

*Mise en relation de données  
hétérogènes pour le renforcement  
des systèmes de sécurité alimentaire  
– Cas de la production agricole en  
Afrique de l’Ouest*

Hugo Deléglise

*R. Interdonato , M.Roche, M. Teisseire, E. Maître D’Hotel, A. Bégué*

# 1. Introduction

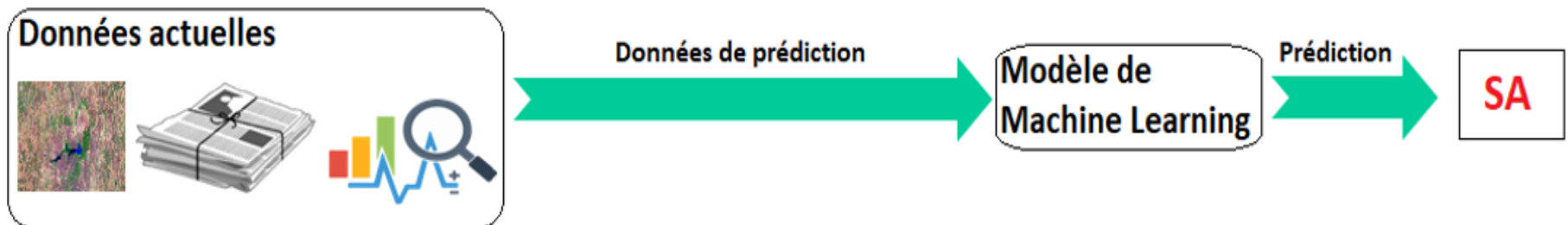
- **Pénuries alimentaires** restent un problème d'actualité en Afrique de l'Ouest
- Outil **opérationnel** d'aide à la décision pour systèmes de suivi de la **sécurité alimentaire** (SA)
- Nécessité de développer un **outil** de prédiction des pénuries alimentaires **rapide, fiable et gratuit**
- Capacité d'analyser conjointement des **données hétérogènes** car les causes de l'insécurité alimentaire sont **multifactorielles**

# 1. Introduction

- Définition **SA** : « *La SA existe lorsque tous les êtres ont, à tout moment, un accès physique et économique à une nourriture suffisante, saine et nutritive* » (source : WFP)
- 4 « Piliers » de la SA :
  - **Disponibilité** ; dépend de qualité des récoltes, qualité des relations avec les pays d'importation, etc.
  - **Accès** ; dépend du réseau routier, de la situation économique des ménages, etc.
  - **Stabilité d'accès** ; affectée par des chocs (sécheresses, guerres, etc.)
  - **Utilisation** ; dépend des habitudes nutritionnelles des ménages
- ➔ Divers indicateurs permettent d'évaluer la SA (score de diversité alimentaire (SDA), proportion des dépenses alimentaires, etc.)
- ➔ Divers sources de données doivent être exploitées pour **prédire** les indicateurs de SA et **expliquer** leurs multiples causes

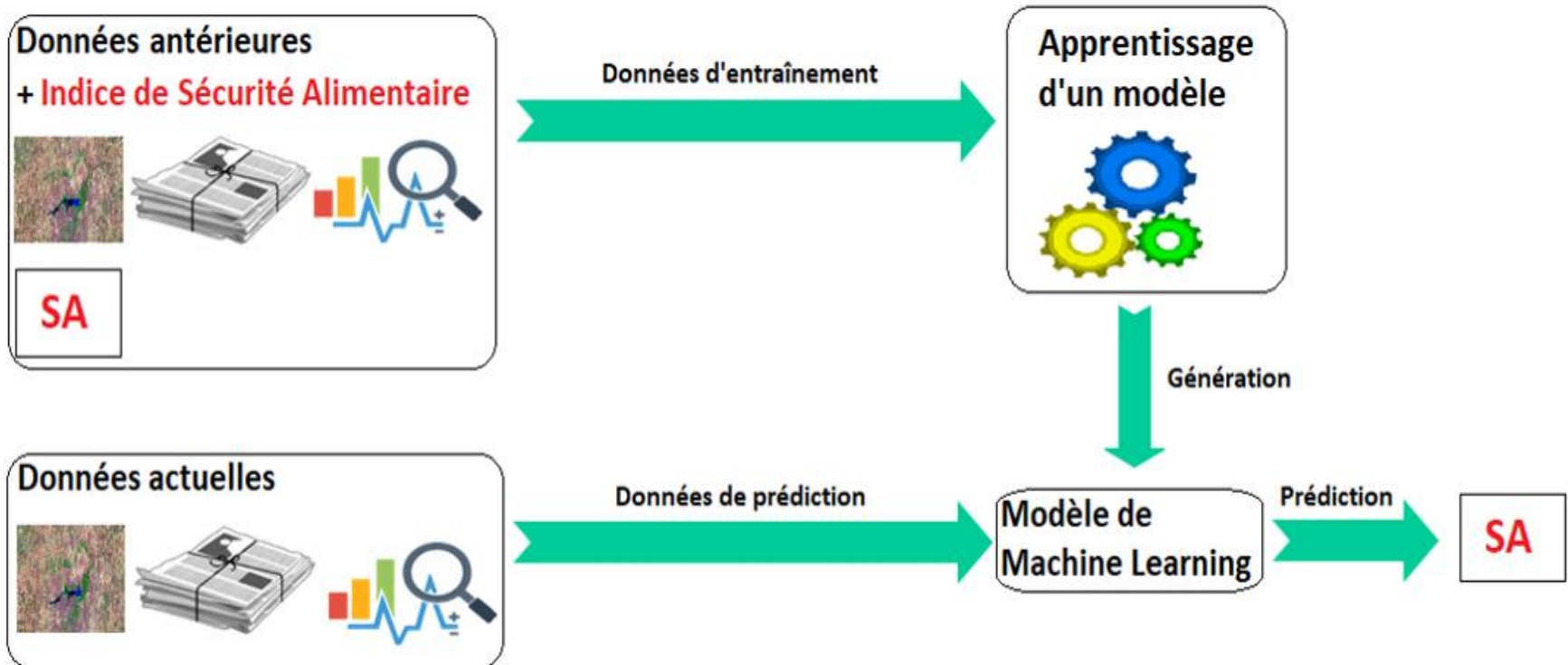
## 2. Matériel & méthodes

- Création d'un algorithme de **machine learning** (ML) supervisé capable de prédire les situations d'insécurité alimentaire



# 2. Matériel & méthodes

- Création d'un algorithme de **machine learning** (ML) supervisé capable de prédire les situations d'insécurité alimentaire



# 2. Matériel & méthodes

## Difficultés :

- Disponibilité des données :
  - Obtenir des indicateurs de SA pertinents et fiables
  - Obtenir des données explicatives porteuses d'informations et de qualité
- Méthodologie de traitement des données :
  - lier des sources de données hétérogènes
  - Traiter des séquences temporelles de variables répétées dans le temps (prendre en compte la temporalité)

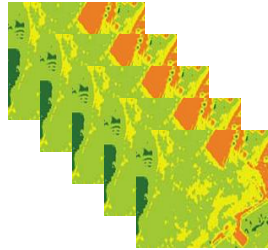
# 2. Matériel & méthodes

## Sources de données

## Extraction variables

## Prédiction

Images satellite :



Données textuelles :



Données quantitatives :



Enquêtes ménages :



0,2	0,1	0,6	0,5
0,0,6	0,1	0,8	0,0
0,0,2	0,3	0,5	0,9
0,0,9	0,5	0,5	0,1
0,0,6	0,2	0,3	0,6
0,0,7	0,7	0,2	0,9
0,1	0,4	0,6	0,3

0,2	0,1	0,6	0,5
0,0,6	0,1	0,8	0,0
0,0,2	0,3	0,5	0,9
0,0,9	0,5	0,5	0,1
0,0,6	0,2	0,3	0,6
0,0,7	0,7	0,2	0,9
0,1	0,4	0,6	0,3

0,2	0,1	0,6	0,5
0,0,6	0,1	0,8	0,0
0,0,2	0,3	0,5	0,9
0,0,9	0,5	0,5	0,1
0,0,6	0,2	0,3	0,6
0,0,7	0,7	0,2	0,9
0,1	0,4	0,6	0,3

0,6	0,4	0,8	0,0
0,5	0,2	0,3	0,4
0,8	0,5	0,9	0,2
0,5	0,5	0,5	0,3
0,4	0,7	0,2	0,6

**Modèles de ML :**  
(Régressions, SVM, random forest, Réseaux de neurones, etc.)

# 3. Résultats et discussion

## Données exploitées (Burkina Faso)

### Indicateurs de SA :

- SDA ; *Enquête permanente agricole (EPA)* ; Décembre 2009
- Valeur (F-CFA) des dépenses alimentaires des ménages ;  
*Living Standard Measurement Study (LSMS)* ; Septembre 2014

### Variables explicatives :

- NDVI (végétation) ; *Modis* ; 16 jours ; 250m
- Estimation des pluies ; *Trmm* ; fréq. 10 jours ; 6km
- Biomasse ; *Proba-V* ; fréq. 1an ; province
- Prix du maïs ; *SONAGES* ; fréq. 1 mois ; par marché
- Densité de population ; *AfriPop* ; 2014 ; 100m



# 3. Résultats et discussion

Apprentissage d'un Random Forest pour la prédiction du SDA et de la dépense alimentaire des ménages sur les 45 provinces.

**Dépenses alimentaires des ménages (LSMS ; 2014) :**

Variables explicatives	Nb Vars	R <sup>2</sup> (RF)
NDVI(t)	9	0,21
NDVI(t) ; pluies(t)	21	0,39
NDVI(t,t-1,t-2) ; pluies(t,t-1,t-2)	63	0,47
NDVI(t,t-1,t-2) ; pluies(t,t-1,t-2) ; prix mais(t,t-1)	81	0,50
NDVI(t,t-1,t-2) ; pluies(t,t-1,t-2) ; prix mais(t,t-1) ; densité pop	86	0,51
NDVI(t,t-1,t-2) ; Pluies(t,t-1,t-2) ; prix mais(t,t-1) ; densité pop ; indicSA	91	<b>0,56</b>

**SDA (EPA ; 2009) :**

Variables explicatives	Nb Vars	R <sup>2</sup> (RF)
NDVI(t)	14	0,22
NDVI(t) ; pluies(t)	35	0,25
NDVI(t,t-1,t-2) ; pluies(t,t-1,t-2)	105	0,29
NDVI(t,t-1,t-2) ; pluies(t,t-1,t-2) ; prix mais(t,t-1)	125	0,31
NDVI(t,t-1,t-2) ; Pluies(t,t-1,t-2) ; prix mais(t,t-1) ; biomasse (t,t-1,t-2)	128	<b>0,33</b>

# 3. Résultats et discussion

- Variables pertinentes :
  - NDVI et pluies durant la période de croissance
  - Prix du maïs des mois précédents l'enquête
- Dépendances avec l'indicateur de SA
  - SDA corrélé + avec les NDVI et pluies
  - SDA corrélé - avec les prix du maïs

➔ Perspectives ?

# 4. Conclusion & perspectives

- Données de télédétection ; économiques ; démographiques porteuses d'informations complémentaires pour expliquer les indicateurs de SA
- ➔ **Intégrer** de nouvelles variables au modèle (OpenStreetMap ; rendements, variables extraites d'articles de journaux, etc.)
- ➔ **Exploiter** un modèle de prédiction adapté au traitement de données hétérogènes et de séries temporelles (nouvelles variables de type séquentiel, index de phénologie, etc.)