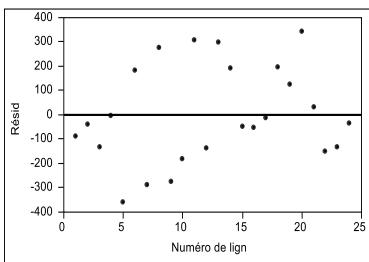


Figure 4.28 a : Graphique du résidu par ligne

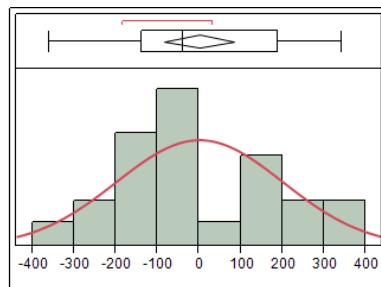


Figure 4.28 b : Distribution des résidus

### Analyse de l'année 2007

Pour cette année, nous disposons de 23 observations ayant une moyenne de  $1119 \text{ kg.MS.ha}^{-1}$ . La mise en œuvre de la procédure de sélection de variables du type pas à pas à travers descendante a permis d'obtenir l'équation suivante ayant deux variables pertinentes que sont INT, TWR.  $\mathbf{Y = -1495,72 + 505,03 INT + 2,16TWR}$

L'estimation des paramètres du modèle liés à ces variables et la constante montre des probabilités significatives (Tableau 4.42). la corrélation des valeurs observées en fonction des valeurs prévues donne un  $R^2$  de 0,61 et un RMSE de 100 kg.(Figure 4.29). Il est caractérisé par une erreur (RMSE) relative de 20 % (Tableau 4.43).

Tableau 4.42 : Estimations des coefficients

Variables	Estimation	Erreurs standard	Rapport t	Prob.> t
Constante	-1495,72	682,19	-2,19	0,04*
INT	505,03	90,25	5,60	<, 0001**
TWR	2,16	1,17	1,85	0,0786

\*significatif au seuil de 5 pour 100 ; \*\* significatif au seuil de 1 pour 10000

Tableau 4.43 : Résumé de l'ajustement

Paramètres	Valeurs
R <sup>2</sup>	0,61
R <sup>2</sup> ajusté	0,57
R <sup>2</sup> validation	0,53
Racine de l'erreur quadratique moyenne	20 %
RMSE Validation	226,29 kg
Moyenne de la réponse	1119,17
Observations (ou sommes pondérées)	23

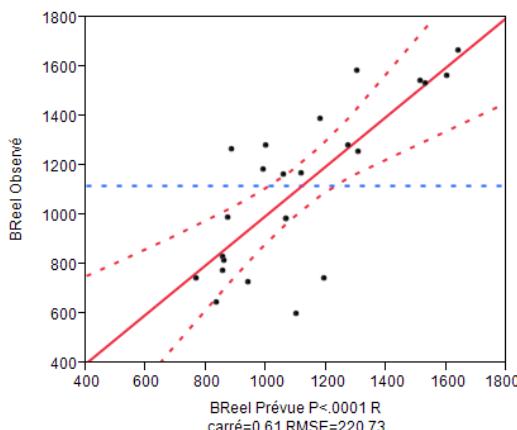


Figure 4.29 : Les valeurs observées en fonction des valeurs prévues pour l'année 2007

### Analyse des résidus

Il ressort de l'analyse, que les résidus de l'année 2007 indiquent un écart type de 210 kg (Tableau 4.44). Le diagnostic du modèle de cette année montre une répartition régulière (Figure 4.30 a) et une distribution normale des résidus (Figure 4.30 b), un indice DW de 2,4

Tableau 4.44 : Estimations des coefficients

Type	Coefficient	Estimation	Inférieur à	Supérieur à
			95 %	95 %
Position	$\mu$	0	-91	91
Dispersion	$\sigma$	210	163	298

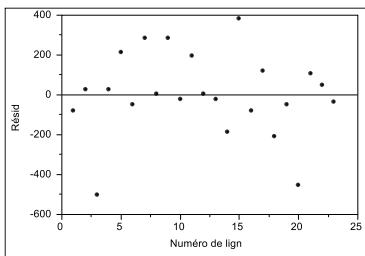


Figure 4.30 a : Graphique du résidu par ligne

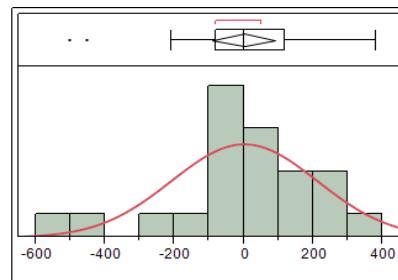


Figure 4.30 b : Distribution des résidus

### Analyse de l'année 2008

Pour cette année, nous disposons de 33 observations ayant une moyenne de 483 Kg MS.ha<sup>-1</sup>. La mise en œuvre de la procédure de sélection de variables du type pas à pas à travers la méthode descendante a permis d'obtenir l'équation suivante ayant une seule variable pertinente qu'est le PVA. **Y = -1125,59 + 14,50 PVA**

L'estimation des paramètres du modèle liés à cette variable et la constante montre des probabilités significatives (Tableau 4.45). La corrélation des valeurs observées en fonction des valeurs prévues montre un R<sup>2</sup> de 0,59 et un RMSE de 283 kg.(Figure 4.31). Il est caractérisé par une erreur (RMSE) relative de 58 % (Tableau 4.46).

Tableau 4.45 : Estimations des coefficients

Terme	Estimation	Erreur standard	Rapport t	Prob.> t
Constante	-1125,59	248,11	-4,54	<, 0001*
PVA	14,50	2,20	6,62	<, 0001*

\*significatif au seuil de 1 pour 10000

Tableau 4.46 : Résumé de l'ajustement

Paramètres	Valeurs
R <sup>2</sup>	0,58
R <sup>2</sup> ajusté	0,57
R <sup>2</sup> validation	0,53
Racine de l'erreur quadratique moyenne	58 %
RMSE Validation	292,11
Moyenne de la réponse	483,20
Observations (ou sommes pondérées)	33

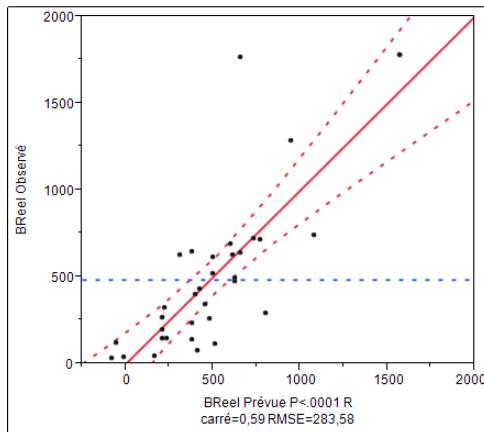


Figure 4.31 : Les valeurs observées en fonction des valeurs prévues pour l'année 2008

### Analyse des résidus

Il ressort de l'analyse, que les résidus de l'année 2008 donnent un écart type de 279 kg (Tableau 4.47). Le diagnostic du modèle de cette année montre une répartition régulière (Figure 4.32 a) et une distribution normale des résidus (Figure 4.32 b).

Tableau 4.47 : Estimations des coefficients

Type	Coefficient	Estimation	Inférieur à 95 %	Supérieur à 95 %
Position	$\mu$	0	-99	99
Dispersion	$\sigma$	279	224	369

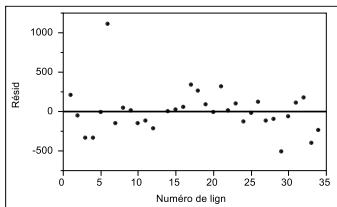


Figure 4.32 a : Graphique du résidu par ligne

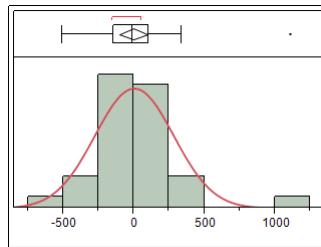


Figure 4.32 b : Distribution des résidus

### Analyse de l'année 2009

Pour l'année 2009, nous ne disposons que de 16 observations ayant une moyenne de  $490 \text{ kg.MS.ha}^{-1}$ . La mise en œuvre de la procédure de sélection de variables du type pas à pas à travers la méthode descendante a permis d'obtenir l'équation suivante ayant deux variables pertinentes que sont ETAF et EVA.  $Y = -1551,76 + 3,61 \text{ ETAF} + 31,13 \text{ EVA}$

L'estimation des paramètres du modèle liés à ces variables et la constante montre des probabilités significatives (Tableau 4.48). la corrélation des valeurs observées en fonction des valeurs prévues donne un  $R^2$  de 0,82 et un RMSE de 112 kg.(Figure 4.33). Il est caractérisé par une erreur (RMSE) relative de 22 % (Tableau 4.49).

Tableau 4.48 : Estimations des coefficients

Variables	Estimation	Erreurs standard	Rapport t	Prob.> t
Constante	-1551,76	300,12	-5,17	<,0002**
ETAF	3,61	1,64	2,20	<,04*
EVA	31,13	5,04	6,18	<,0001***

\*significatif au seuil de 5 pour 100 ; \*\*significatif au seuil de 5 pour 10000 ; \*\*\*significatif au seuil de 1 pour 10000

Tableau 4.49 : Résumé de l'ajustement

Paramètres	Valeurs
R <sup>2</sup>	0,82
R <sup>2</sup> ajusté	0,79
R <sup>2</sup> validation	0,75
Racine de l'erreur quadratique moyenne	22 %
RMSE Validation	119,54
Moyenne de la réponse	489,82
Observations (ou sommes pondérées)	16

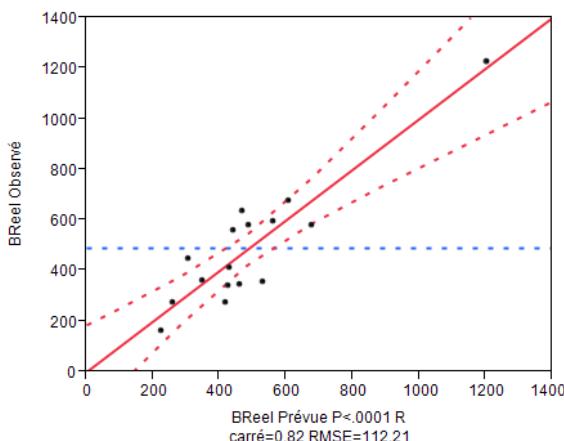


Figure 4.33 : Les valeurs observées en fonction des valeurs prévues pour l'année 2009

## Analyse des résidus

Il ressort de l'analyse, que les résidus de l'année 2009 donnent un écart type de 104 kg (Tableau 4.50). Le diagnostic du modèle de cette année donne une répartition régulière (Figure 4.34 a) et une distribution normale de ces résidus (Figure 4.34 b) comme l'atteste l'un indice DW de 2.

Tableau 4.50 : Estimations des coefficients

Type	Coefficient	Estimation	Inférieur à 95 %	Supérieur à 95 %
Position	$\mu$	0	-56	56
Dispersion	$\sigma$	104	77	162

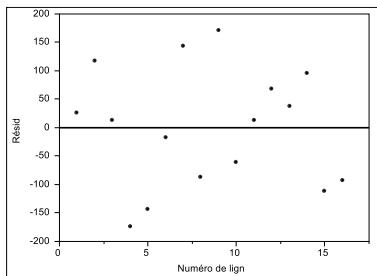


Figure 4.34 a : Graphique du résidu par ligne

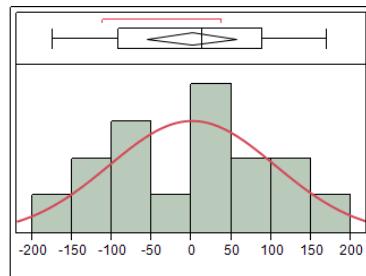


Figure 4.34 b : Distribution des résidus

## Analyse de l'année 2010

Pour cette année 2010, nous disposons de 32 observations ayant une moyenne de  $1005 \text{ Kg.MS.ha}^{-1}$ . La mise en œuvre de la procédure de sélection de variables du type pas à pas à travers la méthode descendante a permis d'obtenir l'équation suivante avec deux variables pertinentes que sont INT, SDA.  **$Y = -4931,74 + 1057,45 \text{ INT} + 151,27 \text{ SDA}$**

L'estimation des paramètres du modèle liés à ces variables et la constante montre des probabilités significatives (Tableau 4.51). la corrélation des valeurs observées en fonction des valeurs prévues montre un  $R^2$  de 0,68 et un RMSE de 348 kg.(Figure 4.35). Il est caractérisé par une erreur (RMSE) relative de 34 % (Tableau 4.52).

Tableau 4.51 : Estimations des coefficients

Variables	Estimation	Erreurs standard	Rapport t	Prob.> t
Constante	-4931,745	1719,608	-2,87	0,007**
INT	1057,4551	152,2813	6,94	<,0001***
SDA	151,27711	71,16638	2,13	0,04*

\*significatif au seuil de 5 pour 100 ; \*\*significatif au seuil de 1 pour 100 ; \*\*\*significatif au seuil de 1 pour 10000

Tableau 4.52: Résumé de l'ajustement

Paramètres	Valeurs
R <sup>2</sup>	0,68
R <sup>2</sup> ajusté	0,67
R <sup>2</sup> Validation	0,62
Racine de l'erreur quadratique moyenne	34 %
RMSE validation	361
Moyenne de la réponse	1004,98
Observations (ou sommes pondérées)	32

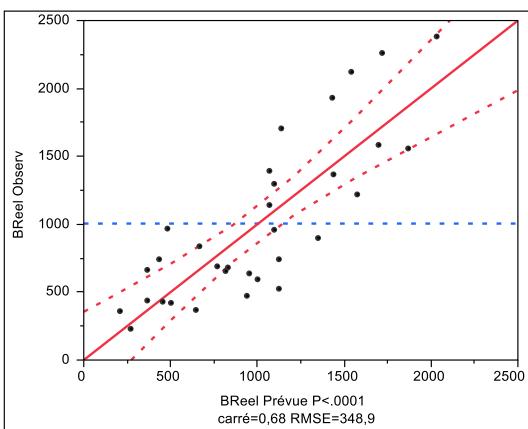


Figure 4.35 : Les valeurs observées en fonction des valeurs prévues pour l'année 2010

## Analyse des résidus

Il ressort de l'analyse, que les résidus de l'année 2010 donnent un écart type de 337 kg (Tableau 4.53). Le diagnostic du modèle de cette année montre une répartition régulière (Figure 4.36 a) et une distribution normale des résidus (Figure 4.36 b), un indice DW de 1,99.

Tableau 4.53 : Estimations des coefficients

Type	Coefficient	Estimation	Inférieur à 95 %	Supérieur à 95 %
Position	$\mu$	-0	-122	122
Dispersion	$\sigma$	337	271	449

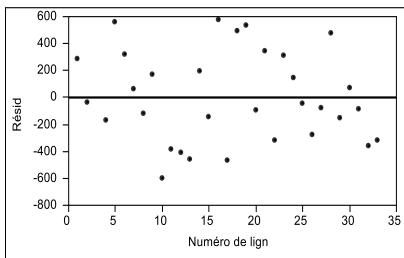


Figure 4.36 a : Graphique du résidu par ligne

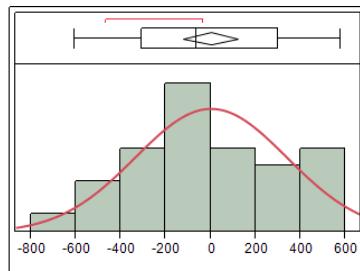


Figure 4.36 b : Distribution des résidus

## Analyse de l'année 2011

Pour cette année, nous disposons de 39 observations. La mise en œuvre de la procédure de sélection de variables du type pas à pas par la méthode descendante pour l'année 2011 a permis d'obtenir l'équation suivante avec quatre variables pertinentes que sont MAX, WDEFI, ETAF, PEA.  **$Y = 1850,33 + 2329,23 \text{ MAX} - 27,77 \text{ WDEFI} + 6,63 \text{ ETAF} - 89,84 \text{ PEA}$**

L'estimation des paramètres du modèle liés à ces variables et la constante montre des probabilités significatives (Tableau 4.54). la corrélation des valeurs observées en fonction des valeurs prévues montre un  $R^2$  de 0,76 et un RMSE de 138 kg.(Figure 4.37). Il est caractérisé par une erreur (RMSE) relative de 29 % (Tableau 4.55).

Tableau 4.54 : Estimations des coefficients

Variables	Estimation	Erreurs standard	Rapport t	Prob.> t
Constante	1850,34	679,48	2,72	0,01**
Max	2329,23	307,74	7,57	<, 0001****
WDEFI	-27,78	11,35	-2,45	0,02*
ETAF	6,63	2,06	3,21	0,002**
PEA	-89,85	27,81	-3,23	0,002**

\*significatif au seuil de 5 pour 100 ; \*\*significatif au seuil de 1 pour 100 ;  
\*\*\*significatif au seuil de 5 pour 1000 ; \*\*\*\*significatif au seuil de 1 pour 10000

Tableau 4.55 : Résumé de l'ajustement

Paramètres	Valeurs
R <sup>2</sup>	0,76
R <sup>2</sup> ajusté	0,73
R <sup>2</sup> validation	0,65
Racine de l'erreur quadratique moyenne	29 %
RMSE validation	154,5 kg.ha <sup>-1</sup>
Moyenne de la réponse	482,14 kg.ha <sup>-1</sup>
Nombre d'observations	39

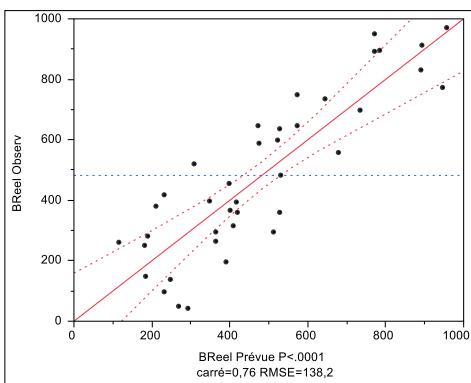


Figure 4.37 : Les valeurs observées en fonction des valeurs prévues pour l'année 2011

## Analyse des résidus

Il ressort de l'analyse, que les résidus de l'année 2011 donnent un écart type de 131 kg (Tableau 4.56). Le diagnostic du modèle de cette année montre une répartition régulière (Figure 4.38 a) et une distribution normale des résidus (Figure 4.38 b), un indice DW de 1,92

Tableau 4.56 : Estimations des coefficients

Type	Coefficient	Estimation	Inférieur à 95 %	Supérieur à 95 %
Position	$\mu$	-0	-42	42
Dispersion	$\sigma$	131	107	169

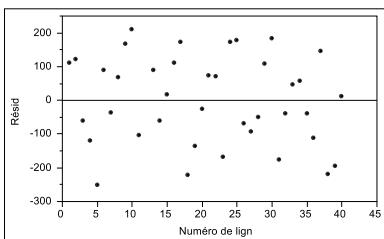


Figure 4.38 a : Graphique du résidu par ligne

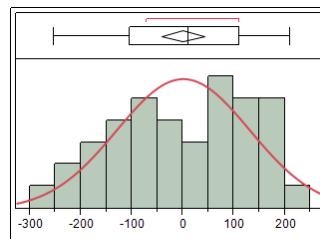


Figure 4.38 b : Distribution des résidus

## Analyse de l'année 2012

Pour cette année 2012, nous disposons de 24 observations ayant une moyenne de 1277 KG.MS.ha<sup>-1</sup>. La mise en œuvre de la procédure de sélection de variables du type pas à pas à travers la variante descendante a permis d'obtenir l'équation suivante ayant trois variables pertinentes que sont INDXLATEST, WDEFR, PVA.

$$\mathbf{Y = 720,92 -16,19 INDXLATEST 6,95 WDEFR 19,71 PVA}$$

L'estimation des paramètres du modèle liés à ces variables et la constante montre des probabilités significatives (Tableau 4.57). la corrélation des valeurs observées en fonction des valeurs prévues montre un R<sup>2</sup> de 0,78 et un RMSE de 308kg. (Figure 4.39). Il est caractérisé par une erreur (RMSE) relative de 24 % (Tableau 4.58).

Tableau 4.57 : Estimations des coefficients

Variables	Estimation	Erreurs standard	Rapport t	Prob.> t
Constante	720,92	785,57	0,92	<,37
INDXLATES	-16,19	7,58	-2,14	<,0451*
T				
WDEFR	6,95	2,14	3,25	<,0040**
PVA	19,72	2,88	6,85	<,0001***

\*significatif au seuil de 5 pour 100 ; \*\*significatif au seuil de 5 pour 1000 ;

\*\*\*significatif au seuil de 1 pour 10000

Tableau 4.58 : Résumé de l'ajustement

	Paramètres	Valeurs
R <sup>2</sup>		0,78
R <sup>2</sup> ajusté		0,74
R <sup>2</sup> validation		0,65
Racine de l'erreur quadratique moyenne		24 %
RMSE validation		347,88
Moyenne de la réponse		1276,77
Observations (ou sommes pondérées)		24

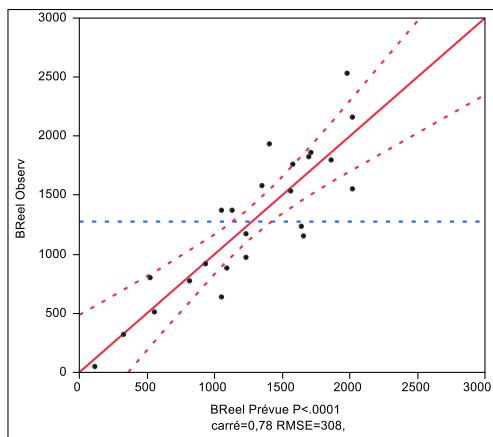


Figure 4.39 : Les valeurs observées en fonction des valeurs prévues pour l'année 2012

## Analyse des résidus

Il ressort de l'analyse, que les résidus de l'année 2012 indiquent un écart type de 288 kg (Tableau 4.59). Le diagnostic du modèle de cette année montre une répartition régulière (Figure 4.40 a) et une distribution normale des résidus (Figure 4.40 b) attesté par un indice DW de 2,10.

Tableau 4.59 : Estimations des coefficients

Type	Coefficient	Estimation	Inférieur à 95 %	Supérieur à 95 %
Position	$\mu$	-0	-121	121
Dispersion	$\sigma$	288	224	404

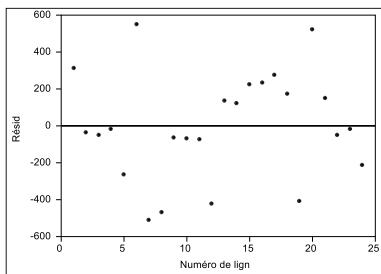


Figure 4.40 a : Graphique du résidu par ligne

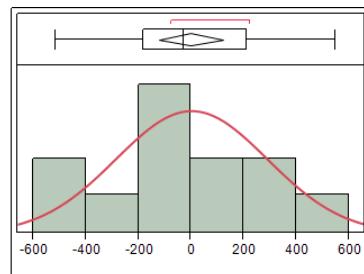


Figure 4.40 b : Distribution des résidus

**Annexe 4.4: L'évolution du bilan fourrager sur la période de 2000 à 2013 au Niger (Ces déficits sont consécutifs à des sécheresses).**

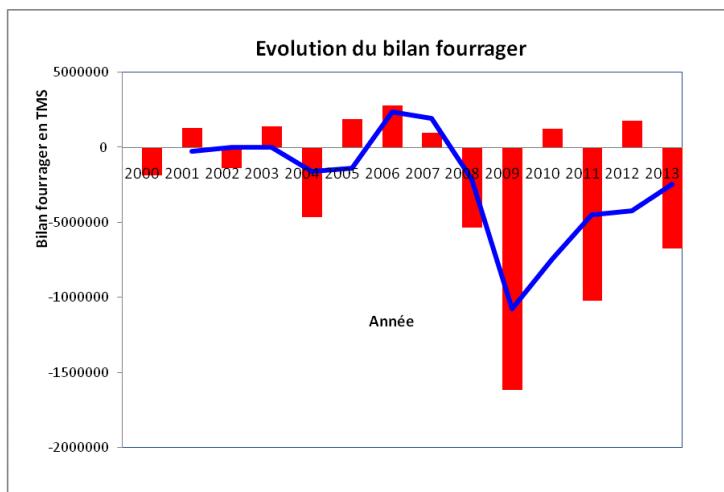


Figure 4.41 : L'évolution du bilan fourrager sur la période de 2000 à 2013 au Niger (MEIA)