

Documentation Complète

Dataset Forest Digital Twin

Écosystèmes Forestiers du Maroc

Systeme de Surveillance et d'Analyse
des Écosystèmes Forestiers Marocains

Date : 22 août 2025

Données Géospatiales Multi-Sources
Surveillance Satellite
Intelligence Artificielle

Table des matières

1	Vue d'ensemble	3
2	Structure des fichiers	3
2.1	Fichiers d'échantillonnage spatial (11 fichiers)	3
2.1.1	Régions principales (3 fichiers)	3
2.1.2	Haut Atlas subdivisé (4 fichiers)	3
2.1.3	Argan subdivisé (4 fichiers)	3
2.2	Fichiers de séries temporelles (5 fichiers)	3
3	Bandes Spectrales Sentinel-2	4
3.1	Bandes du Spectre Visible (400-700 nm)	4
3.2	Bandes Red Edge (700-800 nm)	4
3.3	Bandes Proche Infrarouge (800-1000 nm)	5
3.4	Bandes SWIR (1100-2200 nm)	5
4	Indices de Végétation	6
4.1	NDVI - Normalized Difference Vegetation Index	6
4.2	EVI - Enhanced Vegetation Index	6
4.3	NDMI - Normalized Difference Moisture Index	7
4.4	NBR - Normalized Burn Ratio	7
4.5	GNDVI - Green Normalized Difference Vegetation Index	7
4.6	SAVI - Soil Adjusted Vegetation Index (Argan uniquement)	8
5	Série Temporelle NDVI (2020-2024)	8
5.1	Variables temporelles annuelles	8
5.2	Métriques dérivées calculables	8
6	Variables Topographiques	9
6.1	Elevation (Altitude)	9
6.2	Slope (Pente)	9
6.3	Aspect (Orientation)	9
7	Variables Climatiques	10
7.1	Temperature (Température moyenne de l'air)	10
7.2	Precipitation (Précipitations cumulées)	10
8	Régions Géographiques	10
8.1	Rif (34.5°N - 35.5°N, 5.5°W - 4.0°W)	10
8.2	Moyen Atlas (33.0°N - 34.0°N, 5.5°W - 4.0°W)	11
8.3	Haut Atlas (30.5°N - 32.0°N, 8.0°W - 5.0°W)	11
8.4	Mamora (34.0°N - 34.3°N, 6.7°W - 6.2°W)	11
8.5	Argan (30.0°N - 31.0°N, 9.5°W - 8.0°W)	11

9	Guide d'Interprétation	12
9.1	Forêt en Bonne Santé	12
9.2	Forêt sous Stress Modéré	12
9.3	Forêt en Danger Critique	13
9.4	Syndromes de Dégradation	13
10	Seuils d'Alerte	13
10.1	Seuils critiques par région	13
10.2	Système d'alerte multi-niveaux	13
11	Applications Machine Learning	14
11.1	Features de base (Variables principales)	14
11.2	Features ingénierie avancée	14
11.3	Modèles recommandés par objectif	15
12	Spécifications Techniques	15
12.1	Sources et traitements	15
12.2	Contrôle qualité	15
12.3	Standards et formats	15

1 Vue d'ensemble

Ce dataset contient des données géospatiales multi-sources pour la surveillance et l'analyse de la santé des écosystèmes forestiers marocains. Il combine des observations satellitaires Sentinel-2, des données climatiques et topographiques pour créer une base de connaissances complète sur 5 régions forestières principales.

Caractéristiques du Dataset

- **Période d'observation** : Mars 2024 - Octobre 2024
- **Série temporelle** : 2020 - 2024 (NDVI)
- **Résolution spatiale** : 30 mètres
- **Système de coordonnées** : EPSG :4326 (WGS84)
- **Total des échantillons** : ~2,900 points géoréférencés
- **Couverture** : Principales forêts du Maroc (Rif, Atlas, Mamora, Argan)

2 Structure des fichiers

2.1 Fichiers d'échantillonnage spatial (11 fichiers)

2.1.1 Régions principales (3 fichiers)

- `samples_rif_2024.csv` - ~**500 points** - Chaîne du Rif (forêts méditerranéennes)
- `samples_moyen_atlas_2024.csv` - ~**500 points** - Moyen Atlas (forêts de montagne)
- `samples_mamora_2024.csv` - ~**500 points** - Forêt de Mamora (chêne-liège)

2.1.2 Haut Atlas subdivisé (4 fichiers)

- `samples_haut_atlas_ouest_2024.csv` - ~**200 points** - Partie occidentale
- `samples_haut_atlas_est_2024.csv` - ~**200 points** - Partie orientale
- `samples_haut_atlas_nord_2024.csv` - ~**200 points** - Versant nord
- `samples_haut_atlas_sud_2024.csv` - ~**200 points** - Versant sud

2.1.3 Argan subdivisé (4 fichiers)

- `samples_argan_nord_ouest_2024.csv` - ~**150 points** - Secteur nord-ouest
- `samples_argan_nord_est_2024.csv` - ~**150 points** - Secteur nord-est
- `samples_argan_sud_ouest_2024.csv` - ~**150 points** - Secteur sud-ouest
- `samples_argan_sud_est_2024.csv` - ~**150 points** - Secteur sud-est

2.2 Fichiers de séries temporelles (5 fichiers)

- `timeseries_samples_rif_2024.csv` - Évolution NDVI 2020-2024 Rif
- `timeseries_samples_moyen_atlas_2024.csv` - Évolution NDVI 2020-2024 Moyen Atlas

- `timeseries_samples_mamora_2024.csv` - Évolution NDVI 2020-2024 Mamora
- `timeseries_samples_haut_atlas_2024.csv` - Évolution NDVI 2020-2024 Haut Atlas
- `timeseries_samples_argan_2024.csv` - Évolution NDVI 2020-2024 Argan

3 Bandes Spectrales Sentinel-2

3.1 Bandes du Spectre Visible (400-700 nm)

Bande	Longueur d’onde	Résolution	Description	Utilité forestière
B2	490 nm (Bleu)	10m	Diffusion at-mosphérique	<ul style="list-style-type: none">— Détection de l’eau— Discrimination végétation/sol— Calcul de l’EVI
B3	560 nm (Vert)	10m	Pic de réflectance verte	<ul style="list-style-type: none">— Vigueur de la végétation— Contenu en chlorophylle— Base du GNDVI
B4	665 nm (Rouge)	10m	Absorption chlorophylle	<ul style="list-style-type: none">— Photosynthèse active— Base du NDVI— Stress végétal

3.2 Bandes Red Edge (700-800 nm)

Bande	Longueur d’onde	Résolution	Description	Utilité forestière
B5	705 nm (RE1)	20m	Début red edge	<ul style="list-style-type: none">— Détection précoce du stress— Contenu en azote— État physiologique

Bande	Longueur d'onde	Résolution	Description	Utilité forestière
B6	740 nm (RE2)	20m	Centre red edge	<ul style="list-style-type: none"> — Concentration LAI — Biomasse — Santé foliaire
B7	783 nm (RE3)	20m	Fin red edge	<ul style="list-style-type: none"> — Contenu en eau des feuilles — Sénescence — Maladies

3.3 Bandes Proche Infrarouge (800-1000 nm)

Bande	Longueur d'onde	Résolution	Description	Utilité forestière
B8	842 nm (NIR)	10m	NIR principal	<ul style="list-style-type: none"> — Structure cellulaire — Densité végétale — Base indices végétation
B8A	865 nm	20m	NIR étroit	<ul style="list-style-type: none"> — Vapeur d'eau atmosphérique — Classification espèces — Biomasse

3.4 Bandes SWIR (1100-2200 nm)

Bande	Longueur d'onde	Résolution	Description	Utilité forestière
B11	1610 nm (SWIR1)	20m	Humidité	<ul style="list-style-type: none"> — Stress hydrique — Contenu en eau — Base du NDMI

Bande	Longueur d'onde	Résolution	Description	Utilité forestière
B12	2190 nm (SWIR2)	20m	Zones brûlées	<ul style="list-style-type: none"> — Détection incendies — Minéraux du sol — Base du NBR

4 Indices de Végétation

4.1 NDVI - Normalized Difference Vegetation Index

Formule et Caractéristiques

Formule : $NDVI = \frac{B8-B4}{B8+B4}$

Plage : -1 à +1

Interprétation écologique :

- < 0 : Eau, neige, nuages
- 0 - 0.2 : Sol nu, roches, zones urbaines
- 0.2 - 0.4 : Végétation clairsemée ou stressée
- 0.4 - 0.6 : Végétation modérément dense
- 0.6 - 0.8 : Végétation dense et saine
- > 0.8 : Végétation très dense (forêt tropicale)

Applications forestières :

- Suivi de la déforestation
- Détection de stress (sécheresse, maladies)
- Estimation de la biomasse
- Phénologie et saisonnalité

4.2 EVI - Enhanced Vegetation Index

Formule et Caractéristiques

Formule : $EVI = 2.5 \times \frac{B8-B4}{B8+6 \times B4 - 7.5 \times B2 + 1}$

Plage : -1 à +1

Avantages sur NDVI :

- Moins sensible à la saturation en zones denses
- Correction des effets atmosphériques et du sol
- Meilleur pour les forêts denses du Rif et Moyen Atlas
- Plus sensible aux variations de structure végétale

4.3 NDMI - Normalized Difference Moisture Index

Formule et Caractéristiques

Formule : $NDMI = \frac{B8-B11}{B8+B11}$

Plage : -1 à +1

Interprétation hydrique :

- < **-0.2** : Stress hydrique extrême
- **-0.2 - 0** : Stress hydrique modéré à sévère
- **0 - 0.2** : Hydratation normale
- **0.2 - 0.4** : Bonne hydratation
- > **0.4** : Saturation en eau

Applications critiques :

- Prédiction du risque d'incendie
- Détection précoce de la sécheresse
- Monitoring stress hydrique en zones arides (Argan)
- Vulnérabilité aux parasites et maladies

4.4 NBR - Normalized Burn Ratio

Formule et Caractéristiques

Formule : $NBR = \frac{B8-B12}{B8+B12}$

Plage : -1 à +1

Classification des brûlures :

- < **-0.2** : Zone récemment brûlée (destruction sévère)
- **-0.2 - 0.1** : Sol nu ou zones brûlées anciennes
- **0.1 - 0.3** : Végétation basse/repousse post-incendie
- > **0.3** : Végétation non brûlée

4.5 GNDVI - Green Normalized Difference Vegetation Index

Formule et Caractéristiques

Formule : $GNDVI = \frac{B8-B3}{B8+B3}$

Plage : -1 à +1

Spécificités :

- Plus sensible à la chlorophylle que NDVI
- Détection précoce du stress nutritionnel
- Discrimination fine des espèces forestières
- Monitoring de la phénologie automnale

4.6 SAVI - Soil Adjusted Vegetation Index (Argan uniquement)

Formule et Caractéristiques

Formule : $SAVI = \frac{B8-B4}{B8+B4+0.5} \times 1.5$

Plage : -1.5 à +1.5

- Adaptation aux zones arides :
- Correction de l’influence du sol nu
 - Spécialement calibré pour l’écosystème arganier
 - Améliore la précision en zone de couverture végétale faible

5 Série Temporelle NDVI (2020-2024)

5.1 Variables temporelles annuelles

Variable	Description	Contexte climatique Maroc
NDVI_2020	NDVI moyen année 2020	Année de référence, pluviométrie normale
NDVI_2021	NDVI moyen année 2021	Début de période de sécheresse
NDVI_2022	NDVI moyen année 2022	Pic de sécheresse historique au Maroc
NDVI_2023	NDVI moyen année 2023	Récupération partielle, pluviométrie irrégulière
NDVI_2024	NDVI moyen année 2024	État actuel, monitoring continu

5.2 Métriques dérivées calculables

Listing 1 – Calcul des métriques temporelles

```
# Tendance lin aire sur 5 ans
ndvi_trend = (NDVI_2024 - NDVI_2020) / 4

# Volatilit interannuelle
ndvi_volatility = std([NDVI_2020, NDVI_2021, NDVI_2022, NDVI_2023,
    NDVI_2024])

# R silience post-s cheresse
resilience_index = (NDVI_2024 - NDVI_2022) / NDVI_2022

# Ann es de d clin cons cutives
decline_years = count(NDVI[i] $<$NDVI[i-1] for consecutive years)

# Impact s cheresse 2022
drought_impact = (NDVI_2020 - NDVI_2022) / NDVI_2020
```

6 Variables Topographiques

6.1 Elevation (Altitude)

- **Unité** : Mètres au-dessus du niveau de la mer
- **Source** : SRTM 30m Global DEM
- **Plage Maroc** : 0 - 4,165m (Jbel Toubkal)

Zonation altitudinale forestière :

- **<800m** : Forêts thermophiles (chêne-liège, arganier, oléastre)
- **800-1500m** : Forêts mésophiles (chêne vert, thuya, caroubier)
- **1500-2500m** : Forêts montagnardes (cèdre de l'Atlas, sapin du Maroc)
- **2500-3200m** : Limite forestière (genévriers, épineux)
- **>3200m** : Végétation alpine clairsemée

6.2 Slope (Pente)

- **Unité** : Degrés (0-90°)
- **Calcul** : Dérivée de l'élévation SRTM

Classification géomorphologique :

- **0-5°** : Terrain plat (vallées, plateaux)
- **5-15°** : Pente douce (coteaux, piémonts)
- **15-30°** : Pente modérée (versants)
- **30-45°** : Pente forte (escarpements)
- **>45°** : Pente très forte (falaises, ravins)

6.3 Aspect (Orientation)

- **Unité** : Degrés azimutaux (0-360°)
- **Référence** : 0°/360°=Nord, 90°=Est, 180°=Sud, 270°=Ouest

Orientation	Angle	Ensoleillement	Humidité	Végétation typique
Nord	315-45°	Faible	Élevée	Cèdre, sapin, hêtre
Est	45-135°	Matinal	Modérée	Forêts mixtes
Sud	135-225°	Maximum	Faible	Chêne vert, pin, thuya
Ouest	225-315°	Après-midi	Variable	Formations de transition

7 Variables Climatiques

7.1 Temperature (Température moyenne de l'air)

- **Unité** : Kelvin (conversion : °C = K - 273.15)
- **Source** : ERA5 Daily (ECMWF)
- **Résolution** : 0.25° (~27km)
- **Période** : Moyenne mars-octobre 2024
- Zonation thermique forestière Maroc :**
 - <10°C : Étage montagnard supérieur
 - 10-15°C : Étage montagnard moyen
 - 15-20°C : Étage montagnard inférieur/méditerranéen
 - 20-25°C : Étage thermo-méditerranéen
 - >25°C : Étage aride/semi-aride

7.2 Precipitation (Précipitations cumulées)

- **Unité** : Millimètres (mm)
- **Source** : CHIRPS Daily v2.0
- **Résolution** : 0.05° (~5km)
- **Période** : Cumul mars-octobre 2024

Isohyète	Régime	Bioclimat	Formations forestières
<200mm/an	Aride	Saharien	Acacias épineux, steppes
200-400mm/an	Semi-aride	Steppique	Arganier, alfa, jujubier
400-600mm/an	Sub-humide	Méditerranéen	Thuya, chêne vert, oléastre
600-800mm/an	Humide	Méditerranéen humide	Chêne-liège, chêne zéen
>800mm/an	Très humide	Méditerranéen perhumide	Cèdre, sapin, chêne tauzin

8 Régions Géographiques

8.1 Rif (34.5°N - 35.5°N, 5.5°W - 4.0°W)

Caractéristiques biogéographiques :

- **Écosystème** : Forêts méditerranéennes humides
- **Espèces dominantes** : Chêne-liège (*Quercus suber*), Pin maritime (*Pinus pinaster*)
- **Altitude** : 200 - 2,000m
- **Climat** : Méditerranéen humide (600-1200mm/an)
- **Sols** : Schistes, grès, calcaires
- **Menaces** : Incendies, surpâturage, dépérissement du chêne-liège

8.2 Moyen Atlas (33.0°N - 34.0°N, 5.5°W - 4.0°W)

Caractéristiques biogéographiques :

- **Écosystème** : Forêts de montagne continentales
- **Espèces dominantes** : Cèdre de l'Atlas (*Cedrus atlantica*), Chêne vert (*Quercus ilex*)
- **Altitude** : 800 - 3,000m
- **Climat** : Continental montagnard (400-800mm/an)
- **Sols** : Calcaires, basaltes, marnes
- **Menaces** : Dépérissement du cèdre, sécheresse, surexploitation

8.3 Haut Atlas (30.5°N - 32.0°N, 8.0°W - 5.0°W)

Caractéristiques biogéographiques :

- **Écosystème** : Forêts de haute montagne arides
- **Espèces dominantes** : Thuya (*Tetraclinis articulata*), Genévrier (*Juniperus phoenicea*)
- **Altitude** : 1,000 - 4,000m
- **Climat** : Montagnard aride (200-600mm/an)
- **Sols** : Calcaires, schistes, conglomérats
- **Menaces** : Érosion, surpâturage, changement climatique

8.4 Mamora (34.0°N - 34.3°N, 6.7°W - 6.2°W)

Caractéristiques biogéographiques :

- **Écosystème** : Forêt de plaine subéreuse
- **Espèces dominantes** : Chêne-liège (*Quercus suber*) en monoculture
- **Altitude** : 0 - 200m
- **Climat** : Méditerranéen atlantique (400-600mm/an)
- **Sols** : Sables, argiles, limons
- **Menaces** : Dépérissement, urbanisation, gestion intensive

8.5 Argan (30.0°N - 31.0°N, 9.5°W - 8.0°W)

Caractéristiques biogéographiques :

- **Écosystème** : Forêt aride endémique (Réserve de Biosphère UNESCO)
- **Espèces dominantes** : Arganier (*Argania spinosa*) endémique
- **Altitude** : 0 - 1,500m
- **Climat** : Aride à semi-aride (150-400mm/an)
- **Sols** : Calcaires, schistes, alluvions
- **Menaces** : Surexploitation, défrichement, sécheresse

9 Guide d'Interprétation

9.1 Forêt en Bonne Santé

Signatures spectrales et indicateurs temporels

Signatures spectrales :

- $\text{NDVI} > 0.6$
- $\text{NDMI} > 0.2$
- $\text{EVI} > 0.4$
- $\text{NBR} > 0.3$

Indicateurs temporels :

- Tendence NDVI stable ou positive
- Faible variabilité interannuelle ($\text{CV} < 0.1$)
- Récupération rapide post-stress

9.2 Forêt sous Stress Modéré

Signatures spectrales et indicateurs temporels

Signatures spectrales :

- $\text{NDVI} : 0.4 - 0.6$
- $\text{NDMI} : 0 - 0.2$
- $\text{EVI} : 0.2 - 0.4$

Indicateurs temporels :

- Tendence NDVI légèrement négative (-0.01 à $-0.05/\text{an}$)
- Variabilité modérée ($\text{CV} : 0.1 - 0.2$)
- Récupération lente

9.3 Forêt en Danger Critique

Signatures spectrales et indicateurs temporels

Signatures spectrales :

- NDVI < 0.4
- NDMI < 0
- NBR < 0.1

Indicateurs temporels :

- Tendence NDVI fortement négative (<-0.05/an)
- Forte variabilité (CV > 0.2)
- Absence de récupération

9.4 Syndromes de Dégradation

Syndrome	NDVI	NDMI	NBR	EVI	Cause probable
Sécheresse chronique	↓	↓↓	→	↓	Déficit hydrique prolongé
Post-incendie récent	↓↓	↓	↓↓	↓↓	Combustion récente
Maladie/Parasites	↓	→	→	↓	Agents pathogènes
Déforestation	↓↓	↓	↓	↓↓	Coupe/défrichement
Surpâturage	↓	↓	→	↓	Pression pastorale

10 Seuils d'Alerte

10.1 Seuils critiques par région

Région	NDVI critique	NDMI critique	EVI critique	Action recommandée
Rif	< 0.5	< 0.1	< 0.3	Surveillance incendies renforcée
Moyen Atlas	< 0.45	< 0.05	< 0.25	Monitoring dépérissement cèdre
Haut Atlas	< 0.4	< 0	< 0.2	Lutte anti-érosive
Mamora	< 0.4	< 0	< 0.2	Diagnostic sanitaire

10.2 Système d'alerte multi-niveaux

Niveau	Couleur	Condition	Action
Normal	Vert	Tous indices >seuils	Monitoring de routine
Attention	Jaune	1 indice <seuil	Surveillance renforcée
Alerte	Orange	2+ indices <seuils	Investigation terrain
Critique	Rouge	Tous indices <seuils	Intervention urgente

11 Applications Machine Learning

11.1 Features de base (Variables principales)

Listing 2 – Features principales pour ML

```
features_spectral = ['B2', 'B3', 'B4', 'B5', 'B6', 'B7', 'B8', 'B8A', 'B11', 'B12']
features_indices = ['NDVI', 'EVI', 'NDMI', 'NBR', 'GNDVI', 'SAVI']
features_topo = ['elevation', 'slope', 'aspect']
features_climate = ['temperature', 'precipitation']
features_temporal = ['NDVI_2020', 'NDVI_2021', 'NDVI_2022', 'NDVI_2023', 'NDVI_2024']
```

11.2 Features ingénierie avancée

Listing 3 – Ingénierie des features

```
# Ratios spectraux
B8_B4_ratio = B8 / B4 # Vegetation index simple
B11_B8_ratio = B11 / B8 # Moisture content
B12_B11_ratio = B12 / B11 # Burn sensitivity

# Indices temporels
ndvi_trend = (NDVI_2024 - NDVI_2020) / 4
ndvi_volatility = np.std([NDVI_2020, NDVI_2021, NDVI_2022, NDVI_2023, NDVI_2024])
drought_impact = (NDVI_2020 - NDVI_2022) / NDVI_2020
resilience_index = (NDVI_2024 - NDVI_2022) / NDVI_2022

# Interactions topographiques
slope_aspect_north = slope * np.cos(np.radians(aspect)) # Northness
slope_aspect_east = slope * np.sin(np.radians(aspect)) # Eastness
elevation_precipitation_ratio = elevation / precipitation
topographic_wetness_index = np.log((precipitation + 1) / (slope + 0.1))

# Indices composites
forest_health_score = (NDVI + NDMI) / 2
stress_index = 1 - (NDMI + 1) / 2 # 0=no stress, 1=max stress
fire_risk_index = (1 - NDMI) * (1 - NBR) * temperature / 300
```

11.3 Modèles recommandés par objectif

Objectif	Features prioritaires	Algorithme	Métriques
Classification santé	Indices + Temporel	Random Forest	F1-score, Accuracy
Détection incendies	NDMI, NBR, Météo	XGBoost	Précision, Rappel
Prédiction stress	NDMI, Tendances	LSTM/GRU	RMSE, MAE
Segmentation espèces	Red Edge, Topo	CNN	IoU, Dice
Analyse tendances	Série temporelle	Prophet, ARIMA	MAPE, R^2

12 Spécifications Techniques

12.1 Sources et traitements

Donnée	Source	Résolution	Prétraitement
Imagerie	Sentinel-2 L2A (ESA)	10-20m	Composite médian, masquage nuages
Météo	ERA5 Daily (ECMWF)	0.25°	Moyenne temporelle
Précipitations	CHIRPS v2.0 (UCSB)	0.05°	Cumul période
Topographie	SRTM GL1 v3 (NASA)	30m	Dérivées (pente, exposition)

12.2 Contrôle qualité

- **Précision géospatiale** : $\pm 15\text{m}$ (Sentinel-2)
- **Complétude temporelle** : $>90\%$
- **Validation radiométrique** : Cohérence inter-bandes
- **Filtrage atmosphérique** : Algorithme Sen2Cor
- **Détection nuages/ombres** : QA60 quality flags

12.3 Standards et formats

- **Projection** : EPSG :4326 (WGS84 Geographic)
- **Format fichiers** : CSV (UTF-8)
- **Encodage numérique** : Float64 (variables continues)
- **Métadonnées** : ISO 19115-2 (géospatial)
- **Interopérabilité** : OGC Standards (WMS, WFS)