Crop Monitoring as an
E-agricutural tool in
Developping Countries

DATABASE CONTAINING NDVI, DMP AND WHEAT STATISTICS FOR MOROCCO

Reference: ***E-AGRI D43.1 Database containing NDVI, DMP and wheat statistics for Morocco.***

Author(s): Riad Balaghi, Herman Eerens, Qinghan Dong

Version: 1.0

Date: 26/02/2014

DOCUMENT CONTROL

Signatures

Author(s) :

Reviewer(s) :

Approver(s) :

Issuing authority :

Change record

Release	Date	Pages	Description	Editor(s)/Reviewer(s)
1.0	26/02/2014	32	Database containing NDVI, DMP and wheat statistics for Morocco	

TABLE OF CONTENT

ACRONYMS & CLOSSARY 6

1. EXECUTIVE SUMMARY 7

2. INTRODUCTION 8

3. NDVI AND DMP STATISTICS 9

3.1. Some backgrounds 9

3.1.1. SPOT-VEGETATION (VGT) & CTIV 9

3.1.2. JRC-MARS & VITO's Application Unit 9

3.1.3. Some Technical Remarks 11

3.2. Ancillary data & study area 12

3.2.1. Global ancillary data 12

3.2.2. Study area 12

3.3. Preparation & adaptation of the RS-data 15

3.4. Final databases with RUM values 15

3.5. DATA SETS DELIVERED 17

4. WHEAT STATISTICS 20

LIST OF FIGURES

Figure 1: Moroccan study area. Morocco (blue) and its provinces (red) and the analysed image block, in this case with the cropland AF (A). The 38 considered provinces and the Region-IDs (B). Digital terrain model GTOPO30 (C). The GLC2000 land use map. White=cropland in the northwest, green/brown=forests and shrubs mainly in the Atlas mountains, grey=desertic zones in the south and east, red=urban areas (D). 13

Figure 2: Typical cereal cropping season in Morocco in terms of phenology, rainfall and temperature (figure provided by INRA-Morocco). 14

Figure 3: Example of the three different formats for the final tables with RUM-values..... 16

LIST OF TABLES

Table 1: Image types (DT=Datatype: 1=unsigned byte with V=0→255, 2=signed short integer with V=-32768→+32767). 11

Table 2: The considered Moroccan provinces and their Region_ID numbers. 13

ACRONYMS & CLOSSARY

AFI	Area Fraction Image
AFS	Area Frame Sampling
DMP	Dry Matter Productivity
DSS	Direction de la Stratégie et des Statistiques
E-AGRI	A research Project on Crop Monitoring as an E-agriculture tool in Developing Countries www.e-agri.info
fAPAR	Fraction of solar radiation absorbed by plants
GLC2000	Global Land Cover for the year 2000
GlCropV2	Land cover map made of in collaboration with VITO
INRA	Institut National de la Recherche Agronomique - Morocco
LAI	Leaf Area Index
NDVI	Normalized Difference Vegetation Index
SPOT-VEGETATION	Programme conceived to allow daily monitoring of terrestrial vegetation cover through remote sensing, at regional to global scales http://www.vgt.vito.be/pages/mission.htm
SRTM	Shuttle Radar Topography Mission: www2.jpl.nasa.gov/srtm
VITO	Flemish institute for technological research, Belgium

1. EXECUTIVE SUMMARY

The objectives of this deliverable are to:

- Collect 10 days SPOT-VEGETATION NDVI or DMP from 1999 to 2009 ;
- Define spatial extent of agricultural areas of Morocco based on GlobCover map ;
- Extract 10 days SPOT-VEGETATION NDVI or DMP from 1999 to 2009, corresponding to agricultural areas of Morocco ;
- Calculate mean NDVI or DMP values for agricultural areas of Morocco at national level and provincial level ;
- Collect wheat area and production statistics at provincial level ;
- Build database containing NDVI or DMP and wheat statistics.

All these deliverables have been completely achieved:

1. SHP-files with the boundaries of the country and its 38 provinces converted to the INSPIRE-LAEA projection (compatible with the VGTp-imagery). The associated DBF-file also includes the Region_ID numbers.
2. Files MAROC*n*.csv, with *n*=1/2/3, containing the RUM-values in the three different formats.
3. Official statistics: File “cereales 1979-2012_Bis.xls” was provided by INRA. The TXT-files were derived from it. They contain one line for each Region x Year, and each such line comprises three comma-separated values: Region_ID (table 2), Year and official yield (in qx/ha). Negative values are missing. There are four such files for each of the crops: BLE_DUR (durum wheat), BLE_TEN (soft wheat), ORGE (barley) and CER_AUT (total of soft wheat, durum wheat and barley).
4. Area and yield data of the three main cereal crops, soft wheat, durum wheat and barley, were graciously provided by “la Direction de la Stratégie et des Statistiques”. They are available for 40 provinces of the country for the period of 1978-2012. Production at a province level is obtained by multiplying the yield value with the area estimated by DSS.

2. INTRODUCTION

This deliverable is part of Deliverable D.43 "*Yield estimation for wheat based on remote sensing in Morocco*". It consists in buiding remote sensing and crops statistis databases to be used to predict crop yield at the provincial level in the main production region in Morocco (see Deliverable D.43.2).

The objectives of this deliverable are to :

- Collect 10 days SPOT-VEGETATION NDVI and DMP from 1999 to 2009 ;
- Define spatial extent of agricultural areas of Morocco based on GlobCover map ;
- Extract 10 days SPOT-VEGETATION NDVI and DMP from 1999 to 2009, corresponding to agricultural areas of Morocco ;
- Calculate mean NDVI and DMP values for agricultural areas of Morocco at national level and provincial level ;
- Collect wheat area and production statistics at provincial level ;
- Build database containing NDVI and DMP and wheat statistics.

3. NDVI AND DMP STATISTICS

3.1. Some backgrounds

3.1.1. SPOT-VEGETATION (VGT) & CTIV

The CTIV (Centre de Traitement d'Images Végétation), hosted at VITO, is responsible for the processing, archiving and distribution of all VGT-information. The two major products are:

- **VGT-P:** Individual registrations ("segments"), calibrated and geo-corrected, but without atmo-correction, hence they contain TOA-reflectances (top-of-atmosphere).
- **VGT-S10:** 10-daily global composites with TOC (top-of-canopy) reflectances and NDVI. The S10 are derived from all P-segments registered in the course of a given dekad (10 days). For each pixel, the "best available" observation is selected and atmo-corrected. The global S10 also comprises a Status Mask image (SM) which indicates the conditions of the selected observation: clear, cloud, snow/ice, error, etc...

The VGT data series started in April 1998 and has been continued until today without any interruption. All data are expressed in the WGS84-Lon/Lat system with a resolution of 1°/112 (i.e. about 1 km around a great circle).

3.1.2. JRC-MARS & VITO's Application Unit

Since 2000 VITO's "Application Unit" systematically delivers a number of VGT-products in near-real time to the MARS-project (Monitoring Agricultural ResourceS) of the EU-JRC (MARSOP-contracts). Below we only discuss some elements which are relevant for E-AGRI. On behalf of JRC-MARS the VGT-products of CTIV are adapted as follows:

- All images are converted to the ENVI-format (*.img with the metadata expressed in associated *.hdr files).
- The (byte) NDVI-images are rescaled (V=digital number):

  CTIV: NDVI = -0.10 + 0.04*V, with V=0-255
  MARS: NDVI = -0.08 + 0.04*V, with V=0-250

  The upper range (V=251-255) is used to label pixels with aberrant observations (251=error/missing, 252=cloud, 253=snow, 254=water, 255=other exception). This

information is extracted from the Status Mask and from the GLC2000 land cover map (for the distinction between land and sea).

- New images are derived with two biophysical parameters: fAPAR (fraction of absorbed PAR, 400-700 nm) and DMP (Dry Matter Productivity). DMP is computed with a Monteith approach from fAPAR and external meteorological information (daily solar radiation and T_{min}/T_{max}).
- In some cases, smoothed versions are computed for NDVI and fAPAR, using a modified version of the Swets-algorithm which inspects each pixel's time profile, detects all the cloudy observations and interpolates more appropriate values. In these cases, DMP is computed from the smoothed fAPAR (+ meteo).
- At the end, databases are derived with so-called RUM-values (Regional Unmixed Means).

Slightly different approaches are followed for the two actions of the JRC-MARS project:

- FoodSec (VGT): All the VGT-S10 are collected and archived on a global scale, although some dedicated regions of interest (ROIs) must be treated in more detail (e.g. IGAD, SADC, Mercosur,etc.). All images remain expressed in the original LonLat-system of CTIV. fAPAR is directly derived from the S10 using a "light" version of the famous CYCLOPES algorithm. Project partner MeteoConsult provides the daily meteodata requested for the DMP-computations. They are derived from ECMWF information and expressed in global grids with a resolution of 0.25°. For FoodSec, no smoothing is made of NDVI and fAPAR. RUM-databases are computed for all variables (NDVI, fAPAR, DMP) using the FAO-GAUL1 map (3363 administrative regions) and different crop masks, mainly derived from the GlobCover land use map.
- Agri4Cast (VGTp): This action only deals with the pan-European continent and all the imagery is expressed in the equal-area INSPIRE-LAEA projection with 1 km spatial resolution. In this case, the dekadal composites are derived via the alternative approach developed by JRC-IES (Nadine Gobron) which starts from the VGT-P segments. Each segment is corrected for BRDF-effects, atmo-corrected and then an fAPAR-layer is added. Afterwards, the S10-composite is derived using an fAPAR-MVC rule. NDVI is derived from the atmo-corrected reflectances in RED and NIR. To distinguish the results from the standard VGT-products of CTIV, we label them as "VGTp". In this case, smoothed versions are generated as well for NDVI, fAPAR and DMP. The daily meteodata are delivered by project partner Alterra via grids with a resolution of 25 km. Finally, RUM-databases are computed for all variables (non-smoothed and smoothed

NDVI, fAPAR, DMP) using the EU-NUTS3 map (1929 administrative regions) and different crop masks, mainly derived from CORINE/CLC2000.

3.1.3. Some Technical Remarks

Table 1 summarizes the involved image types. All images derived for JRC-MARS follow the fixed naming convention: ***spyyttv*.img/hdr**:

- *s* = sensor: s=v for VGT (FoodSec), s=u for VGTp (Agri4Cast)
- *p* = periodicity: in our case we always have p=t (10-daily)
- *yy* = year with for VGT yy=98,99,00,01,...,11
- *tt* =dekad in year: tt=01, 02,..., 36. NB: Every month is subdivided in three dekads, the first two always comprise 10 days (1-10, 11-20), while the third one has variable length (8/9 days for February, 10 or 11 days for the other months).
- *v* = suffix for the "variable" expressed by the concerned image (see **Table 1**).

For instance: vt9810i contains the flagged (but non-smoothed) NDVI from the standard VGT-S10 (FoodSec) for the first dekad of April 1998. And ut0001b is the smoothed fAPAR for the first dekad of 2000 as derived for AGri4cast (VGTp).

Table 1: *Image types (DT=Datatype: 1=unsigned byte with V=0→255, 2=signed short integer with V=-32768→+32767).*

v	CONTENTS	DT	SCALING	FLAGS
i	NDVI non-smoothed	1	NDVI [-] = -0.08 + 0.004•V (V=0-250)	251=error/missing, 252=cloud, 253=snow/ice, 254=water, 255=other exception
k	NDVI smoothed			
a	fAPAR non-smoothed	1	fAPAR [%] = 0.5•V (V=0-200)	
b	fAPAR smoothed			
p	DMP noon-smoothed	2	DMP [kgDM/ha/day] = 0.01•V (V=0-32767)	-1=error/missing, -2=cloud, -3=snow, -4=water, -5=other exception
y	DMP smoothed			

3.2. Ancillary data & study area

3.2.1. Global ancillary data

Three sources, all compatible with the global 1 km spatial resolution imagery of SPOT-VGT:

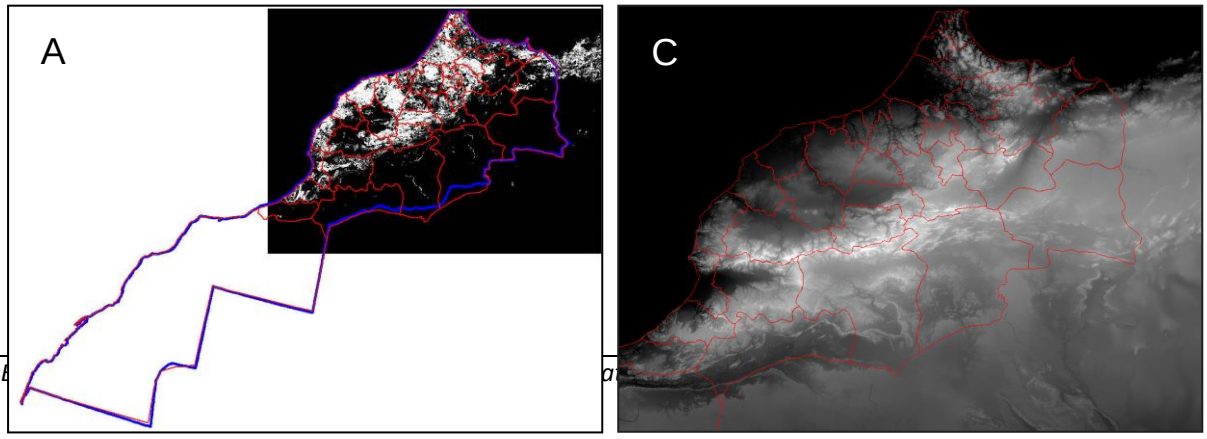
- The digital terrain model GTOPO30.
- The GLC2000 land use map (hard classification: each 1 km pixel is assigned to 1 class).
- Early 2011, JRC-FoodSec created a new global 0/1-cropmask at 250 meters spatial resolution. It was compiled from different sources: GlobCover V2.2, CORINE-2000, AfriCover, the SADC data set and the USGS Cropland Use Intensity data set. The mask covers exactly the same region as the global SPOT-VGT, but the resolution is four times higher, i.e. 1°/112/4 or roughly 250 meters along a great circle. The image is framed such that each VGT-pixel exactly covers 4x4 mask pixels. From this 0/1-mask we derived a 1 km spatial resolution "Area Fraction Image" (AFI) which indicates for each 1 km pixel the area fraction covered by cropland. This AFI is called GlCropV2 and is used as well for E-AGRI – at least as long as no better alternatives are available.

3.2.2. Study area

The VGT-data for the Moroccan study area was extracted from the European VGTp-S10 of Agri4Cast by VITO. Hence, they are expressed in the INSPIRE-LAEA projection system with a resolution of 1 km. As shown in **Figure 1**, the selected ROI only covers the northern parts of the country (see **Table 2** for provinces names and their Region_ID numbers).

It is bounded as follows:

- X: from 2 275 000 m to 3 385 000 m
- Y: from 765 000 m to 1 575 000 m
- Nr. of columns=1110, nr. of records=810, nr. of pixels=899 100.



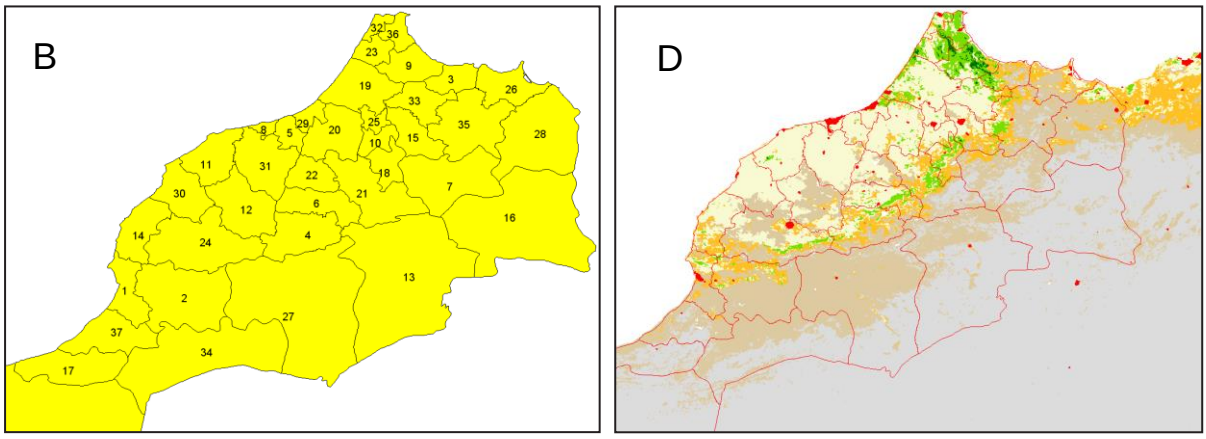


Figure 1: *Moroccan study area. Morocco (blue) and its provinces (red) and the analysed image block, in this case with the cropland AF (A). The 38 considered provinces and the Region-IDs (B). Digital terrain model GTOPO30 (C). The GLC2000 land use map. White=cropland in the northwest, green/brown=forests and shrubs mainly in the Atlas mountains, grey=desertic zones in the south and east, red=urban areas (D).*

Table 2: *The considered Moroccan provinces and their Region_ID numbers.*

ID	NAME	ID	NAME	ID	NAME	ID	NAME
1	AGADIR	11	EL JADID	21	KHENIFRA	31	SETTAT
2	TAROUDAN	12	EL KELAA	22	KHOURIBG	32	TANGIER
3	AL HOCEI	13	ERRACHID	23	LARACHE	33	TAOUNATE
4	AZILAL	14	ESSAOUIR	24	MARRAKEC	34	TATA
5	BEN SLIM	15	FES	25	MEKNES	35	TAZA
6	BENI MEL	16	FIGUIG	26	NADOR	36	TETOUAN
7	BOULEMAN	17	GUELMIM	27	OUARZAZA	37	TIZNIT
8	CASABLAN	18	FRANE	28	OUJDA	38	LAAYOUNE
9	CHEFCHAO	19	KENITRA	29	RABAT		
10	EL HAJEB	20	KHEMISSE	30	SAFI		

INRA-Morocco provided the following data sets:

- Official yield statistics (yields, areas, productions) for four crops (soft wheat, durum wheat, barley and autumn cereals), 34 years (1979→2012) and 38 provinces (see §4).
- SHP-files with the boundaries of the 38 concerned provinces.

NB: The 38 provinces are shown in **Figure 1** and **Table 2**. Note that we added the ID-numbers (1-38). The association between these Region-IDs and the Region names will be used throughout the analyses and it should not be lost.

INRA also provided relevant information on the normal wheat phenology (see **Figure 2**). In Morocco, wheat growth is mainly determined by the weather circumstances in the period October-March.

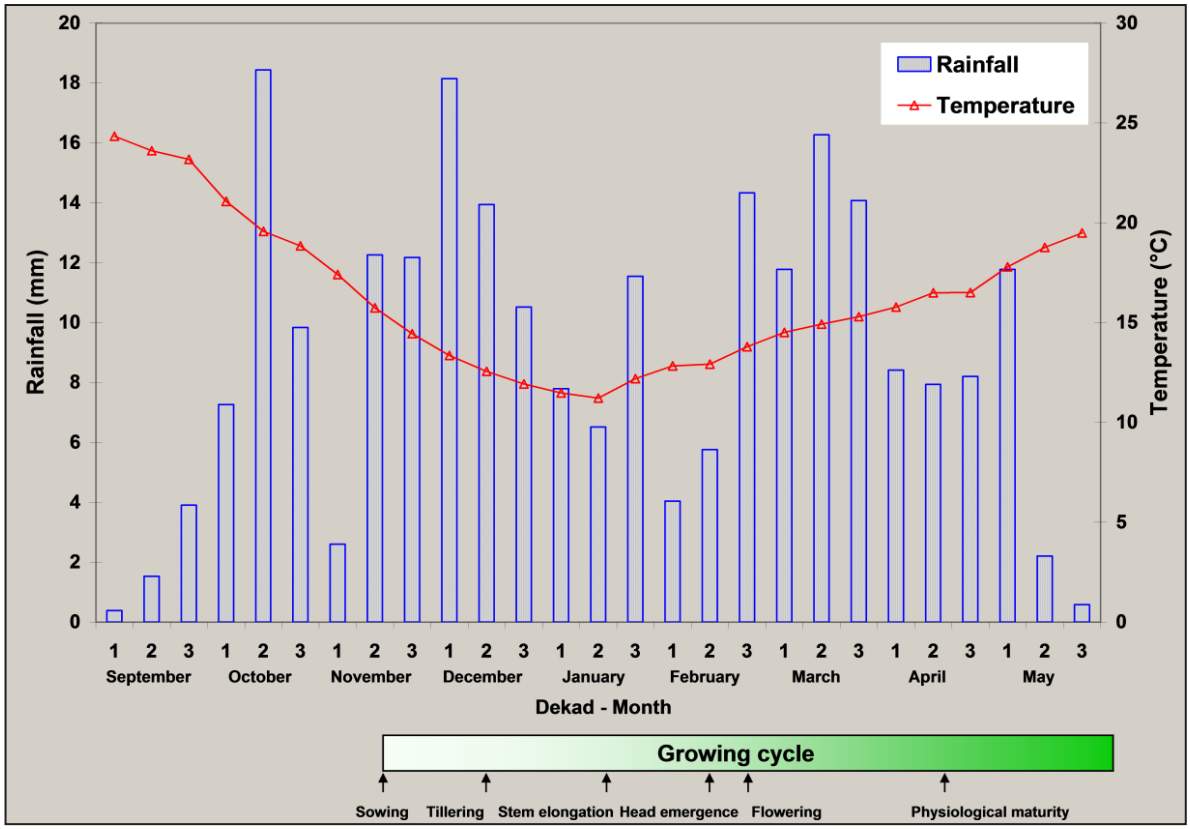


Figure 2: *Typical cereal cropping season in Morocco in terms of phenology, rainfall and temperature (figure provided by INRA-Morocco).*

3.3. Preparation & adaptation of the RS-data

The characteristics of the Moroccan ROI were described in §3.2.2. Again, we first extracted the considered ROI from the global ancillary images: GLC2000, GTOPO30 and the cropland AFI (GlCropV2). And the SHP-file with the province boundaries was converted to raster format. In the resulting Regions-IMG, each pixel contains the REG_ID number (1-38) of the province to which it belongs. See **Figure 1** and **Table 2**.

Morocco RS-imagery was extracted by VITO from our European VGTp-archive, which already contained the five target-variables: i-NDVI, a-FAPAR, k-NDVI, b-FAPAR and y-DMP. All the imagery covers the full period 1999-2012 (14 years, always with 36 dekads).

The RUM-extraction was done as followings:

- Separately for the five concerned variables: i-NDVI, a-FAPAR, k-NDVI, b-FAPAR, y-DMP (see **Table 1** – hence not for p-DMP).
- The mean value for all "cropland pixels" in each of the 38 provinces.
- As "cropland" were considered all pixels having an area fraction of 100% in the cropland AFI (see **Figure 1**).

The "dedicated program" was then used to convert initial RUM-databases into easier-to-handle format (see below in §3.4).

3.4. Final databases with RUM values

The final RUM-values are provided by VITO in three different forms and files (see **Figure 3**). This gives six files: X1.csv, X2.csv, X3.csv, with X= Maroc. All are ASCII-TXT and work with "comma-separated values" (CSV) which facilitates the import in EXCEL of other spreadsheets.

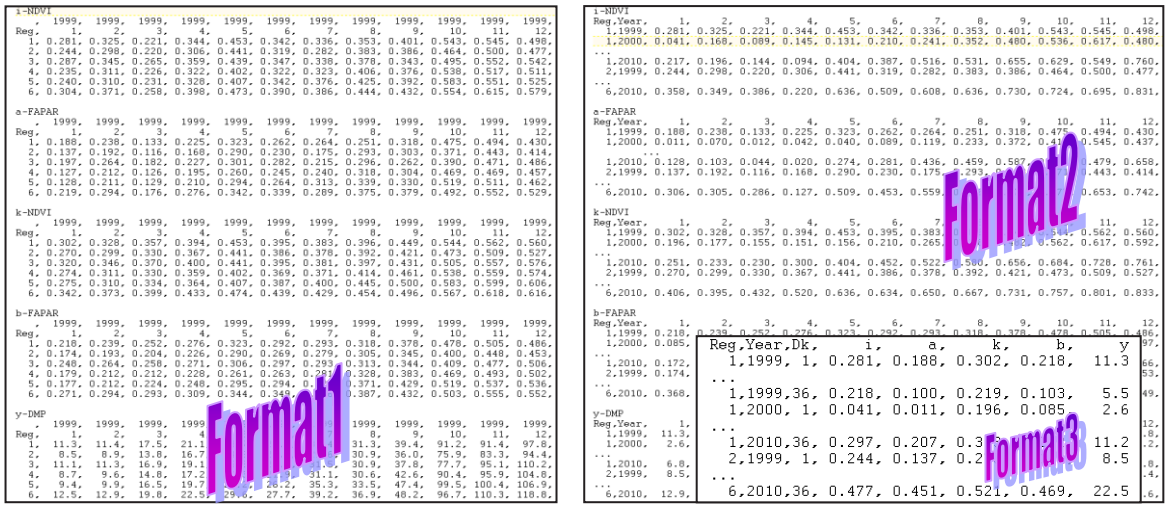


Figure 3: *Example of the three different formats for the final tables with RUM-values.*

Format 1:

- Five tables, one for each variable (i, a, k, b, y).
- Per table: one line per region. It starts with Region_ID and then contains the values for all subsequent dekads, in all years.

Format 2:

- Five tables, one for each variable (i, a, k, b, y).
- Per table: one line per region x year.
- Each line starts with the concerned Region_ID and year. Then follow the values for all the 36 dekads.

Format 3:

- One single table.
- Separate lines for each combination of region x year x dekad
- Each line starts with Region_ID, year and dekad-number. Then follow the corresponding values of the five target variables: i-NDVI, a-FAPAR, k-NDVI, b-FAPAR, y-DMP.

Remarks on missing values:

- All negative values must be considered as missing values.
- They only occur in Morocco, especially for the provinces of Figuig (Region_ID=16), Tata (34) and Laayoune (38).

3.5. DATA SETS DELIVERED

The following datasets were delivered.

5. SHP-files with the boundaries of the country and its 38 provinces converted to the INSPIRE-LAEA projection (compatible with the VGTp-imagery). The associated DBF-file also includes the Region_ID numbers (see **Figure 1** and **Table 2**).
6. Files MAROC*n*.csv, with *n*=1/2/3, containing the RUM-values in the three different formats (see §3.2).
7. Official statistics: File "cereales 1979-2012_Bis.xls" was provided by INRA. The TXT-files were derived from it. They contain one line for each Region x Year, and each such line comprises three comma-separated values: Region_ID (table 2), Year and official yield (in qx/ha). Negative values are missing. There are four such files for each of the crops: BLE_DUR (durum wheat), BLE_TEN (soft wheat), ORGE (barley) and CER_AUT (total of soft wheat, durum wheat and barley).

Reg_ID	Year	Dek	i-NDVI	a-fAPAR	k-NDVI	b-fAPAR	y-DMP
1	1999	1	0.214	9.820	0.229	10.880	13.9
1	1999	2	0.231	11.690	0.242	11.820	14.2
1	1999	3	0.257	13.210	0.263	13.310	20.2
1	1999	4	0.266	14.130	0.279	14.570	22.2
1	1999	5	0.282	15.370	0.288	15.720	27.9
...	...	...	...	...	...	...	...
1	2010	36	0.277	16.520	0.298	18.100	35.0
37	1999	1	0.177	7.060	0.194	8.560	10.5
...	...	...	...	...	...	...	...
37	2010	36	0.341	19.970	0.344	19.910	21.5

Reg_ID	Year	Mean k-NDVI	Barley Yield

		Feb-Mar	Qx/ha
1	1999	0.300	4.1
2	1999	0.267	8.5
3	1999	0.308	8.6
...	...	...	...
37	1999	0.564	4.2
1	2000	0.222	7.1
...	...	...	...
37	2009	0.608	12.0
...	...	...	...
36	2010	0.526	17.7
37	2010	0.439	10.2

VARIABLE	ORIGINAL	SMOOTHED
NDVI	i-NDVI	k-NDVI
fAPAR	a-fAPAR	b-fAPAR
DMP		y-DMP

SHIFT	NONE		± 3 DEKADS	
CROP	**FEB-MARCH**	**FEB-APRIL**	FEB-MARCH	FEB-APRIL
Durum wheat	68	76	70	82
Soft wheat	68	76	67	78
Barley	61	81	59	72
Autumn cereals	72	84	65	80

4. WHEAT STATISTICS

Area and yield data of the three main cereal crops[1], soft wheat, durum wheat and barley, were graciously provided by "la Direction de la Stratégie et des Statistiques" (DSS) (Balaghi et al., 2013). They are available for 40 provinces of the country for the period of 1978-2012. Production at a province level is obtained by multiplying the yield value with the area estimated by DSS.

Area estimation for cereal crops in Morocco is made every year by DSS between February 10th and March 30th, using a sampling method of 3,000 unit areas representing 19 million hectares (see Deliverable D54.2). Starting since 2008, DSS has renewed the sampling procedure to integrate modern techniques of satellite remote sensing and GIS which improved precision of estimators. A GIS application has been specifically developed for this purpose with capability of automated steps of the sampling procedure.

Within the sampled areas, sub-sampled plots are harvested and their yields directly measured. Production of a sampled area is the product of measured yield on sub-sampled plots and area represented by the sample. Data on production and area is then aggregated by province.

Monitoring of vegetation, area and yield estimations are carried out by DSS along the cereal crop cycle in three phases:

- **Phase 1:** A survey on the evolution of harvest, is done in February, to evaluate crop growth stages and vegetative stand of crops ;
- **Phase 2:** A survey on land cover, done between February and June, to estimate cereal areas ;
- **Phase 3:** A survey on expected production, done in April (1 to 2 months before harvest) to estimate production of the three main cereals: soft wheat, durum wheat and barley.

Official agricultural statistics are delivered in Excel format, at province administrative level; which makes the province, the smallest territorial area for forecasting. Polygons delimiting provinces are available in a GIS format.

[1] These cereals are sown in the fall (autumn) and are sometimes called autumnal cereals as compared to spring cereals like corn or sorghum.