ESTIMATING MAIZE GRAIN YIELD IN SCARCE FIELD-DATA ENVIRONMENT : AN APPROACH COMBINING REMOTE SENSING AND CROP MODELLING IN BURKINA FASO

Leroux L., Baron C., Castets M., Escorihuela M-J., Diouf A.A., Bégué A., Lo Seen D.

AfricaGIS2017, November,2017, Ethiopia

MAIZE (*ZEA MAYS L.*) : KEYSTONE OF FOOD SECURITY

- Most produced crop in the world

- **In West Africa :**
  - Staple crop
  - Providers of health benefits and vital nutrients
  - 30 kg/capita/year

MAIZE (*ZEA MAYS L.*) : KEYSTONE OF FOOD SECURITY

- Most produced crop in the world

- **In West Africa** :
  - Staple crop
  - Providers of health benefits and vital nutrients
  - 30 kg/capita/year

KEYSTONE TO SUPPORT FOOD REQUIREMENTS AND FOOD SECURITY

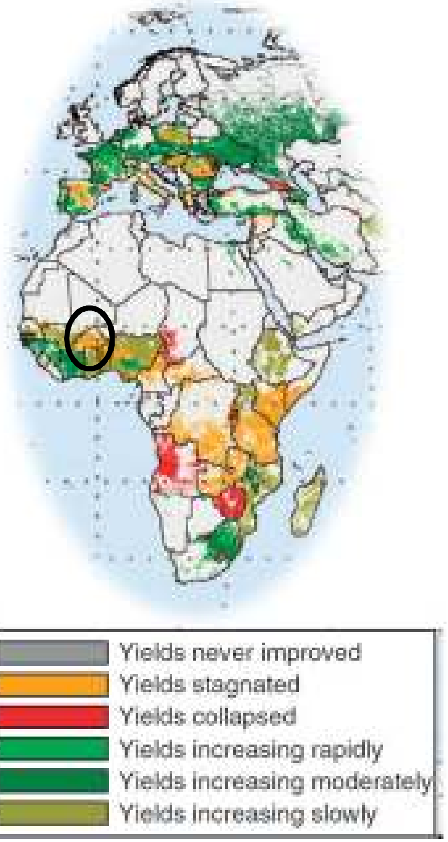
MAIZE (*ZEA MAYS* L.) : KEYSTONE OF FOOD SECURITY

- Most produced crop in the world

- **In West Africa :**
  - Staple crop
  - Providers of health benefits and vital nutrients
  - 30 kg/capita/year

KEYSTONE TO SUPPORT FOOD REQUIREMENTS AND FOOD SECURITY

MAIZE YIELD TRENDS [RAY ET AL., 2012]



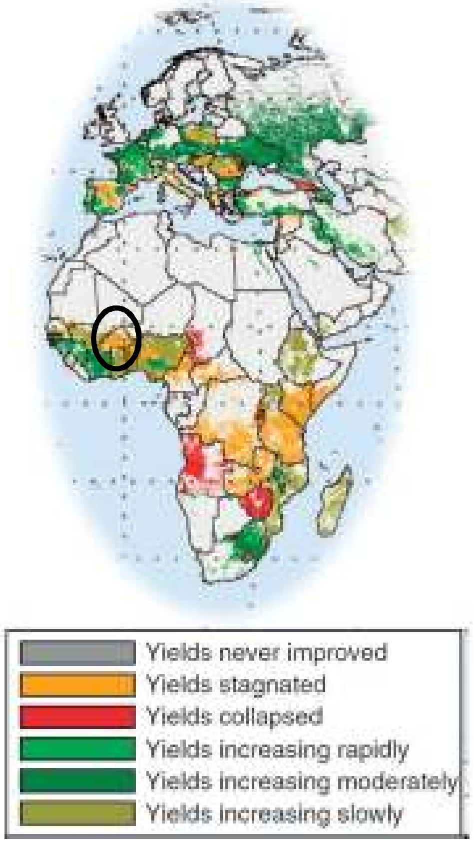
- Stagnation of maize yields
- Socio-economic x biophysical limitations

MAIZE (*ZEA MAYS L.*) : KEYSTONE OF FOOD SECURITY

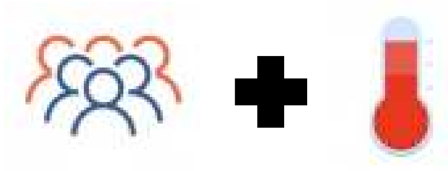
- Most produced crop in the world
- **In West Africa :**
  - Staple crop
  - Providers of health benefits and vital nutrients
  - 30 kg/capita/year

KEYSTONE TO SUPPORT FOOD REQUIREMENTS AND FOOD SECURITY

MAIZE YIELD TRENDS [RAY ET AL., 2012]



- Stagnation of maize yields
- Socio-economic x biophysical limitations



DECLINE IN PER CAPITA FOOD PRODUCTION [RAY ET AL., 2013]

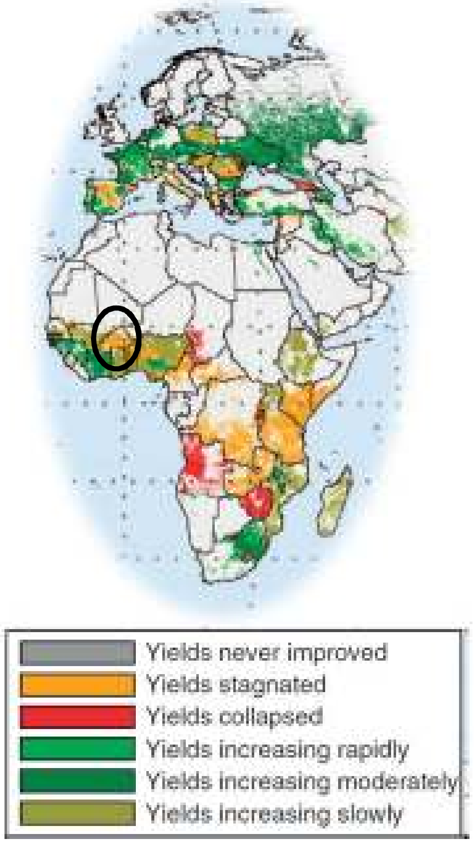
MAIZE (*ZEA MAYS L.*) : KEYSTONE OF FOOD SECURITY

- Most produced crop in the world

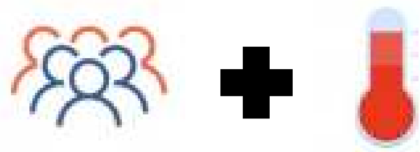
- **In West Africa :**
  - Staple crop
  - Providers of health benefits and vital nutrients
  - 30 kg/capita/year

KEYSTONE TO SUPPORT FOOD REQUIREMENTS AND FOOD SECURITY

MAIZE YIELD TRENDS [RAY ET AL., 2012]



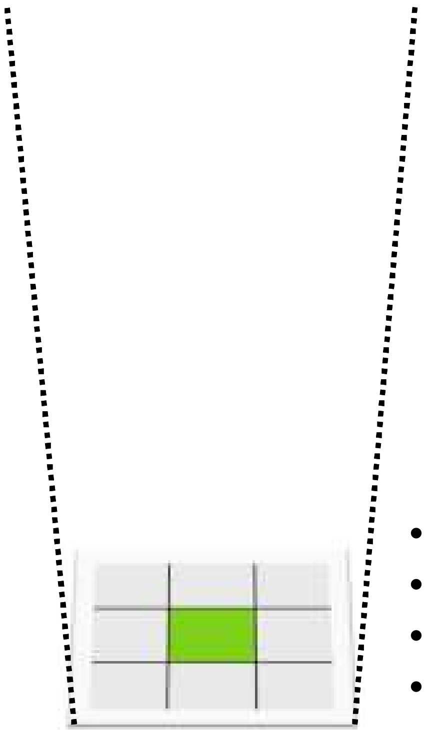
- Stagnation of maize yields
- Socio-economic x biophysical limitations



DECLINE IN PER CAPITA FOOD PRODUCTION [RAY ET AL., 2013]

Timely and reliable information on maize crop yields is needed to provide timely estimates of food shortage and support decision-making

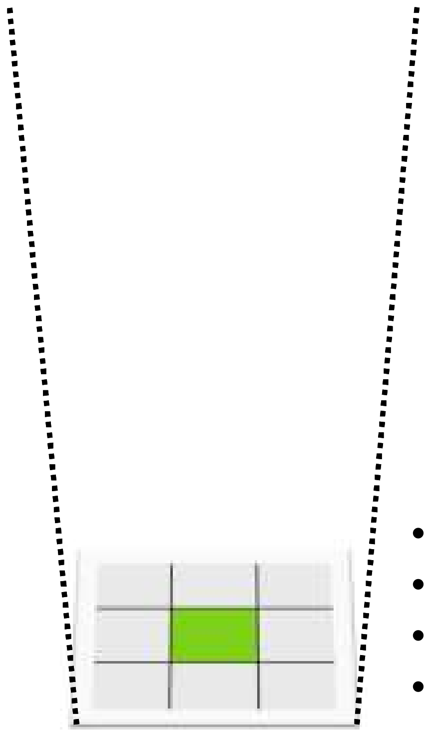
YIELD ESTIMATION METHODS



FIELD-BASED SURVEY

- Expensive (time & labor)
- Sampling methods
- Inaccessibility
- Difficulties to upscale to large areas

YIELD ESTIMATION METHODS



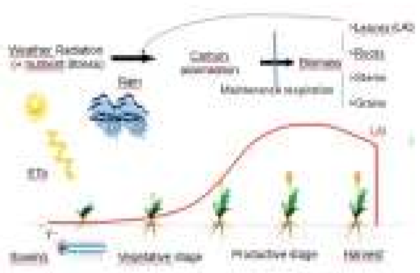
FIELD-BASED SURVEY

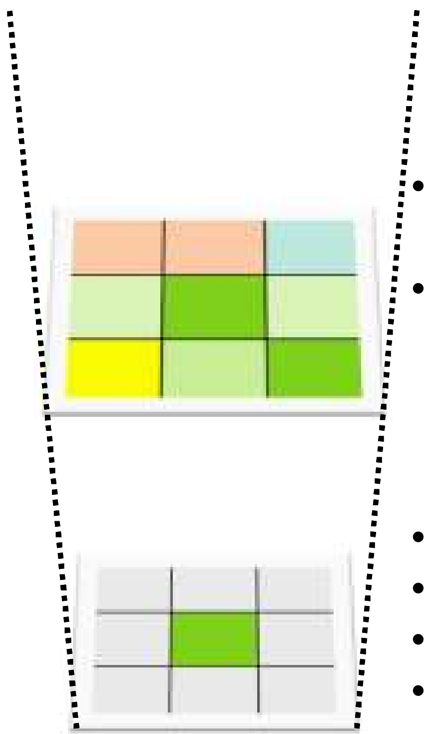
- Expensive (time & labor)
- Sampling methods
- Inaccessibility
- Difficulties to upscale to large areas

LACK OF GROUND DATA OR UNRELIABLE DATA

YIELD ESTIMATION METHODS

CROP GROWTH MODEL





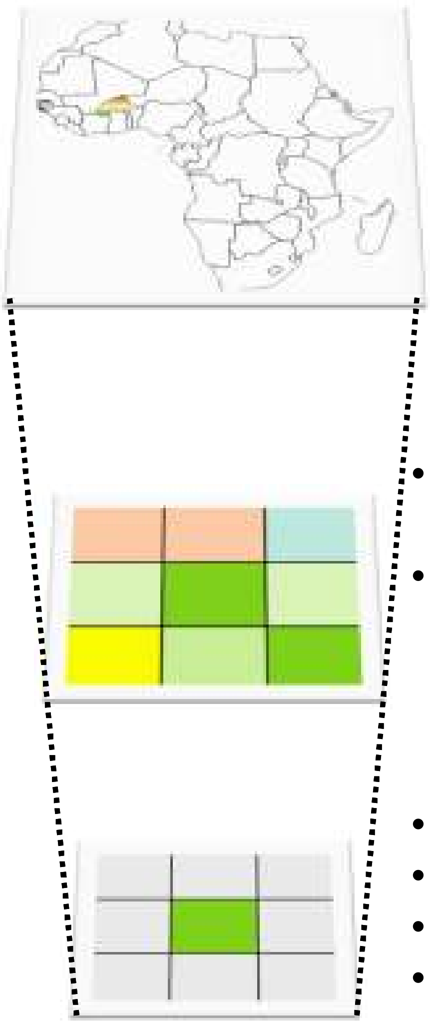
- Approximation of the reality on the ground
- Potential yields under water or nutrient limitation

FIELD-BASED SURVEY

- Expensive (time & labor)
- Sampling methods
- Inaccessibility
- Difficulties to upscale to large areas

LACK OF GROUND DATA OR UNRELIABLE DATA

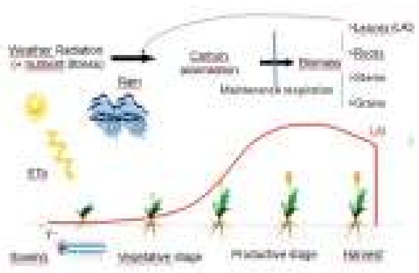
YIELD ESTIMATION METHODS



REMOTE SENSING

- Timely and exhaustive information on vegetation cover
- Biomass=f(Vegetation Indices)
- Empirical model calibrated with agricultural statistics **BUT** available ~ 3 months after the end of the cropping season

CROP GROWTH MODEL



- Approximation of the reality on the ground
- Potential yields under water or nutrient limitation

FIELD-BASED SURVEY

- Expensive (time & labor)
- Sampling methods
- Inaccessibility
- Difficulties to upscale to large areas

LACK OF GROUND DATA OR UNRELIABLE DATA

OBJECTIVES

IMPROVE MAIZE YIELDS ESTIMATION USING A CROP MODEL TO GENERATE DIFFERENT COMPONENTS OF YIELDS AS PROXY OF IN SITU OR AGRICULTURAL STATISTICS DATA, AND COMBINING THEM WITH REMOTE SENSING DATA

Uncalibrated approach [Lobell et al., 2015, Burke et al., 2017, Sibley et al., 2014]

OBJECTIVES

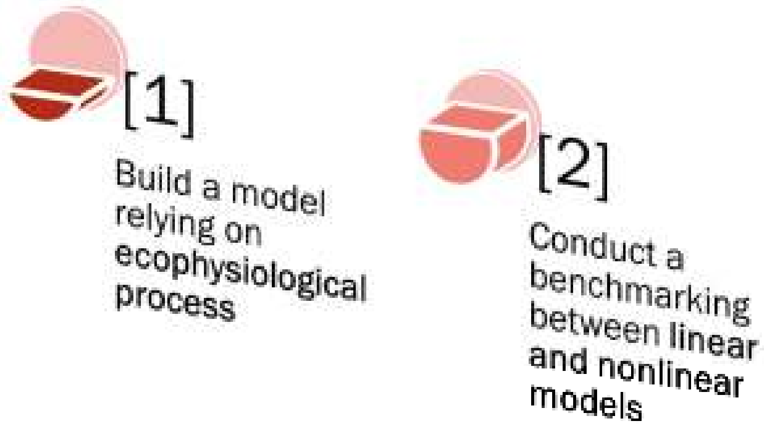
IMPROVE MAIZE YIELDS ESTIMATION USING A CROP MODEL TO GENERATE DIFFERENT COMPONENTS OF YIELDS AS PROXY OF IN SITU OR AGRICULTURAL STATISTICS DATA, AND COMBINING THEM WITH REMOTE SENSING DATA

Uncalibrated approach [Lobell et al., 2015, Burke et al., 2017, Sibley et al., 2014]

OBJECTIVES

IMPROVE MAIZE YIELDS ESTIMATION USING A CROP MODEL TO GENERATE DIFFERENT COMPONENTS OF YIELDS AS PROXY OF IN SITU OR AGRICULTURAL STATISTICS DATA, AND COMBINING THEM WITH REMOTE SENSING DATA

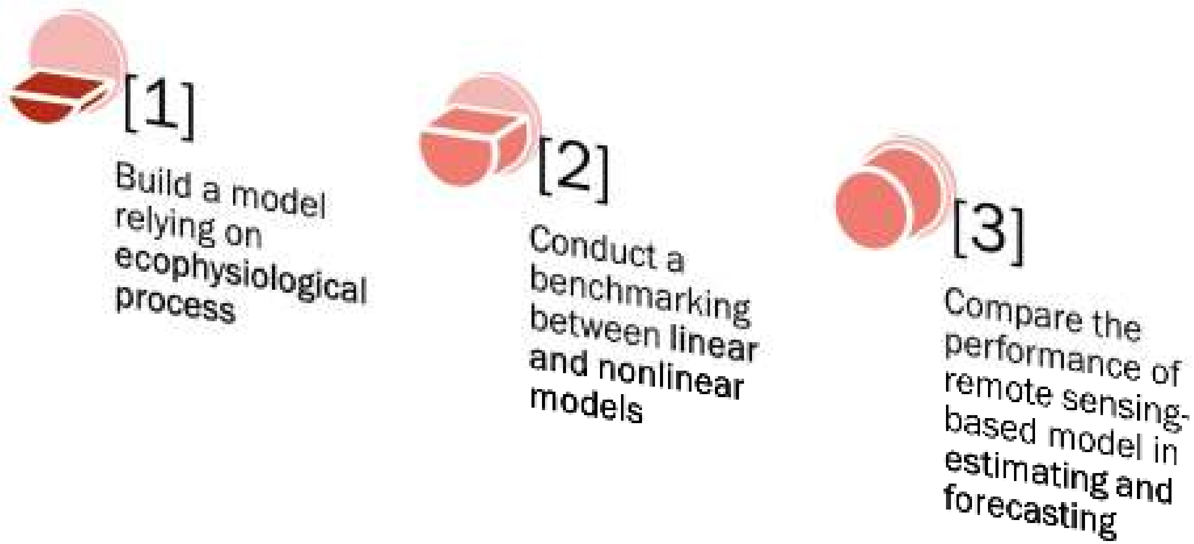
Uncalibrated approach [Lobell et al., 2015, Burke et al., 2017, Sibley et al., 2014]



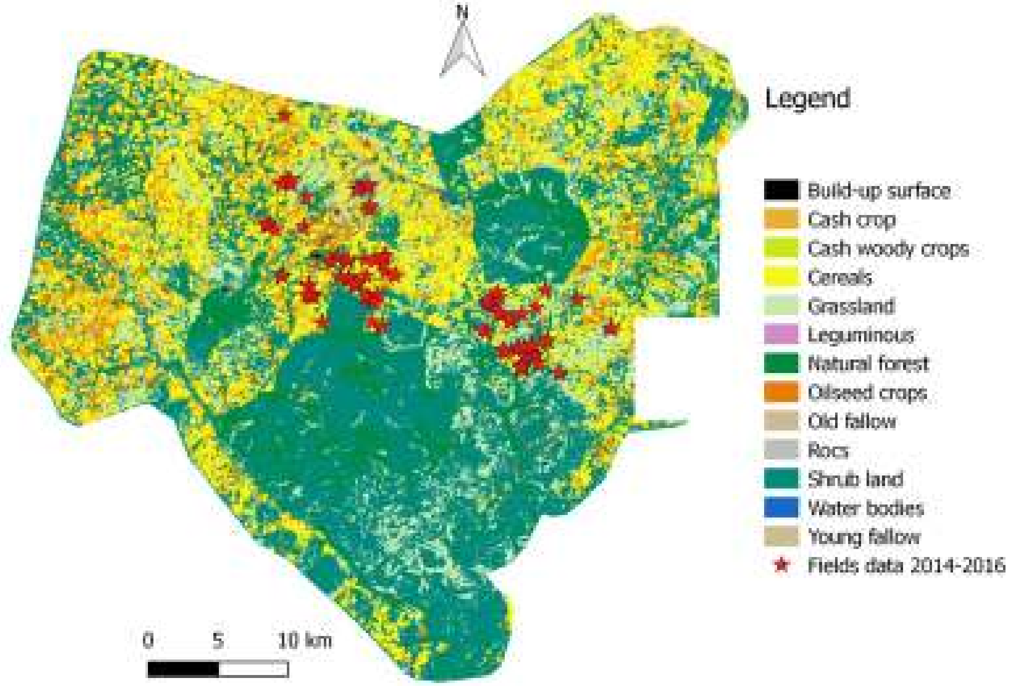
OBJECTIVES

IMPROVE MAIZE YIELDS ESTIMATION USING A CROP MODEL TO GENERATE DIFFERENT COMPONENTS OF YIELDS AS PROXY OF IN SITU OR AGRICULTURAL STATISTICS DATA, AND COMBINING THEM WITH REMOTE SENSING DATA

Uncalibrated approach [Lobell et al., 2015, Burke et al., 2017, Sibley et al., 2014]



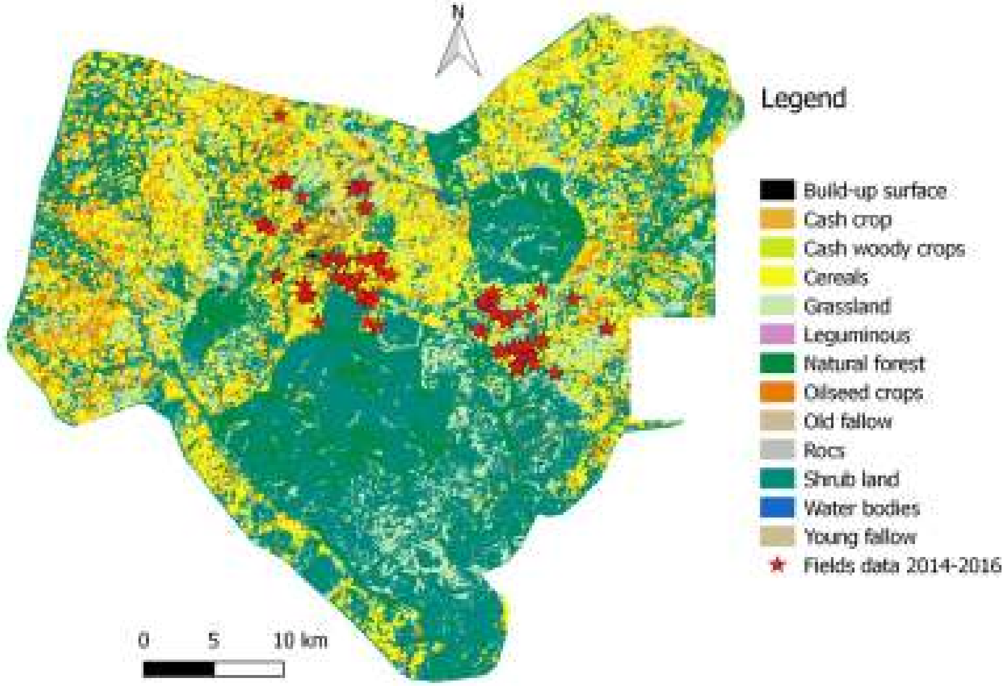
STUDY AREA & FIELDS DATA



- Tuy province of Burkina Faso
- Sudanian climate
- Rainy season : June-September

- Agropastorale activities
- Rainfed crops : mainly maize and cotton

STUDY AREA & FIELDS DATA

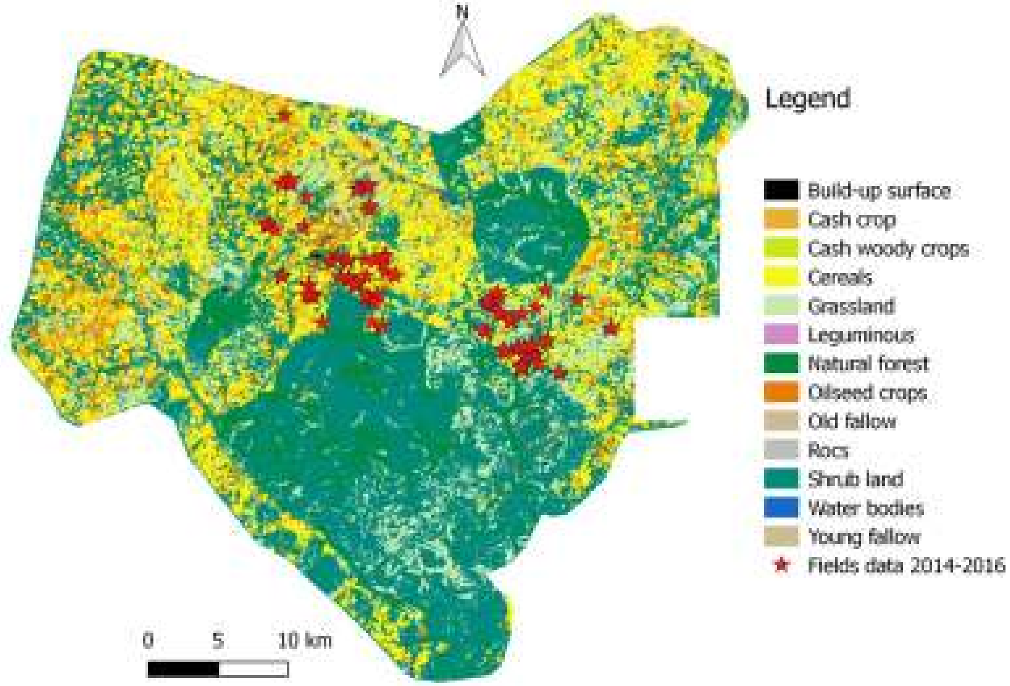


- Tuy province of Burkina Faso
- Sudanian climate
- Rainy season : June-September

- Agropastorale activities
- Rainfed crops : mainly maize and cotton

- Field surveys in 2014-2015-2016
- 114 maize fields
- Agricultural practices and vegetation parameters

STUDY AREA & FIELDS DATA



- Tuy province of Burkina Faso
- Sudanian climate
- Rainy season : June-September

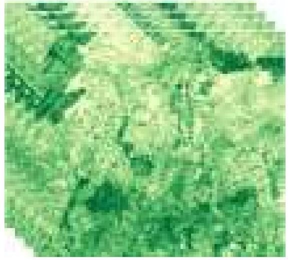
- Agropastorale activities
- Rainfed crops : mainly maize and cotton

- Field surveys in 2014-2015-2016
- **114 maize fields**
- Agricultural practices and vegetation parameters

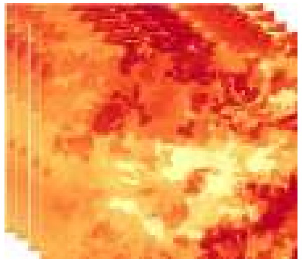
INDEPENDENT DATASET TO TEST THE ROBUSTNESS OF THE APPROACH

DATA

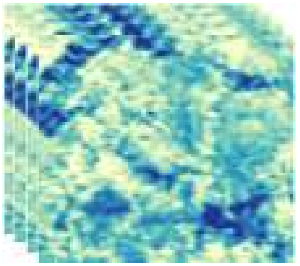
MODIS NDVI TIME SERIES



LAND SURFACE TEMPERATURE TIME SERIES



SMOS SOIL MOISTURE TIME SERIES

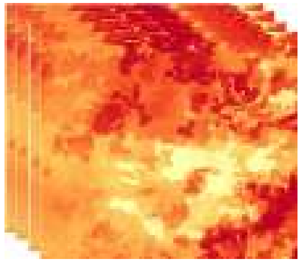


Disaggregated with the DISPATCh method [Merlin et al., 2013]

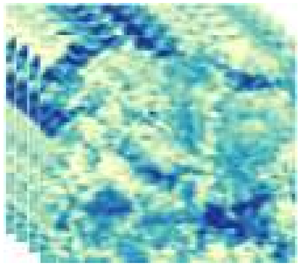
DATA

MODIS NDVI TIME SERIES

LAND SURFACE TEMPERATURE TIME SERIES



SMOS SOIL MOISTURE TIME SERIES



Disaggregated with the DISPATCh method [Merlin et al., 2013]

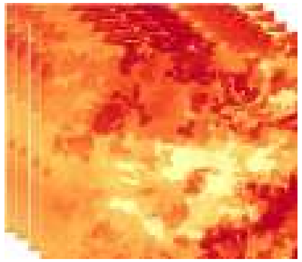
Phenology, Vegetation vigor and drougth/heat stress related indicators

CWSI | TCI | TVDI | SMADI | NDVI | SSM

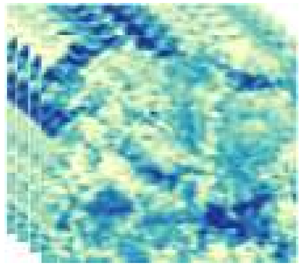
DATA

MODIS NDVI TIME SERIES

LAND SURFACE TEMPERATURE TIME SERIES



SMOS SOIL MOISTURE TIME SERIES



Disaggregated with the DISPATCh method [Merlin et al., 2013]

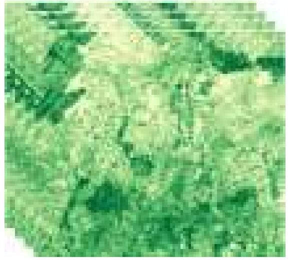
Phenology, Vegetation vigor and drougth/heat stress related indicators

CWSI | TCI | TVDI | SMADI | NDVI | SSM

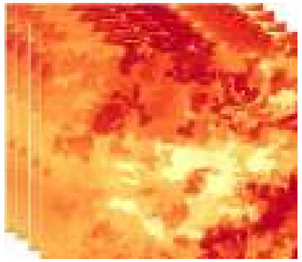
Start of Season | Peak of Season | End of Season

Data

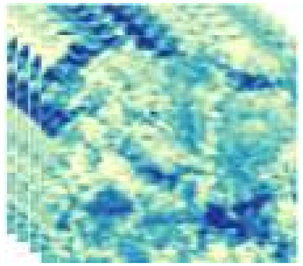
MODIS NDVI time series



Land Surface Temperature time series



SMOS Soil Moisture time series



Disaggregated with the DISPATCh method [Merlin et al., 2013]

Phenology, Vegetation vigor and drougth/heat stress related indicators

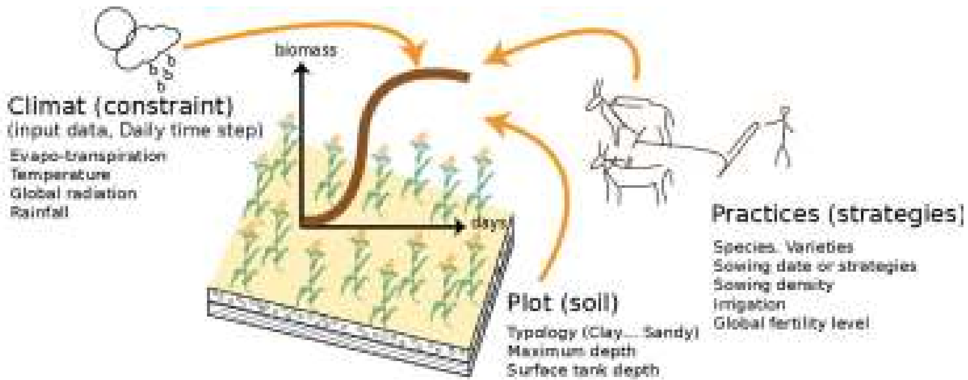
CWSI | TCI | TVDI | SMADI | NDVI | SSM

Start of Season | Peak of Season | End of Season

Vegetative period | Productive period

SARRA-O CROP MODEL [BARON ET AL., 2005; CASTETS ET AL., IN PREP]

SARRA-O CROP MODEL [BARON ET AL., 2005; CASTETS ET AL., IN PREP]

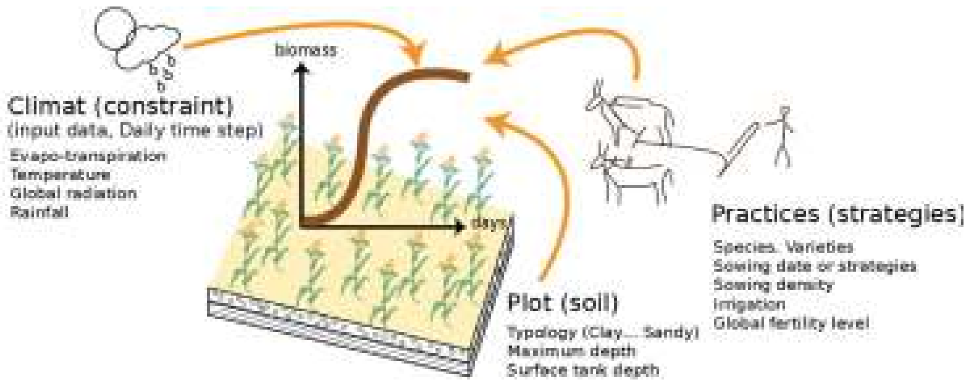


- Sarra-H, a crop model for maize
- Daily time step
- Attainable biomass and yields under **climatic constraints**

- Implementation under the **Ocelet Modelling Plateform**
- **ECMWF** agrometeorological data
- **TAMSAT** rainfall data

- Validated for the Tuy province [Akakpo 2017]

SARRA-O CROP MODEL [BARON ET AL., 2005; CASTETS ET AL., IN PREP]



- Sarra-H, a crop model for maize
- Daily time step
- Attainable biomass and yields under **climatic constraints**

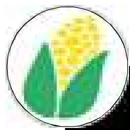
- Implementation under the **Ocelet Modelling Plateform**
- **ECMWF** agrometeorological data
- **TAMSAT** rainfall data

- Validated for the Tuy province [Akakpo 2017]

2011-2015

Aboveground biomass at flowering AGB-F

Crop Water stress from flowering to maturation Cstr



Attainable maize final yield

STATISTICAL MODELS AND STRATEGY

- *Above Ground Biomass at flowering (AGB-F) and water stress index (Cstr)*

AGB-F

Vegetation and drought indices – Vegetative period

STATISTICAL MODELS AND STRATEGY

- *Above Ground Biomass at flowering (AGB-F) and water stress index (Cstr)*

AGB-F	Vegetation and drought indices – Vegetative period
Cstr phase 4-5	Vegetation and drought indices – Productive period

STATISTICAL MODELS AND STRATEGY

- *Above Ground Biomass at flowering (AGB-F) and water stress index (Cstr)*

AGB-F

Vegetation and drought indices – Vegetative period →

Cstr phase 4-5

Vegetation and drought indices – Productive period →

Multi Linear Regression Model (MLR)

Random Forest Model (RF)

STATISTICAL MODELS AND STRATEGY

- *Above Ground Biomass at flowering (AGB-F) and water stress index (Cstr)*

AGB-F → Vegetation and drought indices – **Vegetative period** →

Cstr phase 4-5 → Vegetation and drought indices – **Productive period** →

Multi Linear Regression Model (MLR)

Random Forest Model (RF)

EVALUATION

- 10-fold cross validation
- Cv-R^2,cv-RMSE …

STATISTICAL MODELS AND STRATEGY

- *Above Ground Biomass at flowering (AGB-F) and water stress index (Cstr)*

AGB-F → Vegetation and drought indices – Vegetative period → Multi Linear Regression Model (MLR)

Cstr phase 4-5 → Vegetation and drought indices – Productive period → Random Forest Model (RF)

EVALUATION

- 10-fold cross validation
- Cv-R^2,cv-RMSE …

IMPORTANCE VARIABLES

- LMG method
- Mean decrease in MSE

STATISTICAL MODELS AND STRATEGY

- *Final maize yield*

STATISTICAL MODELS AND STRATEGY

- ***Final maize yield***

ESTIMATION

Final maize yield $\longrightarrow$ AGB-F x Cstr Phase 4-5

STATISTICAL MODELS AND STRATEGY

- *Final maize yield*

ESTIMATION

Final maize yield → AGB-F x Cstr Phase 4-5

FORECASTING

Final maize yield — Vegetation and drought indices – Vegetative period → MLR / RF

STATISTICAL MODELS AND STRATEGY

- *Final maize yield*

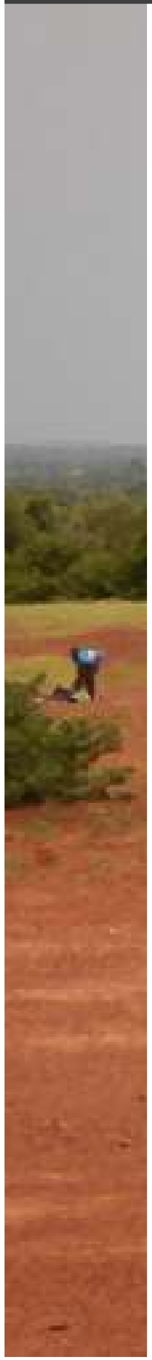
ESTIMATION

Final maize yield → AGB-F x Cstr Phase 4-5

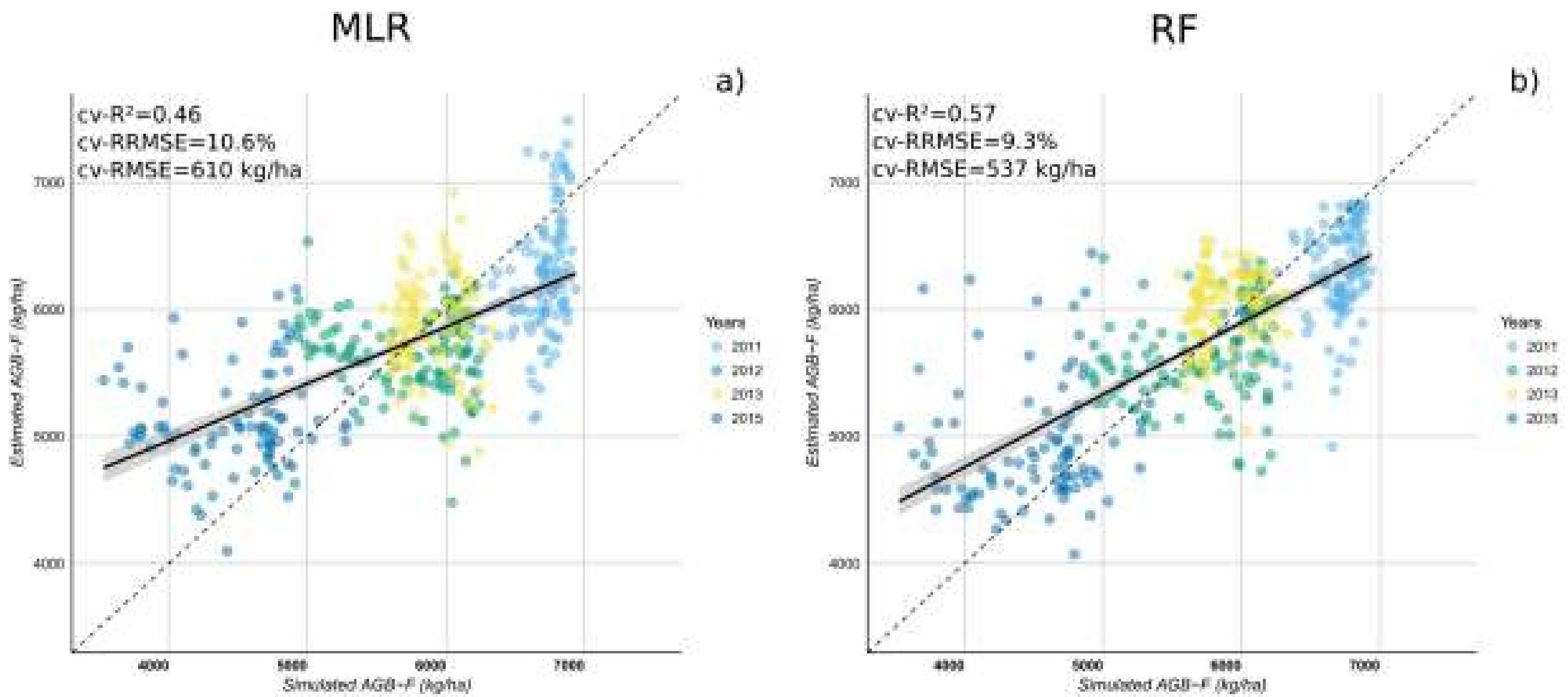
VALIDATION
- Yield from field survey
- Village scale

FORECASTING

Final maize yield — Vegetation and drought indices – Vegetative period → MLR / RF

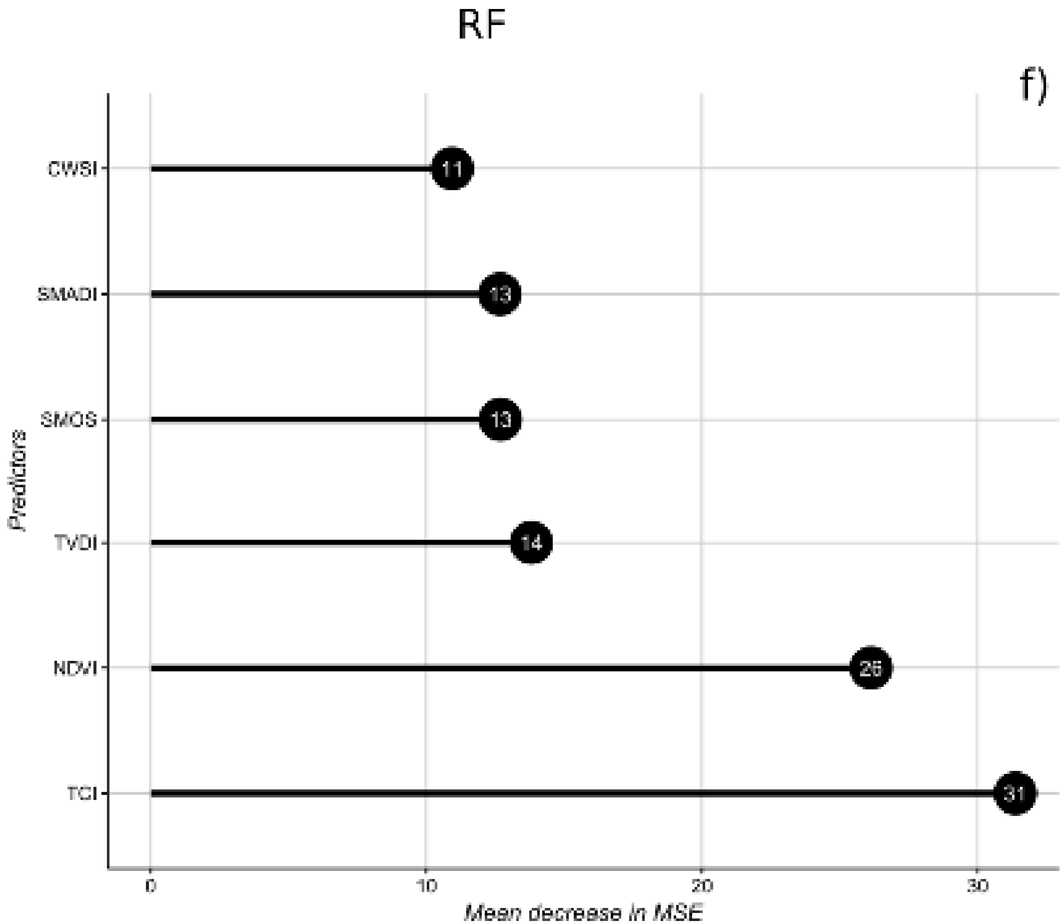


EVALUATION OF AGB-F ESTIMATION



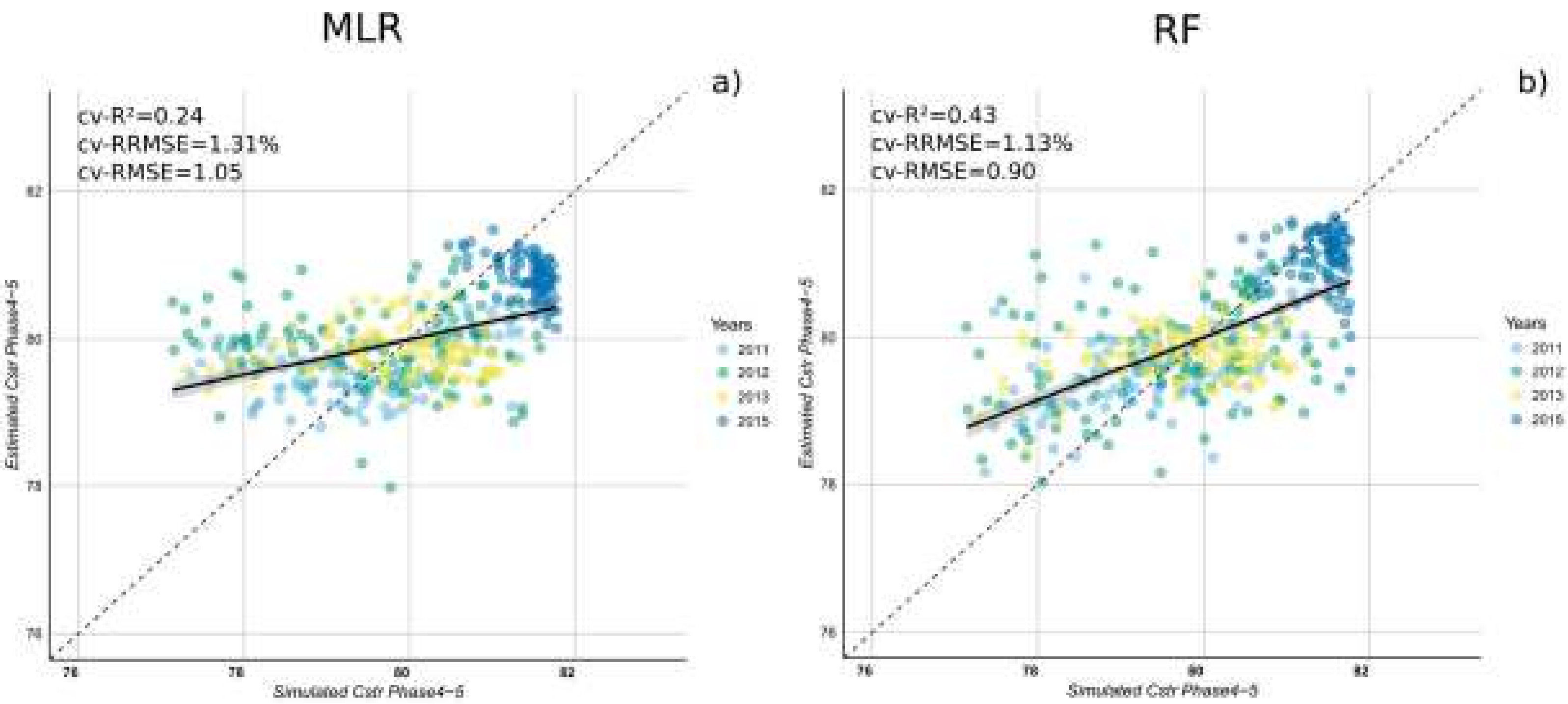
- Overestimation of low values and underestimation of high values
- RF model significantly better than the MLR model (**57%** of the AGB-F variability)

EVALUATION OF AGB-F ESTIMATION



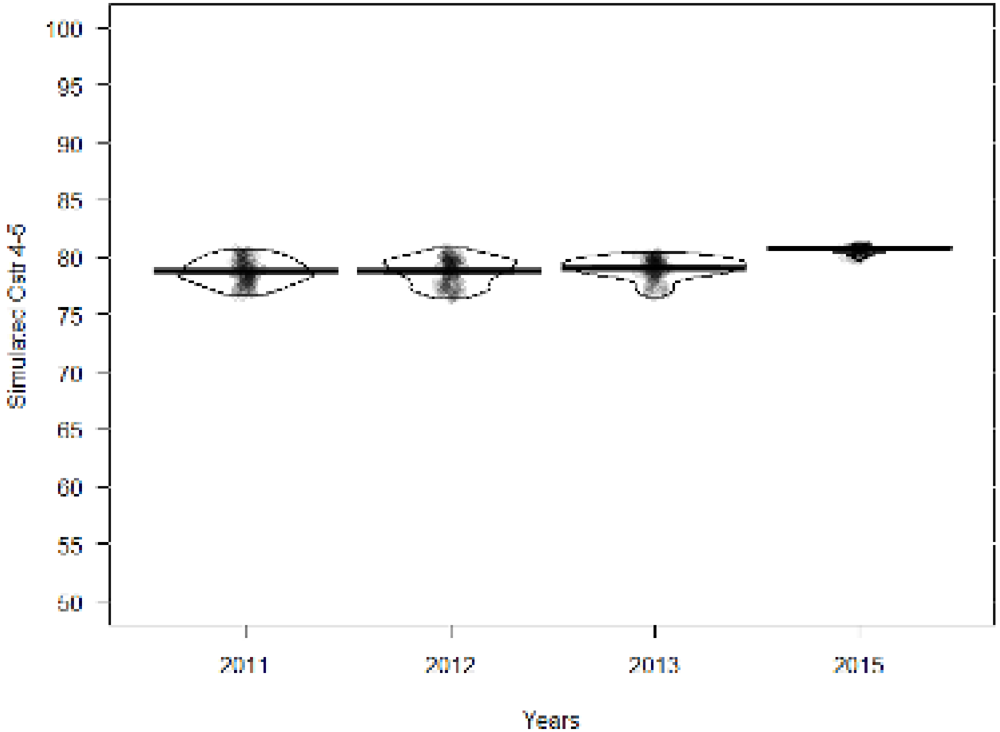
- Overestimation of low values and underestimation of high values
- RF model significantly better than the MLR model (**57%** of the AGB-F variability)
- Importance variable for RF = TCI and NDVI (57%)

EVALUATION OF CSTR PHASE 4-5 ESTIMATION



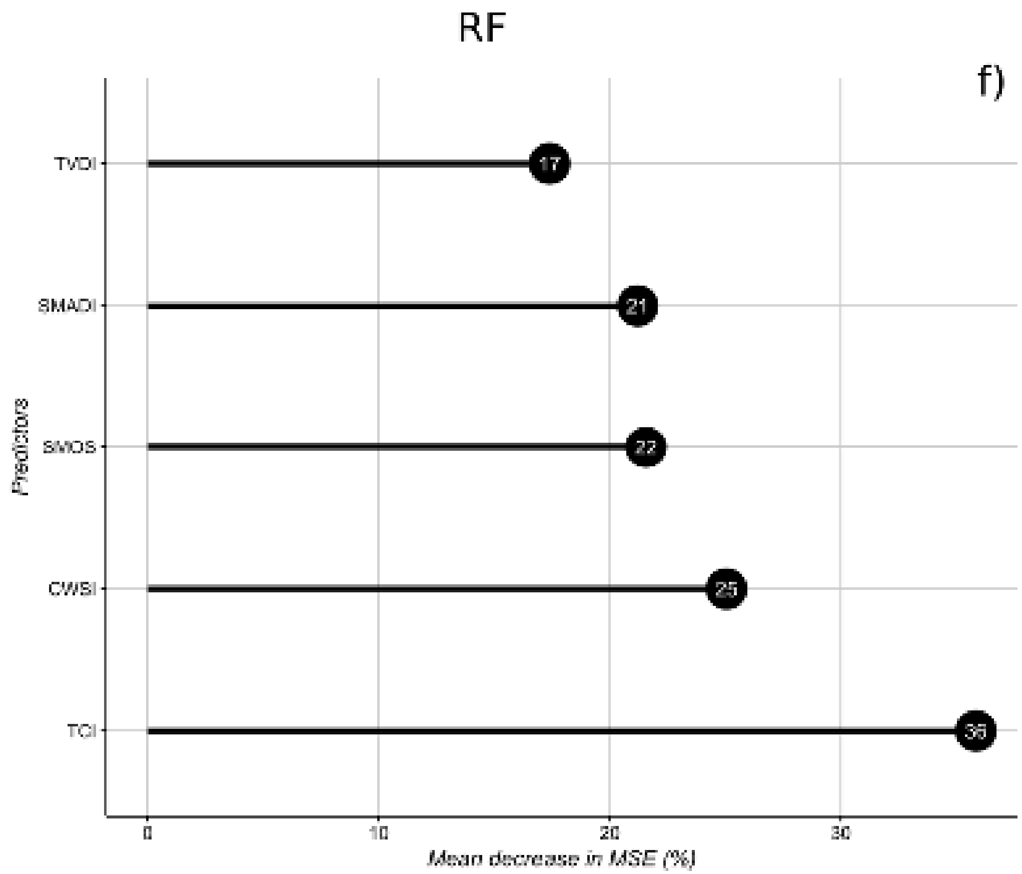
- Low but still significant predictive power
- But with cv-RRMSE < 2%? Low variability observed in SARRA-O Cstr!

EVALUATION OF CSTR PHASE 4-5 ESTIMATION



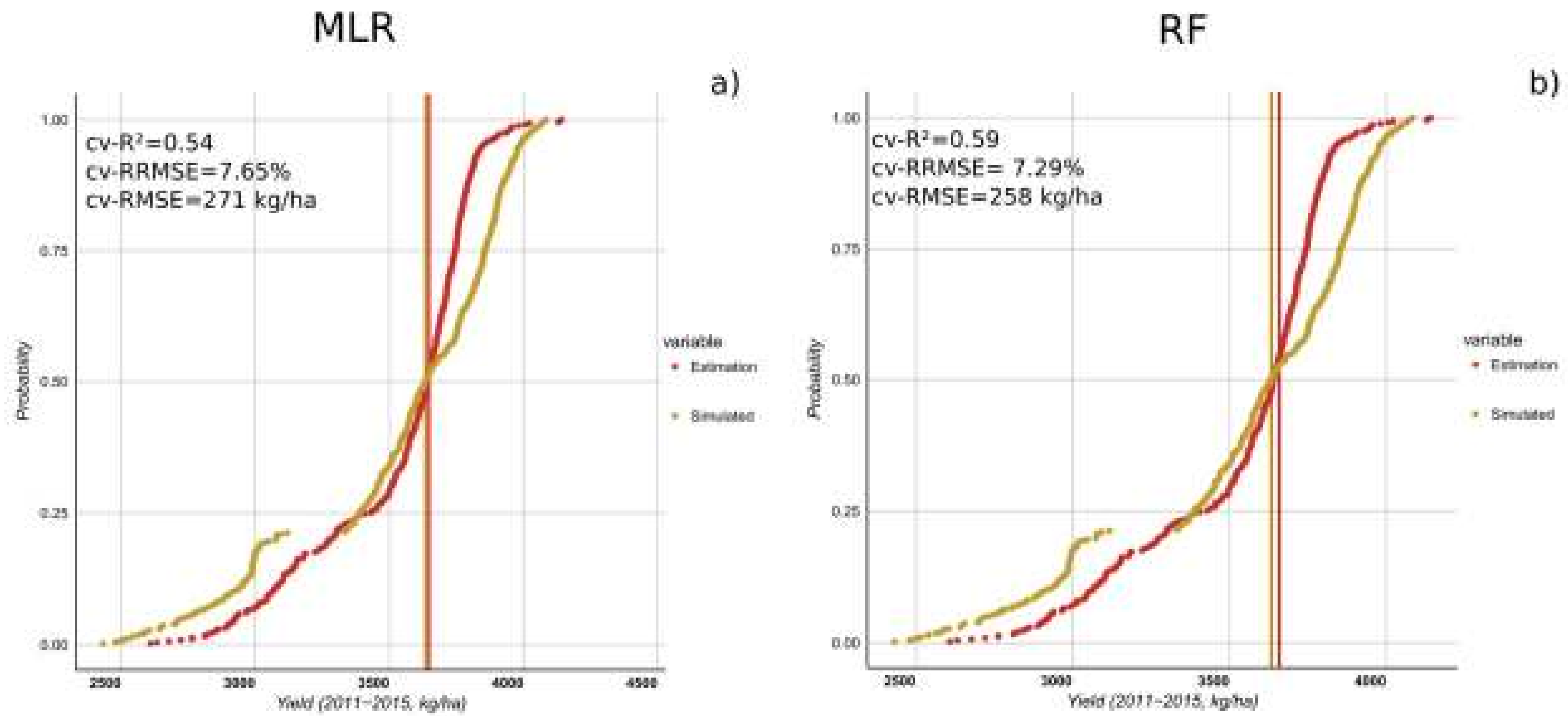
- Low but still significant predictive power
- But with cv-RRMSE < 2%? Low variability observed in SARRA-O Cstr!

EVALUATION OF CSTR PHASE 4-5 ESTIMATION



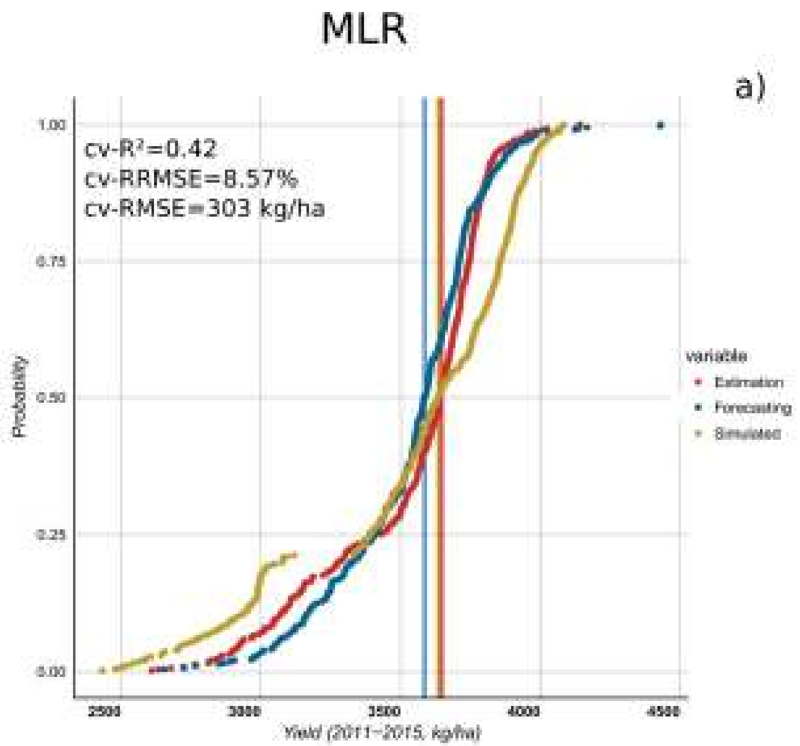
- Temperature Condition Index is the most important driver
- Impact of heat stress on maize : grain number [Eyshi Rezaei et al., 2015]

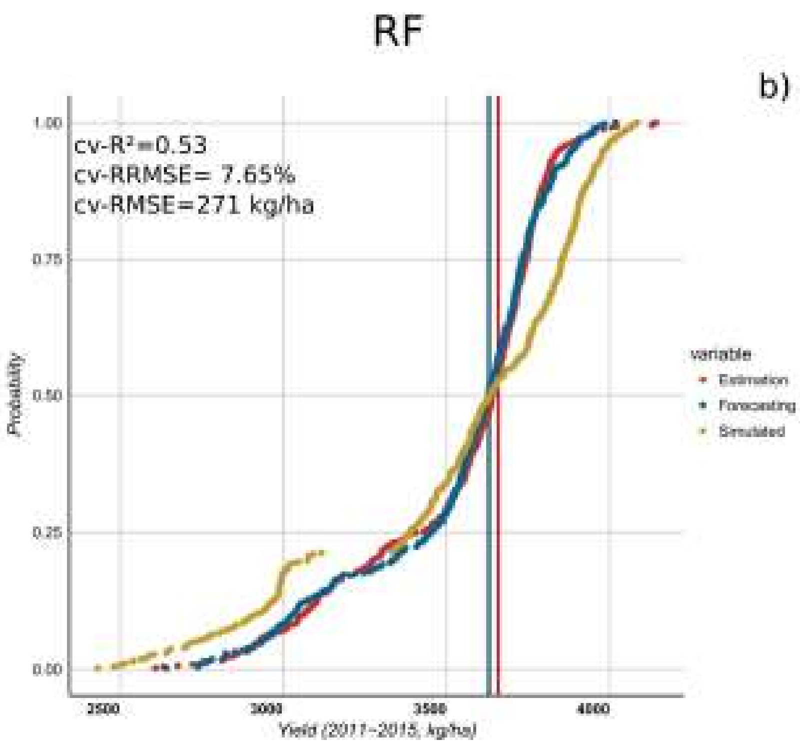
EVALUATION OF MAIZE YIELDS ESTIMATION AT THE END OF THE SEASON



- $Yield = f(AGB - F, Cstr\ Phase\ 4-5)$
- Good potential for maize yield estimation (RMSE<300 kg/ha)
- Good fitting of probability distribution curves :
    - Median SARRA-O : 3634 kg/ha
    - Median MLR : 3648 kg/ha
    - Median RF : 3659 kg/ha

EVALUATION OF EARLY ASSESSMENT OF MAIZE YIELDS





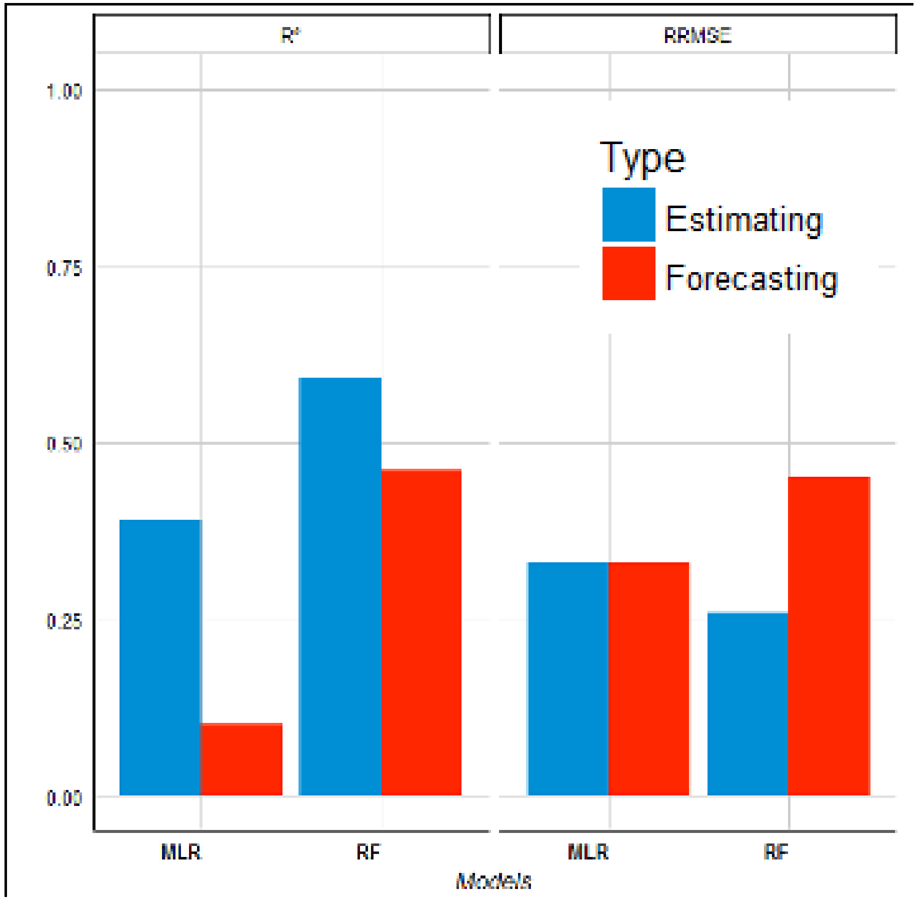
- $Yield = f(Remote\ Sensing\ Indices - Vegetative\ period)$
- Good potential for maize yield forcasting (RMSE<300 kg/ha)
- ~50% of maize yield variability can be explained ~2 months before harvest

VALIDATION OF MAIZE YIELDS WITH GROUND DATA

Independent dataset
*2014, 2015 and **2016***

VALIDATION OF MAIZE YIELDS WITH GROUND DATA

Independent dataset
*2014, 2015 and **2016***



- RF outperforms MLR:
  - Estimation : R^2=0.60
  - Forecasting : R^2=0.46
- High overestimation in forecasting
- MLR : 2016 not accuratly estimated

A GOOD AND EFFECTIVE POTENTIAL OF 'UNCALIBRATED APPROACH' TO ESTIMATE MAIZE YIELD IN SCARCE DATA ENVIRONMENT

A GOOD AND EFFECTIVE POTENTIAL OF 'UNCALIBRATED APPROACH' TO ESTIMATE MAIZE YIELD IN SCARCE DATA ENVIRONMENT

LOOK-BACK ON THE STUDY OBJECTIVES

- Linear *vs* NonLinear models:
  - Higher performance of **RF models** both for estimation and early assessment
  - Complex interaction among biophysical, ecological, physiological and management practices

A GOOD AND EFFECTIVE POTENTIAL OF 'UNCALIBRATED APPROACH' TO ESTIMATE MAIZE YIELD IN SCARCE DATA ENVIRONMENT

LOOK-BACK ON THE STUDY OBJECTIVES

- **Linear *vs* NonLinear models:**
  - Higher performance of **RF models** both for estimation and early assessment
  - Complex interaction among biophysical, ecological, physiological and management practices

- **Estimation vs Early assessment:**
  - Early assessment of maize yields **~ 2 months** before harvesting (RF)
  - Complementary of approaches :
    - **In-season forecasting** : food aids strategies or market and trade information
    - **After harvesting** : agricultural statistics
    - **Both + outputs of crop model** : 'convergence of evidence' in EWS

THANK YOU FOR LISTENING

Leroux Louise – CIRAD UR AÏDA/CSE
louise.leroux@cirad.fr
www.louise.leroux.igeo.fr

MORE INFORMATION :

LEROUX ET AL., 2017, *Maize grain yield estimating in a west African agricultural landscape at the cross-roads between remote sensing, crop modelling and statistical methods: Case study in south-west of Burkina Faso*, TO BE SUBMITTED TO AGRIC.FOR.METEOROL