

DETERMINANTS CLIMATIQUES DE LA VARIABILITE INTERANNUELLE DES RENDEMENTS EN CEREALES ET PREVISION PRECOCE. APPLICATION A LA PROVINCE DE SETTAT (MAROC)

Lionel Jarlan¹, Jaouad Abaoui², Benoit Duchemin¹, Yves Tourre³, Abdelaziz Ouldbba², Sylvain Mangiarotti¹, Hakim Kharrou⁴, Saïd Khabba⁵, Riad Ballaghi⁶, Michel Le Page¹, Abdelghani Chehbouni¹

1 : Centre d'Etudes Spatiales de la Biosphère, Toulouse, France, lionel.jarlan@cesbio.cnes.fr

2 : Direction de la Météorologie Nationale, Casablanca, Maroc

3 : LDEO, Columbia University, New York, Etats-Unis

4 : Office Regional de Mise en Valeur de l'Eau Agricole du Haouz, Marrakech, Maroc

5 : Faculté des Sciences de Semlalia, Université Cadi Ayyad, Marrakech, Maroc

6 : Institut National des Sciences Agronomiques, Rabat, Maroc

Résumé

Compte-tenu de l'importance de la production de blé pour l'économie marocaine, les décideurs doivent disposer d'informations précises sur la production de blé tôt dans la saison afin d'anticiper les besoins en importation. Dans cette étude, nous avons analysé (1) les rendements en blé tendre (1983-2008) de la province de Settat (2) les précipitations et les températures mesurées sur la station de Settat ; (3) les indices de végétation NDVI à 10 jours du capteur AVHRR ; (4) plusieurs indices témoignant de la dynamique atmosphérique et les champs spatio-temporels de géopotentiels à 500 hPa et (5) plusieurs indices témoignant de la dynamique des températures de surface des océans (TSO) ainsi que les champs spatio-temporels de TSO. Il a été démontré que les relations entre les précipitations et les températures à plusieurs stades critiques du développement du blé, d'une part, et les rendements en grain, d'autre part, sont également significatives au niveau de la province agricole. L'Oscillation Nord Atlantique au tout début de la saison de croissance en novembre est également liée aux rendements. Un motif spatial de corrélation intéressant a également été mis en évidence entre les champs de TSO dans l'Atlantique Equatorial et Tropical plus tôt dans l'année, et les rendements. Les potentialités de prévisions saisonnières des rendements en grains basée sur le climat à grande échelle sont analysées en comparant deux approches de régression (MLR, MVS). La prévisibilité est étudiée pour des modèles construits de décembre à mars avec des critères statistiques draconiens pour limiter la complexité des modèles. Le meilleur modèle est naturellement obtenu pour une prévision en mars avec un coefficient ajusté r^2 et une erreur quadratique moyenne de, respectivement, 0,92 et 2,0 quintaux/hectare. En outre, l'approche MVS surpasse MLR pour tous les cas étudiés. Notre conclusion est que l'utilisation d'indices et de variables témoignant de la dynamique atmosphérique et océanique à l'échelle régionale peut permettre une prévision plus précoce qu'avec les prédicteurs classiquement utilisés pour le développement de ce type de modèles.

Mots clés : NDVI, Rendements, Blé, Oscillation Nord Atlantique, Régression Linéaire, Machine à Vecteur de Support.

Abstract

Regarding the importance of wheat production on Moroccan economy, it is important for decision makers to be able to rely on accurate information on wheat production early in the season in order to pre-empt importation needs. We analyzed de-trended (1) wheat yields (1983-2008) from the Settat province (2) monthly rainfall and 10-days temperature from the surrounding weather station; (3) 10-days Vegetation Indices NDVI from the AVHRR Sensor; (4) monthly atmospheric climate indices and monthly 500hPa geopotentials fields and (5) monthly Sea Surface Temperature (SST) fields and indices. It has been demonstrated that the relationship between rainfall and temperature at critical stages of wheat development and wheat yields already observed at the plot scale is also significant at the provincial level. The ONA in the early beginning of the growing season in November has also been found to be significantly related to yields. Interesting correlation patterns for yields predictability with higher lead time are highlighted with SST fields in the equatorial Atlantic during October. The potentiality of using this climate information together with the other more classical factors (temperature, rainfall and NDVI) is evaluated through the comparison of multiple linear regression (MLR) and Support Vector Regression (SVR). The predictability has been studied at different lead time from December to March with draconian statistical criteria for model selection in order to limit the model complexity. The best model is naturally obtained for a prediction from March. Adjusted coefficient r^2 and root mean square error are 0.92 and 2.0 qt/ha, respectively. In addition, the SVR approach outperforms the MLR models for all studied cases. Our conclusion is that the use of regional climate indices and variables may increase the lead time and the skill of wheat yields forecasting in Morocco.

Keywords : NDVI, Rendements, Wheat, North Atlantic Oscillation, Linear Regression, Support Vector Machine.

Revue Française de Photogrammétrie et de Télédétection

n° 204 (octobre 2013)

1. Introduction

La croissance démographique et l'augmentation connexe de la demande alimentaire ont conduit à un besoin impérieux pour des systèmes efficaces de prévision de la production agricole. L'agriculture pluviale domine la production mondiale avec 80% des terres arables et 60% de la production (Harris, 1991). Au Maroc, les céréales constituent la principale culture agricole : le blé tendre représente 54% de la production totale des céréales du pays. Afin de régulariser le marché des céréales, il est important d'anticiper l'importation dans le cas de production insuffisante pour la consommation intérieure ou pour le stockage des semences. La prévision saisonnière constitue, dans ces cas, un outil bien adapté à la prise de décision.

Les développements méthodologiques pour la prévision saisonnière des rendements agricoles ont suivi deux approches principales :

1. des modèles empiriques liant les rendements à des indicateurs des conditions environnementales (température, précipitations, indices de végétation) ;
2. des modèles de cultures alimentés par la prévision climatique saisonnière (Baigorria et al., 2008 ; Phillips et al., 1998).

Concernant ces derniers outils, ils souffrent néanmoins

1. de la forte demande en données d'entrée des modèles (Brisson et al., 2009) ;
2. des incertitudes sur les prévisions climatiques saisonnières (Hansen et al., 1999) et
3. de la disparité des échelles spatiale et temporelle entre les prévisions climatiques et les attentes des modèles de croissance (Hansen, 2004).

Les limitations évoquées ci-dessus conduisent la grande majorité des pays à adopter l'approche empirique. Sur le Maroc en particulier, Balaghi, et al. (2008) ont démontré que des rendements prévisionnels de bonne qualité peuvent être obtenus environ 1 mois avant la récolte.

D'une part, la plupart de ces études reposent sur la régression linéaire multiple classique tandis que les liens entre les rendements et les prédicteurs potentiels sont susceptibles d'être non linéaires. Un candidat potentiel pour une régression non linéaire est la machine à vecteurs de support (MVS) qui est un domaine de recherche en plein essor depuis quelques années (Kecman, 2001). La MVS est une méthode d'apprentissage automatique basée sur la théorie de la minimisation du risque structurel de Vapnik (1989). Les MVS ont récemment été appliquées à la désagrégation de scénarii climatiques (Ghosh et Mujumdar, 2009) et la prévision des débits (Kalra et Ahmad, 2009). A notre connaissance, cette approche n'a jamais été appliquée à la prévision précoce des rendements céréaliers.

D'autre part, la plupart de ces études reposent sur une série de prédicteurs locaux comme les températures, les précipitations et, pour certaines, une information sur l'activité du couvert au travers des indices de végétation mesurés depuis l'espace. C'est notamment le cas des travaux de Balaghi et al. (2008) déjà évoqués. Les liens indirects entre croissance des cultures et pseudo-oscillations climatiques à grande échelle, marqueurs indirects de conditions météorologiques locales, ne sont généralement pas concomitants à cause de l'éloignement de ces phénomènes qui influencent ces conditions par des télé-connections et en raison de la réponse de la végétation aux précipitations qui est décalée et intégrée sur la saison de croissance. Certaines études récentes ont démontré que les rendements pouvaient être prédits plusieurs mois à l'avance en utilisant des informations climatiques à grande échelle comme alternatives aux mesures in situ des conditions météorologiques locales (Letson et al., 2002 ; Martinez et al., 2009). Néanmoins, peu d'études ont exploité l'information prédictive contenue dans les indices et les variables climatiques pour la prévision précoce des rendements agricoles en dehors des zones sensibles à l'ENSO (Hansen, 2004). En Méditerranée du Sud, la raison principale est que les effets de ce phénomène sont probablement noyés dans la forte variabilité de la dynamique climatique des latitudes moyennes (Bolle, 2002). L'impact de l'oscillation nord-atlantique sur le climat marocain et, en particulier, sur les précipitations hivernales, est pourtant bien connu (Knippertz et al., 2003 ; Lamb et Peppler, 1987). En outre, il a également été montré que le réchauffement de l'Atlantique tropical au début de l'automne est favorable à la vigueur de la végétation au printemps le long de la côte atlantique du pays (Jarlan et al., 2013). Outre la précocité des relations rendements-climat, ces indices climatiques sont bien mieux prédits que les conditions météorologiques locales par les modèles de prévision météorologique. Il s'agit d'un aspect particulièrement intéressant pour les pays comme le Maroc où la densité de stations météorologiques est faible.

L'objectif de cette étude est double :

1. Caractériser les liens entre ces signaux climatiques et les rendements agricoles sur une province très productive dominée par la culture du blé pluvial (Settat) ;
2. Proposer des modèles statistiques pour les prédictions précoces de rendements en blé à l'échelle de cette province, en comparant la méthode de régression multi-linéaire classique avec des machines à vecteurs de support pour la régression.

2. Données et Méthodes

Les données utilisées dans cette étude sont constituées des rendements en grain que l'on va chercher à

prédire et un ensemble de variables (indice de végétation et divers indices et champs spatio-temporel témoignant de la dynamique atmosphérique et océanique à l'échelle régionale) dont pourront être déduits des prédicteurs pour le rendement.

2.1. Données utilisées

Rendements agricoles

les statistiques des rendements proviennent du ministère de l'agriculture et de la pêche maritime. Les données consistent en des rendements totaux de blé tendre associés aux surfaces cultivées à l'échelle de la province agricole pour les saisons 1982-1983 à 2007-2008. La saison de croissance commence avec le semis en novembre et la récolte a généralement lieu en mai. La province de Settat est située sur la côte Atlantique. Elle reçoit une pluviométrie moyenne sur la saison de croissance de 282 mm sur la période d'étude et les rendements moyens sont de 11 quintaux par hectare (qt / ha).

Indice de végétation normalisé

L'utilisation du Normalized Difference Vegetation Index (NDVI) pour caractériser la dynamique de la surface végétalisée s'appuie sur la réponse spectrale contrastée des tissus végétaux entre la forte absorption de la chlorophylle dans la bande rouge et la forte réflexion des feuilles dans l'infra-rouge. Les données utilisées dans cette étude sont recueillies par la série de capteurs AVHRR à bord des satellites NOAA (NOAA-7, -9, -11, -14, -16 et -17). Elles sont disponibles sur la période 1982-2008, à 8 km de résolution. La saisonnalité et la tendance sont retirées des données brutes.

Données météorologiques et climatiques

Les données de la station météorologique de la ville de Settat sont utilisées. Les données journalières sont agrégées au pas de temps mensuel pour les précipitations par un simple cumul et une moyenne est effectuée par période de 10 jours pour la température. Par ailleurs, plusieurs indices témoignant de la dynamique atmosphérique et océanique à grande échelle susceptibles d'affecter le climat marocain sont identifiés :

- Les Northern Hemisphere Teleconnection Pattern parmi lesquels l'oscillation nord-atlantique (ONA) et le motif "scandinave" (Scandinavian pattern ou SCA). Certains indices océaniques sont également utilisés : "TNA" (Tropical Atlantique Nord, 15°W-57,5°W et 5,5°N-23,5°N) et "TSA" (Tropical Atlantique Sud, 10°E-30°W et 0°-20°S) et NINO3, marqueur du phénomène ENSO (150°W-90°W et 3°S-3°N).
- Les indices climatiques n'étant qu'un "résumé" d'un phénomène climatique, nous avons également utilisé des champs spatio-temporels témoignant de la dynamique atmosphérique (géopotential à 500 hPa de l'NCEP -Kalnay et al., 1996-) et océanique (température de surface des océans -TSO- de (Reynolds et al., 2002).

L'ensemble des variables décrites ci-dessus sont au pas de temps mensuel à part le NDVI et la température qui sont au pas de temps décadaire (10 jours).

2.2. Méthodes

En vue d'évaluer les liens entre rendements d'une part, et les variables et indices climatiques d'autre part, de simples corrélations retardées (les rendements sont corrélés avec les variables prises à différentes périodes de l'année) sont utilisées entre les rendements en grain et les variables locales et climatiques. La variable cible pour les modèles de prévision est la série de rendements en blé tendre dont nous avons retiré la tendance. Les prédicteurs potentiels correspondent aux variables locales (température, précipitations et indices de végétation), atmosphérique (ONA et SCA) et océanique (TNA et TSA) à grande échelle prise à une certaine période de l'année. Concernant les champs spatio-temporels, des prédicteurs supplémentaires peuvent être construits à partir des cartes de corrélation par une simple moyenne sur une ou plusieurs fenêtres spatiales correspondantes à des motifs spatiaux cohérents de corrélation avec les rendements. Les prédicteurs, pour être sélectionnés (1) doivent présenter une corrélation significative à 95% avec la variable cible (test de Student) et (2) le lien doit avoir un sens physique. En d'autres termes, il se peut qu'étant donné le faible nombre de point des séries temporelles (26 points correspondant aux 26 saisons agricoles étudiées), des valeurs de coefficient de corrélation par chance significatives à 95% puissent être observées mais qu'elles n'aient aucun sens physique. Par exemple, une corrélation significative entre les rendements de blé et l'ONA durant l'été précédent n'aurait aucun sens étant donné que l'ONA n'impacte les précipitations que durant les mois d'hiver au Maroc.

Une fois les N prédicteurs potentiels sélectionnés, les 2N modèles correspondant à toutes les combinaisons de prédicteurs possibles sont construits. Les modèles sont évalués à l'aide du coefficient de détermination r^2 calculés en utilisant l'approche "Leave-one-out" (LOO) qui consiste à retirer aléatoirement n années du jeu de données de calibration, construire un ensemble de modèles (ici 1000) sur le jeu de données restant et effectuer la prévision sur ces n années. Si n=1, il s'agit de l'approche "leave-one-out" très populaire en statistique qui consiste à retirer une seule année. Afin de vérifier la robustesse des modèles, une étude de sensibilité à n est menée dans la dernière partie de ce manuscrit. Ce coefficient de détermination est ensuite corrigé pour tenir compte du nombre de prédicteurs suivants Theil (1961) afin de limiter la complexité des modèles et ainsi améliorer leur capacité de généralisation. Le "meilleur modèle" parmi les 2N possibles maximise le coefficient de détermination " ar^2 " pour r^2 ajusté.

Outre la régression linéaire multiple (RLM) classique détaillée par ailleurs (Draper et Smith, 1998, entre autres), nous utilisons une approche basée sur les machines à vecteur de support pour la régression (MVSF). Il s'agit

d'une technique de régression non linéaire basée sur la minimisation du risque structurel (voir Vapnik, 1989). A la différence d'autres approches telles que la RLM uniquement basée sur la minimisation de la distance aux données d'apprentissage (le risque empirique), les MVSR réduisent la complexité du modèle (i.e. recherche une fonction aussi douce que possible : le risque structurel). Les lecteurs sont invités à se référer à Kecman (2001) pour plus de détails sur la minimisation du risque structurel et les MVSR. Contrairement à la RLM, la mise en place des MVSR requiert la spécification de 3 paramètres algorithmiques qui, dans cette étude, sont identifiés par la maximisation du coefficient r^2 décrit ci-dessus suivant une méthode d'optimisation de type stratégie évolutive (CMA-ES, Hansen et al., 2003).

3. Résultats et discussion

Nous analysons, dans un premier temps, les relations entre les variables classiquement utilisées jusqu'ici (température, précipitation et indice de végétation) pour la prévision saisonnière des rendements en grains. Dans un second temps, les relations entre indices et variables climatiques et les rendements sont caractérisés. Un certain nombre de prédicteurs sont sélectionnés parmi les variables présentant une forte corrélation avec les rendements à une certaine période de l'année. Enfin, le meilleur modèle en termes de coefficient de corrélation ajusté est choisi parmi tous les modèles possibles avec cet ensemble de prédicteurs pré-sélectionnés. La robustesse de ce modèle est évalué et la qualité de ses prédictions analysée.

3.1. Relation entre les rendements et les prédicteurs potentiels

La figure 1 présente les corrélations entre les rendements et, d'une part, les conditions locales (Figure 1a) et, d'autre part, les indices climatiques (Figure 1b).

Conditions locales

Le développement phénologique du blé tendre est lié aux conditions météorologiques locales parmi lesquelles la disponibilité en eau et la température sont les plus importants. La figure 1a montre que des températures douces durant le début de l'hiver sont favorables aux rendements (corrélation positive fin décembre). Des températures anormalement basses durant cette période impacte l'émergence et le tallage en réduisant le nombre de pieds et, *in fine*, le rendement. Inversement, des températures trop élevées au printemps (corrélation négative en mars) pénalisent les rendements car cette période correspond à la phase de remplissage des grains. De même, la corrélation significative du mois de Novembre entre les rendements et les précipitations est à relier avec la phase de germination favorisée par des pluies précoces et régulières. Des pluies précoces favorisent

également un semis précoce bien connu pour être favorable aux rendements. Enfin les corrélations significatives avec les précipitations plus tard dans la saison s'expliquent par la nécessité d'éviter le stress hydrique durant la phase de croissance du blé qui suit le tallage pour obtenir des rendements importants. Enfin, les anomalies de NDVI se trouvent être significativement corrélées aux rendements plus tard dans la saison pendant toute la phase de développement à partir de janvier jusqu'à la sénescence, début Avril. Cette corrélation positive doit être liée au fait qu'une grande surface foliaire et le nombre de timons associés limitent la perte d'eau par évaporation du sol et ainsi améliorent l'efficacité d'utilisation de l'eau par les plantes (Cooper et al., 1987). Les liens mis en évidence entre les rendements d'une part et les conditions météorologiques du début de saison (température et précipitations) sont particulièrement intéressants dans une perspective de prévision précoce.

Climat régional et global

La figure 1b montre le coefficient de corrélation décalée entre les rendements et les indices climatiques. ONA et SCA reflètent la circulation atmosphérique qui peut avoir un impact concomitant sur la température et les précipitations locales, alors que les indices TSA/TNA et NINO3 représentent, respectivement, les conditions de température de surface de l'océan Atlantique et Pacifique qui peuvent être reliés au climat local par téléconnection. Aucun lien n'a été mis en évidence ici avec l'indice NINO3 liée à ENSO mais des travaux antérieurs tentant de mettre en évidence un lien entre les précipitations sur la zone euro-méditerranéenne et l'ENSO ont également été peu concluants. Par contre, la valeur de l'ONA en décembre est corrélée négativement au rendement du blé, probablement à travers l'influence de la température et du stress hydrique sur les premières phases du cycle du blé (germination, tallage) déjà discutée ci-dessus. En effet, la phase positive de l'ONA (l'intensification de la dépression d'Islande et le renforcement de l'anticyclone des Açores) est associée à un déficit de précipitations au-dessus de l'Europe du Sud et du bassin méditerranéen à travers une orientation des flux d'Ouest qui transportent les systèmes synoptiques loin du Nord de l'Afrique. Au cours de la phase négative, les effets sont inversés et cela entraîne des précipitations excédentaires sur l'Afrique du Nord et l'Europe du Sud par le biais de flux d'ouest affaiblis et une direction des flux plus zonale. Sur le Maroc, l'influence de l'ONA sur la côte ouest du pays où la province de Settat étudiée ici est située est forte (Knippertz et al., 2003). Le lien entre les rendements et le SCA en début de saison (figure 1b) est également lié à la nécessité d'une bonne disponibilité en eau pendant la phase de tallage.

Le climat hivernal du bassin méditerranéen est dominé par l'alternance des régimes de temps extra-tropicaux et la subsidence de l'air tropical sec à travers la branche

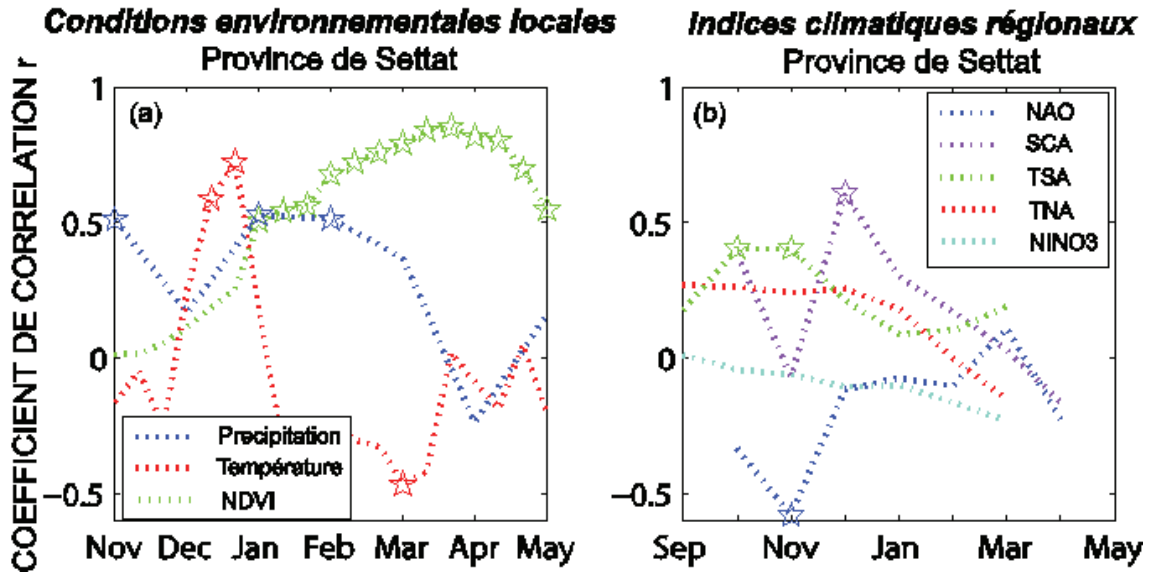


FIGURE 1 : Corrélation retardée entre les rendements en blé tendres et (a) les conditions météorologiques locales et de développement des plantes (température, précipitation et NDVI) et (b) les indices climatiques témoignant de la dynamique atmosphérique (ONA et SCA) et océanique (dans l'Atlantique tropical TNA et TSA et dans le pacifique équatorial NINO3). Les étoiles correspondent aux valeurs significatives à 95% (test de Student).

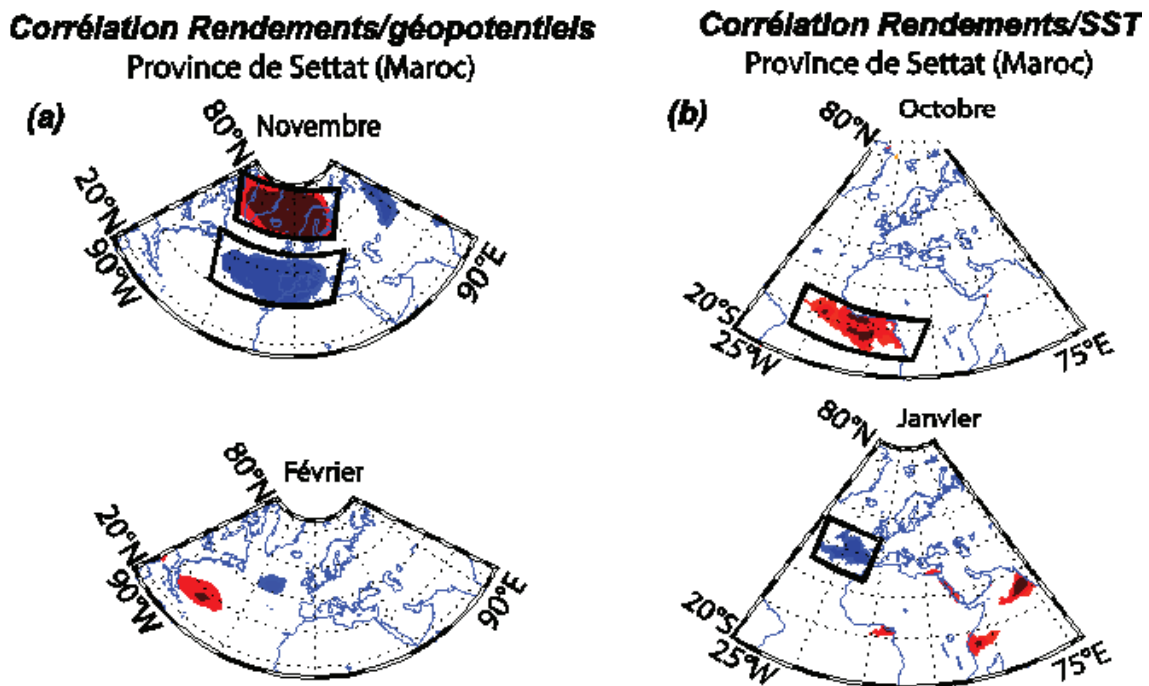


FIGURE 2 : Corrélation retardée entre les rendements en blé tendre et (a) les géopotentiels à 500 hPa pour les mois de Novembre et Février et (b) les températures de surface des océans pour les mois d'octobre et janvier. Uniquement les corrélations significatives à 95% sont tracées. Les fenêtres noires sont les régions sur lesquelles sont moyennées les variables pour construire des prédicteurs spécifiques (voir texte). L'échelle de couleur va du bleu (corrélations négatives significatives) au rouge (corrélations positives significatives).

descendante de la cellule de Hadley. Knippertz et al. (2003) étudient trois cas de pluies précoces et démontrent que les forêts tropicales d'Afrique de l'Ouest et l'Atlantique tropical est une source potentielle d'humidité pour le Maroc. Cassou et Terray (2001) ont montré que les liens observés entre l'Atlantique tropical et les latitudes moyennes pourraient s'expliquer par la perturbation de la cellule de Hadley et la perturbation associée de l'intensité et de l'emplacement du jet aux latitudes moyennes. En accord avec ces études antérieures, il est constaté qu'un réchauffement de l'Atlantique tropical et équatorial au cours du début de l'automne peut être favorable pour les rendements du blé sur le Maroc (corrélation légèrement significative avec l'indice Tropical South Atlantic -TSA- durant les mois d'Octobre et Novembre), toujours probablement grâce aux précipitations accrues sur le Maroc au cours de la première phase de développement du blé.

Les rendements sont ensuite corrélés aux champs de géopotential à 500 hPa et aux températures de surface des océans. La figure 2a montre la corrélation retardée (significative à 95%) entre les rendements et les géopotentiels mensuels. Deux régions homogènes de corrélation sont mises en évidence : un maximum de corrélation positive centré au-dessus de l'Islande et une région de corrélation négative qui s'étend du sud de l'Europe vers l'ouest, au-dessus de l'océan Atlantique. Ce dipôle de corrélation est très proche du dipôle caractéristique de l'oscillation nord-atlantique au niveau 500 hPa et confirme ainsi la corrélation trouvée avec l'indice ONA. Néanmoins, ces motifs spatiaux sont utilisés ici pour construire des prédictors supplémentaires qui permettent de prendre en compte une partie de la variabilité du climat régional qui n'est pas pris en compte par le seul indice ONA. Ces prédictors sont simplement la moyenne des géopotentiels de novembre prise sur les fenêtres spatiales tracées en noir sur la figure 2a.

La même approche est ensuite appliquée pour étudier le lien entre les rendements et la température de surface de l'océan (Figure 2b). Globalement, deux motifs spatiaux de corrélation peuvent être distingués : une zone de corrélation positive située dans la zone équatoriale et tropicale de l'Atlantique sud le long de la côte ouest-africaine en Octobre (qui confirme le lien mis en évidence avec l'indice TSA) et une zone de corrélation négative le long de la péninsule ibérique tard dans la saison en Janvier. Le motif de corrélation positive est probablement lié à l'influence de la TSO de l'Atlantique tropical en fin d'été / début d'automne sur la variabilité atmosphérique de l'hémisphère nord et les phases ONA au cours de l'automne / début de l'hiver suivant par télé-connection comme déjà mentionné ci-dessus. La zone de corrélation négative est probablement liée au premier mode de variabilité interannuelle de la température de surface sur l'Atlantique Nord, montrant une tripôle avec

des anomalies chaudes au large de la côte est des États-Unis et des anomalies froides au nord de 40°N et au sud de 25°N déjà largement documentés (Cayan, 1992). En effet en relâchant la corrélation à 90%, les anomalies chaudes de la côte est des États-Unis sont peu visibles, mais les anomalies négatives au sud de 20°N apparaissent clairement (non montré). De même, deux nouveaux prédictors sont construits à partir de la moyenne des TSO sur les fenêtres de corrélation les plus élevées (en octobre pour l'Atlantique tropical et en janvier pour la bande le long de la péninsule ibérique).

3.2. Prévision précoce des rendements en blé tendre

Influence de la période de prévision

Les modèles de prévisions précoces sont construits tous les mois de Décembre à Mars avec les prédictors disponibles à chaque période ; les paramètres algorithmiques de l'approche par MVSR étant optimisés à chaque fois afin de maximiser le coefficient de corrélation ajusté. La figure 3a montre l'évolution des performances en fonction du mois de prévision en termes de coefficient de corrélation (approche leave-one-out) et d'erreur quadratique. Les performances s'améliorent évidemment en se rapprochant de fin de la saison avec l'augmentation du nombre de prédictors testés mais de bonnes prévisions peuvent être obtenues assez tôt dans l'année ($r^2=0,72$ et $RMSE=3,6$ qt/ha pour les modèles MVSR du mois de janvier par exemple reposant essentiellement sur des prédictors climatiques : l'ONA de Novembre, le SCA de décembre et les TSO de l'océan atlantique équatorial d'octobre). Les meilleurs modèles sont néanmoins obtenus pour une prévision depuis le mois de mars, en particulier grâce à l'apport du NDVI de ce mois, très corrélés aux rendements (figure 1a). Le modèle basé sur les MVSR passe d'ailleurs le critère statistique pour les modèles de ce type en Europe ($RMSE \leq 2$ qt/ha ; Genova et al., 2006). Par ailleurs, les performances des modèles basés sur l'approche MVSR dépassent systématiquement ceux basés sur les RLM.

Robustesse

Afin d'évaluer la robustesse des modèles développés, nous avons effectué une étude de sensibilité aux nombres d'années retirées du jeu de calibration. La figure 3b montre la dégradation des performances en fonction du nombre d'années n retirées. Le cas où $n=1$ correspond aux modèles discutés ci-dessus. La tendance générale est à une dégradation régulière des performances avec n jusqu'à une rupture assez nette pour $n=7$ ce qui correspond à environ 25% du jeu de données (26 années au total). Néanmoins, les résultats jusqu'à 6 années retirées demeurent très satisfaisants. La figure 3b montre également une dégradation des performances des modèles basés sur les MVSR plus fortes que pour les RLM. Nous touchons ici à une limite des MVSR qui nécessitent une calibration fine des paramètres algorithmiques. Dans ce cas, les paramètres ont

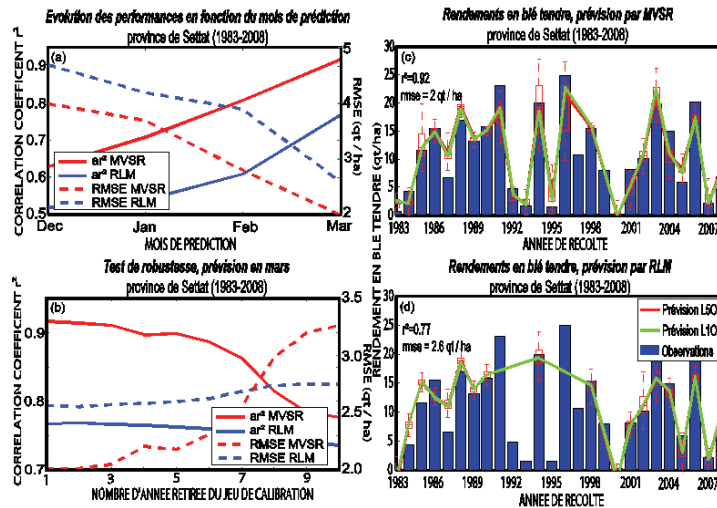


FIGURE 3 : Performance de la prévision précoce pour les deux approches utilisées dans cette étude (régression linéaire multiple -RLM- et machine à vecteur de support pour la régression -MVSR-) : (a) évolution des performances des modèles de prévision en fonction du mois de l'année auquel est effectuée la prévision ; (b) Evolution des performances des modèles en fonction du nombre d'année retiré du jeu de calibration (pour une prévision à partir du mois de mars) ; (c) et (d) séries temporelles de rendements observés et prédits respectivement par la méthode RLM et par la méthode MVSR.

été calibrés pour une approche de type L-1-O et conservés pour l'analyse de sensibilité. Les RLM bénéficie ainsi d'une plus grande facilité de mise en oeuvre. Néanmoins, dans une perspective opérationnelle, le cas où une seule année est retirée correspond typiquement au cas du gestionnaire qui cherchera à prédire la production de l'année à venir en disposant de tout l'historique.

Prévision des rendements

Les figures 3c et 3d présentent les séries temporelles des rendements observés et prédits pour une approche de type L1O, respectivement par la méthode des MVSR et RLM. Les prédictions obtenues et leur dispersion en retirant 5 années du jeu de données (approche L5O) sont également superposées pour comparaison. L'accord global entre observations et prédictions est très bon. Par ailleurs, les modèles sont capables de prédire la production pour des années atypiques telles que 2000 ou 2003. Outre la meilleure performance des MVSR, la figure 3 met un peu par hasard en évidence un autre avantage des MVSR par rapport au RLM qui est la parcimonie de la représentation. Le modèle MVSR est basé sur 4 prédicteurs (l'ONA de Novembre et les TSO de l'océan atlantique équatorial d'octobre, le SCA de décembre et le NDVI de la 3ème décennie de mars) alors que le modèle RLM en utilise 6, parmi lesquels la température des mois de décembre et de mars. Il se trouve que ces données ne sont pas disponibles pour plusieurs années (1991, 1992, 1993, 1996) sur la station de Settat empêchant toute prédiction sur ces années là.

4. Conclusion

Cette étude analyse les liens entre les rendements agricoles sur la province de Settat (Maroc) d'une part et plusieurs variables liées aux conditions météorologiques et environnementales locales (température, précipitation, NDVI) et à la dynamique atmosphérique et océanique à grande échelle vue à travers certains indices (ONA, SCA, TSA) et variables (géopotentiels, TSO) climatiques d'autre part. Les relations bien connues entre les rendements et la température et les précipitations à certains stades critiques du développement du blé déjà mis en évidence par plusieurs auteurs à l'échelle de la parcelle sont retrouvées ici à l'échelle de la province agricole. Par ailleurs, un lien assez fort, probablement à travers les précipitations de début de saison est mis en évidence avec la ONA de Novembre et les TSO d'octobre dans l'océan Atlantique équatorial par téléconnection. Les corrélations significatives observées sont exploitées dans une deuxième partie pour évaluer les potentialités de prévision saisonnière du blé dès le mois de décembre. Deux approches sont comparées : la régression linéaire multiple (RLM) et une approche de régression non-linéaire basée sur les machines à vecteur de support (MVSR). Nous montrons que (1) la régression non linéaire de type MVSR donne de bien meilleurs résultats que la RLM classiquement utilisée pour la construction de ce type de modèles empiriques ; (2) les meilleurs modèles obtenus sont pour une prévision à partir de fin mars ($r^2=0,92$ et $rmse=2$ qt/ha pour le modèle MVSR) mais des prévisions de bonne qualité peuvent être envisagées plus tôt dans l'année. Notre conclusion est que l'utilisation d'indices et de variables témoignant de la dynamique atmosphérique et océanique à l'échelle régionale peut

permettre une prévision plus précoce qu'avec les prédictors classiquement utilisés (température, précipitations et indices de végétation) pour le développement de ce type de modèles. Cette approche est en cours de test sur les provinces agricoles les plus productives en céréales du Royaume.

Remerciements

Ce travail a été conduit dans le cadre du Laboratoire Mixte International TREMA (<http://trema.ucam.ac.ma>) et financé par le CESBIO, le programme MISTRALS et le CNRST.

Références

- Baigorria, G., Jones, J.W. and O'Brien, J.J., 2008. Potential predictability of crop yield using an ensemble climate forecast by a regional circulation model. *Agricultural and Forest Meteorology*, 148(8-9), 1353-1361.
- Balaghi, R., Tychon, B., Eerens, H. and 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(4), 438-452.
- Bolle, H.J., 2002. *Mediterranean Climate : Variability and Trends* (Springer-V., p. 320).
- Brisson, N., Launay, M., Mary, B. and Beaudoin, N., 2009. Conceptual basis, formalisations and parameterization of the SticS crop model. (Collection : Update Sciences and Technologies, Ed.) (Quae., p. 304).
- Cassou, C. and Terray, L., 2001. Dual Influence of Atlantic and Pacific SST on the North Atlantic/Europe Winter Climate. *Geophysical Research Letters* 28(16), 3195-3198.
- Cayan, D., 1992. Latent and Sensible Heat Flux anomalies over the Northern oceans : Driving the Sea Surface Temperature. *Journal of Physical Oceanography* 22, 859-881.
- Cooper, P.J.M., Gregory, P.J., Tuly, D. and Harric, H.C., 1987. Improving Water Use Efficiency of Annual Crops in the Rainfed Farming Systems of West Asia and North Africa. *Exper. Agriculture*, 23, 113-158.
- Draper, N.R. and Smith, H., 1998. *Applied Regression Analysis*, 3rd Edition (p. 736). Wiley Science.
- Genovese, G., Bettio, M. and Fritz, S., 2006. Report on Yields Intercomparison Study (p. 156). Ispra, Italy.
- Ghosh, S. and Mujumdar, P.P., 2009. Climate change impact assessment : Uncertainty modeling with imprecise probability. *Journal of Geophysical Research*, 114(D18).
- Hansen, J., 2004. Linking dynamic seasonal climate forecasts with crop simulation for maize yield prediction in semi-arid Kenya. *Agricultural and Forest Meteorology*, 125(1-2), 143-157.
- Hansen, J.W., Jones, J.W., Kiker, C.F. and Hodges, A.W., 1999. El Niño- Southern Oscillation Impacts on Winter Vegetable Production in Florida. *Journal of Climate*, 12, 92-102.
- Hansen, N., Müller, S.D. and Koumoutsakos, P., 2003. Reducing the Time Complexity of the Derandomized Evolution Strategy With Covariance Matrix Adaptation (CMA-ES). *Evolutionary Computation*, 11(1), 1-18.
- Harris, H.C., 1991. Soil and Crop Management for Improved Water Use Efficiency in Rain-Fed Areas. In H.C. Harris, P.J. Cooper and M. Pala (Eds.), *Implications of climatic variability* (Implicatio., p. 352). Aleppo, Syria : ICARDA.
- Jarlan, L., Driouech, F., Duchemin, B., Tourre, Y., Abaoui, J., Kharrou, H., et al., 2013. Seasonal and interannual vegetation dynamics over Morocco through the NDVI / AVHRR from 1982 to 2008 : linkages with climate signals and potential for seasonal prediction. *International Journal of Climatology*.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., et al., 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77(3), 437-471.
- Kalra, A. and Ahmad, S., 2009. Using oceanic-atmospheric oscillations for long lead time streamflow forecasting. *Water Resources Research*, 45(3), 1-18.
- Kecman, V., 2001. *Learning and Soft Computing : Support Vector Machines, Neural Network and Soft Computing*. Cambridge Massachussets : the MIT Press.
- Knippertz, P., Christoph, M. and Speth, P., 2003. Long-term precipitation variability in Morocco and the link to the large-scale circulation in recent and future climates. *Meteorology and Atmospheric Physics*, 83(1-2), 67-88.
- Knippertz, P., Fink, A.H., Reiner, A. and Speth, P., 2003. Three Late Summer / Early Autumn Cases of Tropical - Extratropical Interactions Causing Precipitation in North-west Africa. *Monthly Weather Review*, 131, 116-135.
- Lamb, P.J. and Pepler, R.A., 1987. North Atlantic Oscillation : Concept and an Application. *Bulletin of the American Meteorological Society*, 68(10), 1218-1225.
- Letson, D., Messina, C., Podesta, G., Royce, F., Andre, R., Hansen, J., Llovet, I., et al., 2002. Use of ENSO-related climate information in agricultural decision making in Argentina ? : a pilot experience. *Agricultural Systems*, 74, 371-392.
- Martinez, C.J., Baigorria, G.A. and Jones, J.W., 2009. Use of climate indices to predict corn yields in southeast. *International Journal of Climatology*, 1691(December 2008), 1680-1691.
- Phillips, J., Cane, M. and Rosenzweig, C., 1998. ENSO, seasonal rainfall patterns and simulated maize yield variability in Zimbabwe. *Agricultural and Forest Meteorology*, 90(1-2), 39-50.
- Reynolds, R.W., Rayner, N.A., Smith, T.M., Stockes, D.C. and Wang, W., 2002. An improved in situ and satellite SST analysis for climate. *Journal of Climate*, 15, 1609-1625.
- Theil, H., 1961. *Economic Forecasts and Policy*. Amsterdam, Holland.
- Vapnik, V.N., 1989. *Statistical Learning Theory*. Wiley-Interscience.