Quantifying spatial and temporal patterns of the water related yield gap, using synoptic data and a dynamic Crop Growth Model (PSn) - A Case Study of Sunflower in Andalucia, Spain

By Tran Van Tuan

I. Introduction

Various approaches and methods have been developed to research on yield gap constraints to improve crop production and assist to gain food security. One of attractive methods is called "Comparative Performance Analysis" (CPA)(De Bie, 2000). The principle of this method is to identify and quantify all factors of land and management that cause a gap between actual and potential yield of a certain crop in an area. In this method, factors as shown in table 1 can be identified and their impact also can be quantified.

Recently, some researchers have used the data from remote sensing (temperature, rainfall, NDVI) for estimating the crop production for food security purposes (Metternicht, 2003; Rugege, Bouma, Skidmore, & Driessen, 2002; Venus, 2000). These researches have been found to be very useful for crop production estimation and contributed to improve the crop growth simulation techniques as well as crop modelling methods. Although it has been very practical of using those RS data for estimating the crop production, these spatial factors (derived from satellite images or from weather station based data) have not been used as based attributes for mapping the yield gap and improving the methodology of yield gap analysis.

Beside the CPA, the "Production Situation Model" (PS-n) can be used with weather based data (temperature, precipitation, relative air humidity and ETo etc) to estimate crop potential and water limited yields in study areas (points). The principle of this model presents a simplified Land Use System in which crop production and yield are solely determined by availability of light, the temperature, the photosynthetic mechanism and other factors.

The main idea of this research is to combine the use of Production Situation Model (with data from weather stations), remote sensing data and techniques to quantify the water related yield gaps of Sunflower in space and time following CPA logic.

II. Methods and Materials

2.1. Study Area

Andalucia is an autonomy region located in the south of Spain. It is composed of 8 provinces including: Sevilla, Malaga, Almeria, Cadiz, Cordoba, Granada, Huelva and Jaen, stretching from Southeast to Southwest of the country. The total natural area is of 87,300 sq km (17.3% of Spanish territory) with total population estimates of 7 million people (20% of Spanish total population).

Andalucia is the home of olive and sunflower cultivation in Spain. Besides that, the main agricultural crops of this region include wheat, barley, fruit, grape and vegetable. Due to the influence of specific climate condition (hot in summer and cold in winter), Andalucia has produced very good products and olive (fruit and oil) and grape wine meanwhile it provides most of sunflower oil in EU. Figure 8 presents the agricultural area in Andalucia which is home of grape, olive, wheat and sunflower.

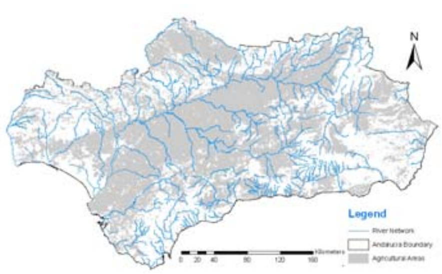


Figure 1: Agricultural map of Andalucia Region

2.2. Field data collection

2.2.1. Sampling scheme

In order to make a random sampling scheme, a crop map of sunflower for the study area was developed. This map was made through image classification of SPOT-VGT NDVI images in combination with crop statistic data and Corine (land cover map) with GIS analysis.

SPOT-NDVI images of Andalucia region with all 300 layers from 1998 to 2006 were classified to 45 classes by unsupervised classification method (the number of classes were identified and decided through divergence analysis). The derived raster map was converted to polygon map and used the mask of agricultural area from the land cover (corine) map of 2000 to select the area that present the crop of sunflower. This map then was intersected with municipality map to identify the area of each NDVI class in each municipality. The attributable of the result was joined with agricultural statistic data of sunflower in Andalucia for multi variable regression analysis in SPSS. A step-wise linear regression model was developed as:

Y = a1*N1 + a2*N2 + … + a45*N45

Where:

Y is the area of sunflower in each municipality (from statistic data)

N1 to N45 is the area of each NDVI class present in each municipality

And a1 to a45 are the coefficients denoting the fraction of sunflower in each NDVI class (0 to 1).

Based on the result of this model, a crop map of sunflower was developed to present the faction of sunflower by every 1 square km in Andalucia (figure 2).

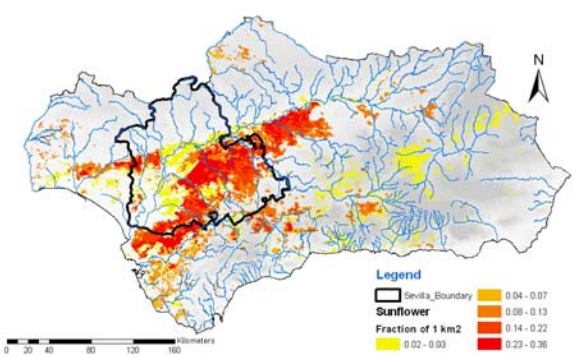
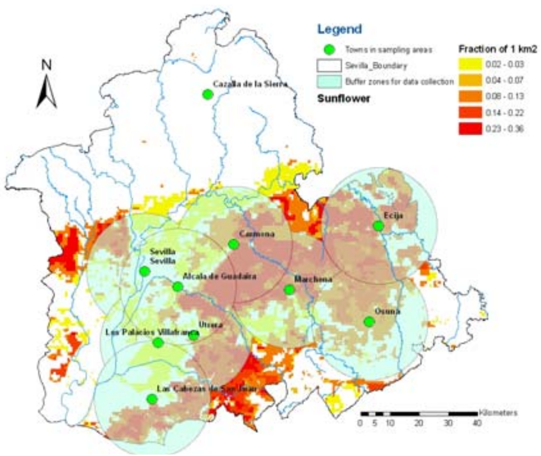


Figure 2: Crop map of rain-fed sunflower in Andalucia

From the crop map of sunflower, it is very clear to observe that most of sunflower is planted in Sevilla province. Based on the result of the crop map and based on the reality that the time is limited for the data collection in the field, Sevilla province has been chosen for sampling schemes of data collection instead of whole region of Andalucia.

Based on the crop map of Sevilla province, the method of strata sampling scheme was applied. The buffer zones were created around the main towns with a distance of 20 km from every town. These buffer zones were chosen to conduct the interview with farmers for data collection (figure 3). Due to high fraction of sunflower in these areas, the interviews have been conducted randomly in any farm of the strata.



2.2.2. Field data collection

Most of field data were obtained by two traditional methods: observation and through interview. The observation was used to collect data that related to land at plot size such as soil quality and current operations on the field. Meanwhile, the interview (by semi-structured method) was applied to obtain the data related to land use operations (management) such as land preparation, soil treatment, planting time, fertilizer application, water shortage and yield etc (the format of interview and observation sheet is presented in the appendices). All the data from interview were used to discuss about the variable of land and land use operation and to develop the assumption for the research. In this fieldwork, statistical data of sunflower actual production in Andalucia (from Ministry of Agriculture, Spain) from 2001 to 2005 were also collected. In this data, the actual production of a certain crop (sunflower, wheat etc) is recorded for different random selected strata of 700 m x 700 m in Andalucia.

Additional data for this research are data from weather stations in Sevilla province. The data were collected are air temperature (daily maximum and minimum), precipitation, Eo, ETo, Air humidity etc. These data later are used as the inputs for PS-n model to estimate PS1 and PS2 yields for all weather station based points. In addition, satellite imageries concerning products of 8-day Gross Primary Productivity (GPP) and daily Land Surface Temperature (LST) in Andalucia region were also downloaded from MODIS (NASA – EOS Gateway - http://delenn.gsfc.nasa.gov/ims-bin/pub/secured/nph-ims.cgi/u39543).

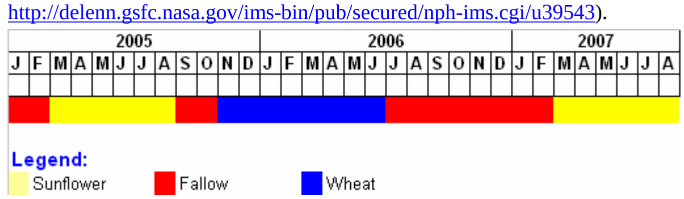


Figure 4: Crop calendar for sunflower in Andalucia (in 23 sites)

All interviewed farmers (23 farms) have reported the same sequence for their crops as a crop rotation in flowing order: Sunflower – Wheat – Sunflower (figure 4). This sequence (rotation) was applied in order to avoid crop diseases (fungus, bacteria etc) and to take the advantage of nutrition from previous crop to support later crops.

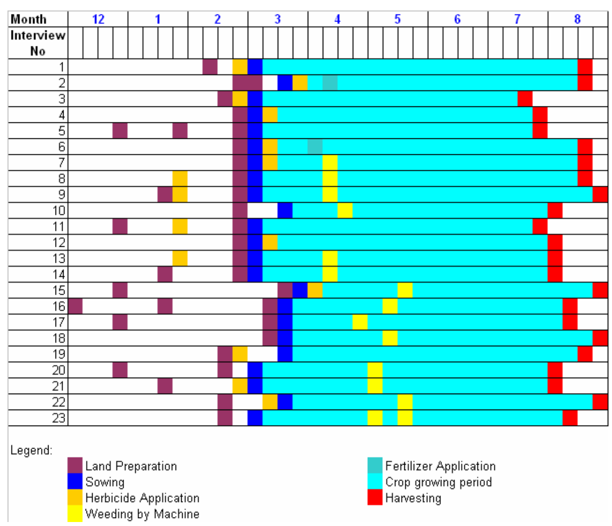


Figure 5: Operation sequence for sunflower cultivation in interviewed farms, Andalucia

In terms of land operational sequences, figure 5 shows that 65% of interviewed farmers sowed the seeds of sunflower at the same time (first week of March), meanwhile, the remaining (35%) sowed one or two weeks after.

2.3. Research method

2.3.1. Research flowchart

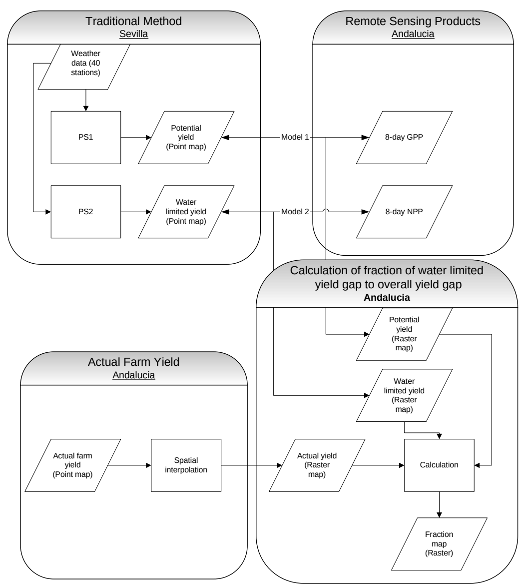


Figure 3: Overview of research method

2.3.2. Software used

The following software and programs are used for data analysis in this research:

- SPSS – version 15.0: for statistic analysis
- ArcGIS – version 9.2: for GIS analysis
- Erdas Imagine – version 9.1: for remote sensing/image processing and analysis
- ArcPad for Windows – version 7.1: for GPS support
- Microsoft Office 2003
- Model selected: Production Situation Model (PSn) (Driessen & Konijn, 1992)

III. Results and Discussion

3.1. Land management for sunflower cultivation in Andalucia

As the traditional CPA method (De Bie, 2000), land management factors were identified and analysed to quantify the degree of influence to final yield gap of sunflower. The interview results have shown that, the land managements including crop calendar and operational sequence are almost homogeneous in interviewed plots.

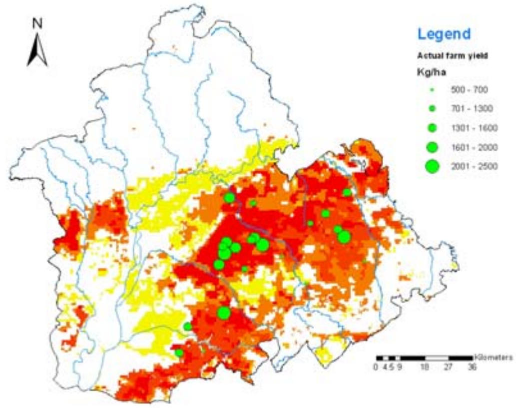


Figure 6: Actual farm yield of sunflower in interviewed farms, year 2007

Figure 6 also shows that the actual farm yields (collected by interviews) are variable from plot to plot. This variation in space can be explained by other factors that may play stronger roles in determining the yield gap of sunflower in Andalucia.

From all analysis, it is possible to come to an assumption that the land management for sunflower cultivation in Andalucia is almost homogeneous in space and time and it does not deicide the yield gap of sunflower in space and time.

3.2. The Heterogeneous of Rainfall Data Used for PSn model

Beside temperature, the rainfall is a very important factor for crop development. Especially, in non-irrigated areas, the rainfall is main source of water to provide moisture for crop's growth. Insufficiency of water in different period of time may provide different effects to the crop and through those effects, the final yields of crop are determined.

In space and time, the rainfall is also a factor that determines the variation of crop yield from plot to plot and year to year. Different amount of rainfall in different plot brings variable yield data that farmer could not modify except through the irrigation intervention.

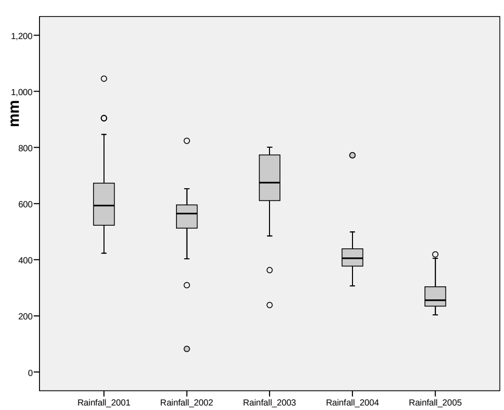


Figure 7: The box-plot shows the distribution of annual rainfall data of 40 weather stations in Sevilla province (2001 to 2005)

3.3. Yield reduction of sunflower due to water shortage

3.3.1. Potential yield of sunflower calculated by PS1 model

Production Situation Model (PS1) was used to calculate the potential yield of sunflower in 40 weather station points in Sevilla province. These yield values were defined by climate factors of solar radiation and temperature (Driessen & Konijn, 1992). Based on the variation of climate in space and time, the potential yields are also variable from plot to plot and from year to year.

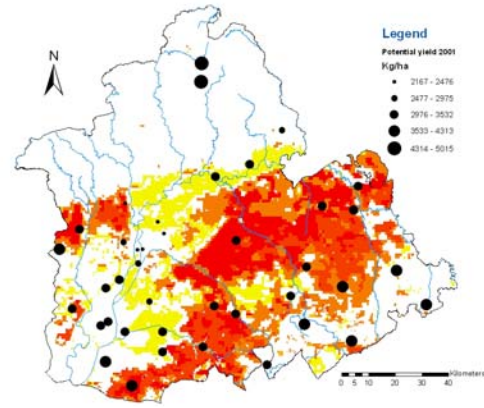


Figure 8: Point map of potential yield of sunflower in Andalucia, year 2001

To extrapolate the point map of potential yield (40 points) to raster map of whole sunflower cultivation in Andalucia, the step-wise multiple regression has been conducted to quantify the correlation of potential yield value and GPP values (figure 9). The regression analysis (linear, step-wise) results show that there is a significant correlation between potential yield of sunflower and GPP values through many 8-day periods (16 periods):

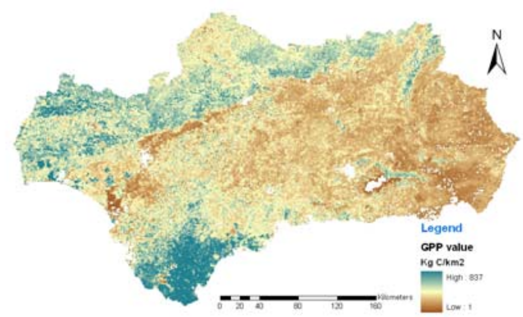


Figure 9: One of GPP images from MODIS, used for potential yield extrapolation (year 2001)

Based on the regression analysis for potential yields and GPPs values of other years, the equation for extrapolation of potential yields for each year was also developed and presented in table 1.

Table 1: Equations for potential yield extrapolation for other years, derived from regression analysis results

Year	Equation for extrapolation	R Square
2001	Potential yield_2001 = 7.401*GPP81 + 11.989*GPP121 + 5.531*GPP177	0.948
2002	Potential yield_2002 = 18.187*GPP113 + 7.855*GPP177	0.941
2003	Potential yield_2003 = 13.19*GPP129 + 12.49*GPP169	0.895
2004	Potential yield_2004 = 12.393*GPP97 + 66.1*GPP177	0.940
2005	Potential yield_2005 = 17.692*GPP89 + 58.544*GPP169	0.851

Table 1 shows that there was a significant correlation between potential yields and GPP values of every year. However, it is not possible to use the coefficients and GPP periods of one year to apply for extrapolation of other years.

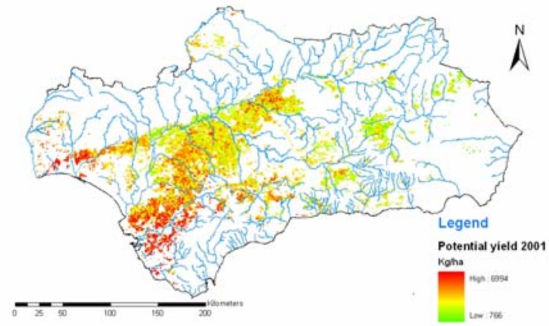


Figure 10: Raster map of potential yield of sunflower in Andalucia, year 2001

In Andalucia, for the year of 2001, the figure 10 shows that the potential yields of sunflower could vary from 463 to 6,995 kg/ha. However, in the high fraction areas for sunflower cultivation, the potential yield could gain from 2,744 to 5,253 kg/ha (based on classified map – shown in Appendices). The variation of potential yields in space will partly contribute to the variation of actual farm yield in reality.

3.3.2. Water limited yield of sunflower

Water limited yield is the second level of potential yield calculation from PSn model. Depending on the availability of water for crop growth though out the relative development stages (RDS), potential yields will be reduced from 0% up to almost 100%. If water was enough for crop growing in all RDS, then PS2 yield = PS1 yield. However, in most of rain-fed crop growing areas, water is always a limited factor and in some periods of RDS, the crop always faced with water stress. In this case, PS2 yield = PS1 yield – yield reduction due to water stress (shortage).

The results of PS2 calculation for water limited yield of sunflower in 40 points of Sevilla province is showing in figure 11. It shows that as well as potential yield, the water limited yield is also variable from point to point within a year. It ranges from 484 to 3167 kg/ha in 40 points in Andalucia for the year 2001. The water limited yields were also calculates for 40 points of year 2002, and year 2005. The PS2 model was failed to calculate water limited yields for year 2003 and year 2004 due to excessive moisture values which provided by weather data.

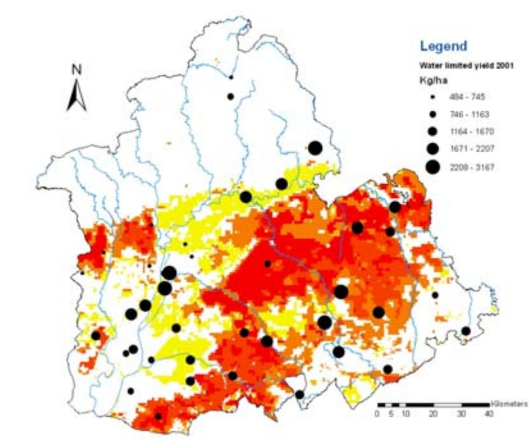


Figure 11: Point map of water limited yield of sunflower in Andalucia, year 2001

To extrapolate the water limited yield from the map of 40 points to raster map for whole sunflower area in Andalucia, the Net Primary Productivity (NPP) from remote sensing was calculated from GPP by following function:

NPP = GPP*cfH2O ***where:*** cfH2O = ETa/ETo

From the analysis results, the cfH2O for every 8-day period was calculated from secondary data of 8-day ETa and 8-day ETo. The image of calculated cfH2O of 8-day period (from day 81 to day 88, year 2001) is presented in figure 12.

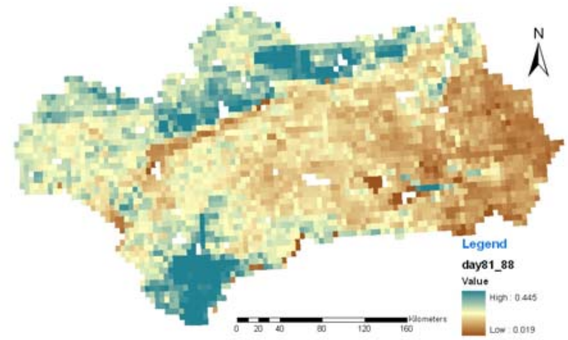


Figure 12: cfH2O calculated from ETa and ETo for 8-day period, year 2001

This value (cfH2O) was used to calculate NPP from GPP image. The image of NPP from day 81 to day 88, year 2001 is shown in figure 13.

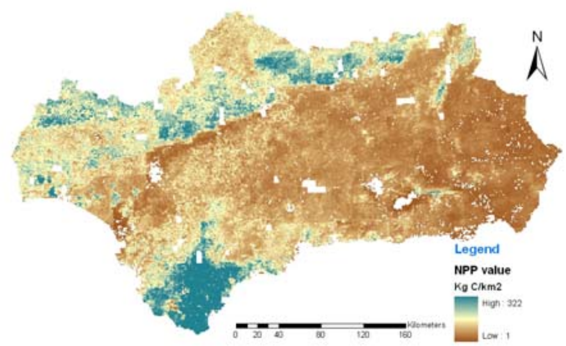


Figure 13: One of 8-day NPP calculated from GPP, year 2001 in Andalucia

The correlation between water limited yield of sunflower and NPP values through 16 eight day period is presented by statistical analysis (linear multiple regression).

Water limited yield (2001) = 29.361*NPP81 + 75.486*NPP153

Apply this equation into the NPP images, the raster map of water limited yield of sunflower year 2001 is produced and presented in figure 14.

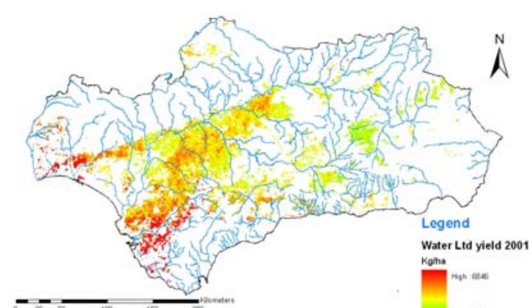


Figure 14: Raster map of water limited yield of sunflower in Andalucia, 2001

Figure 14 shows that the water limited yield of sunflower in Andalucia, year 2001 was distributed from 168 to 6,864 kg/ha. However, in the high fraction areas of sunflower, the water limited yield was mostly in the range of 1000 to 2500 kg/ha.

3.3.3. Yield reduction of sunflower due to water shortage

The result of raster calculation (potential yield minus water limited yield) for every year shows the patterns of water limited yield gap. A raster map of these patterns for year 2001 is presented in figure 15.

This figure shows that the patterns of yield gap caused by water shortage was very variable in space for year 2001(from 32 to 4,574 kg/ha). This gap varied from 2 to 5,599 kg/ha for year 2002 and from 0 to 4,980 kg/ha for year 2005 in Andalucia. Comparing with the crop map of sunflower, in the high sunflower fraction areas, the water limited yield gaps are smaller than other low sunflower fraction areas (classified yield gap map – shown in Appendices). The possibility for sunflower to be cultivated more in high fraction areas may come from the suitability of climate (in terms of water availability) in these areas.

It is also clear to observe that in many areas, there are big numbers of yield reductions due to the water shortage. This result partly contributes significant evidences to prove the hypothesis that the water is the main cause of yield gap for sunflower in Andalucia. The water limited yield reduction map for every year is the result of quantifying the spatial patterns of yield gap caused by water shortage (where and how many kg/ha). Comparing with actual farm yields will help to quantify the fraction of this gap (how many percent in overall yield gap).

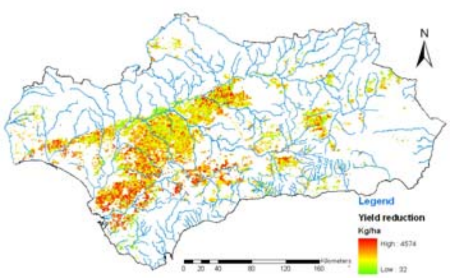


Figure 15: Yield reduction of sunflower in Andalucia due to water shortage, year 2001

3.4. Quantification of overall yield gap of sunflower

3.4.1. Actual farm yield of sunflower

To quantify the overall yield gap of sunflower for every year, the actual farm yield of sunflower is used for following calculation:

Overall yield gap = Potential yield – Actual farm yield

The point data of sunflower actual farm yield for every year (2001 – 2005) were derived from the agricultural yield maps (from Ministry of Agriculture, Spain) and shown in figure 40.

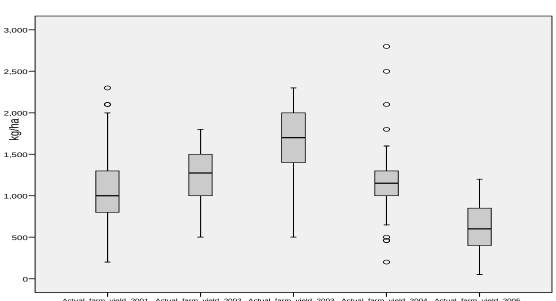


Figure 16: The distribution of actual farm yield of sunflower in Andalucia from 2001 to 2005

Figure 16 shows that the actual farm yield of sunflower in Andalucia is mainly in the range of 500 to 1,500 kg/ha (except year 2003). The actual farm yield is lower than potential yield due to the presence of 2 main yield gaps:

First Gap is caused by the shortage of water (water stress) during growing period.

Second Gap is caused by other land and management factors.

And **Overall gap = First gap + Second gap**

3.4.2. Extrapolation of actual yield gap of sunflower

Spatial interpolation method (Spline tool in ArcGIS) was used to extrapolate the actual farm yields from point map to raster map for each year. For quality control, all the areas with the distance is bigger than 30 km to the actual farm yield points, have been removed from the output raster map. The raster map of actual farm yield for year 2001 is developed and shown in figure 17.

The map shows that in 2001, the actual farm yield was ranging from 21 to 2253 kg/ha. However, in most of the areas, the yield was lower than 1770 kg/ha (classified yield map).

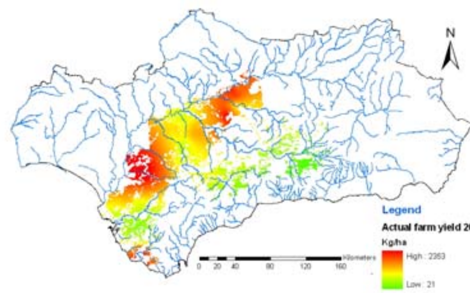


Figure 17: Raster map of actual farm yield of sunflower in Andalucia, 2001

3.4.3. Quantification of overall yield gap of sunflower

The patterns of overall yield gap for each year were quantified by following equation:

Overall yield gap = Potential yield – Actual farm yield

One of the resulted maps, the overall yield gap patterns (total yield reduction) for year 2001 was developed and presented in figure 18.

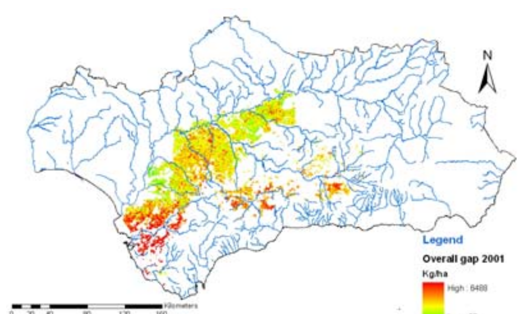


Figure 18: Map of sunflower overall yield gap patterns for year 2001 in Andalucia

In 2001, the yield of sunflower was reduced from 50 to 6488 kg/ha. In most of sunflower areas, the yield reduction by all limiting factors took place from 1300 to 3600 kg/ha (by natural breaks). However, it does not represent the highest yield gap for sunflower in Andalucia for 5-year period. A plot that had a low yield gap in year 2001 may not have the low yield gap in 2002 and vice versa. In order to quantify the patterns of highest overall yield gap in Andalucia for 5-year period, the available results of overall yield gap maps for 3 years (2001, 2002 and 2005) were used to calculate the maximum overall yield gap of each year. The maximum value (yield gap data) of each pixel was selected to produce the map of highest overall yield gap patterns throughout 5 years.

Patterns of highest overall yield gap = Maximum of (overall yield gap_2001, overall yield gap_2002, overall yield gap_2005).

The result of this calculation (map) was then classified into 2 classes of overall yield gaps (low gap and high gap) based on mean value of yield gap data (2,885). The classified map of overall yield gaps of sunflower for 5-year period is shown in figure 19.

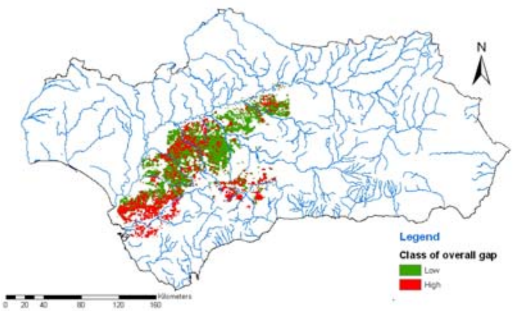


Figure 19: Map of classified maximum overall yield gap for 5 year period in Andalucia

Figure 19 shows that in space (Andalucia) and in time (throughout 5 years), the yield gap of sunflower could be quantified and classified into 2 main classes - the low gap patterns (smaller than 2,391 kg/ha) and the high gap patterns (bigger than 2,885 kg/ha) in Andalucia.

IV. Conclusion and Recommendation

1.1. Conclusions

- ✓ The Production Situation Model can be used to identify the variability of potential yield of sunflower in Andalucia during five years (2001 – 2005). The lowest potential yield of sunflower was 2167 kg/ha in 2001 with average potential yield for five years was 3389 kg/ha. For water limited yield during five years, the lowest yield was in 2005 with 263 kg/ha and the average yield was 1630 kg/ha.
- ✓ The 8 daily product of Gross Primary Productivity (GPP) which cover the space and time of sunflower growth can be used to extrapolate the potential yield of sunflower from point data to raster with the resolution of 1 km^2 a pixel. The regression analysis shows that there was a strong correlation of potential yield and GPP values calculated in the time the sunflower grew. However, this correlation was different from year to year.
- ✓ The Water Sufficiency Coefficient (cfH2O) aggregated for 8-day period can be used to calculate the 8-day basic water limited yield from potential yield instead of using daily data. The results of 2 calculation methods are not much different from each other. Meanwhile the monthly aggregated data of cfH2O is not significant for calculating the monthly water limited yield. This result also shows that, for sunflower, the impact of water stress (drought) on its daily productivity is not much different from 8-daily productivity but this impact is very significant between each month period.
- ✓ During 5 years, 68.99% of sunflower areas in Andalucia had the yield reduction by water shortage from 1,500 to 2,500 kg/ha. There was only 5.07% of the areas that had the yield reduction less than 1,500 kg/ha and 25.94% of the area where the yield reduction was in the range of 2,500 to 4,815 kg/ha. The objective of this research has been achieved by having the final maps of water limited yield gap which have quantified the patterns of water limited yield reduction in space (the variability of yield gap from plot to plot) and in time (the variability of this gap during 5 years).
- ✓ The water shortage (drought) was a main limiting factor causing yield gap of sunflower in Andalucia. This factor contributed more than 60% to the overall yield gap in most area of sunflower cultivation in Andalucia for every year. It is a significant result of this research to consider the intervention for narrowing the yield gap of sunflower in Andalucia by irrigation supply. It is also proves that the weather based estimation data from Production Situation Model together with remote sensing data can be used to quantify a component of yield gap of sunflower (water limited yield gap). By this result, the alternative hypothesis of this research has been accepted.
- ✓ Beside the water shortage, there are other factors causing yield gap for sunflower in Andalucia. These factors also would be variable in space and time. However, in most area, they contributed less than 40% of overall yield gap of sunflower in Andalucia.

1.2. Recommendations

- ✓ The water shortage situation (drought) in Andalucia contributed more than 60% to the overall gap of sunflower (in most area). Another gap component (30 to 40%) could come from other constraints or limiting factors like soil or land management. Supplying enough water for sunflower cultivation should be taken into account in order to improve the productivity of this crop in Andalucia.
- ✓ The maps of quantified patterns of water limited yield gap and overall yield gap of sunflower could be a reference for farmers or decision makers in implementing intervention or making plan to improve the agricultural production in the area.
- ✓ Other studies should be conduct to identify the constraints or other factors that caused the second yield gap of sunflower. Soil and management factors should be taken into account for further researches to quantify the influence of these factors on the variability of sunflower yield gap.