

# EVALUATING THE PERFORMANCE OF A CROP WATER BALANCE MODEL IN ESTIMATING REGIONAL CROP PRODUCTION

**Gabriel B. Senay**

Raytheon, EROS Data Center, Sioux Falls, SD, 57198  
senay@usgs.gov

**James Verdin**

USGS, EROS Data Center, Sioux Falls, SD, 57198  
verdin@usgs.gov

## ABSTRACT

The comparison between a spatially distributed crop index and reported yield was evaluated. The crop index was generated by a crop water balance model that simulates the reduction of crop yield (as a percentage of the potential) due to water deficit. This model is currently operational as a monitoring and forecasting tool for region-wide food security analyses in drought prone countries in Sub-Saharan Africa.

The most important input parameters of the model are Africa-wide satellite-based rainfall estimate images and spatially distributed potential evapotranspiration images calculated from 6-hourly numerical weather model data. District level crop yield data from Ethiopia were used to evaluate the performance of the model. Historical sorghum yield data from 1996-1999 were used to evaluate the performance of a seasonal water requirement satisfaction index (WRSI) for sorghum. WRSI values and reported district yield data were significantly correlated. The WRSI model was particularly successful in capturing the response of the crop during a relatively dry year. In districts that never experienced water deficit during the study period, it was possible to infer the magnitude of yield variability that was caused by factors other than water supply.

## INTRODUCTION

Monitoring the establishment of the crop growing season and the subsequent performance of crops is an important process in the assessment of regional food security conditions in Africa. The Famine Early Warning System Network (FEWS NET) uses data from two operational remote sensing products to monitor agricultural areas for signs of drought on a near-real time, spatially continuous basis. These include dekadal (10-day) AVHRR-Normalized Difference Vegetation Index (NDVI) images, produced by the National Aeronautic and Space Administration (NASA), and Rainfall Estimate (RFE) images prepared by the Climate Prediction Center of the National Oceanic and Atmospheric Administration (NOAA). One of the derivative products currently in use by FEWS NET is the Water Requirement Satisfaction Index (WRSI).

Despite the general agreement between the seasonal WRSI updates and field observations, a comprehensive evaluation of WRSI and reported crop yield has not been made in many countries in Africa where the WRSI image product is being distributed as part of an operational crop monitoring tool. The primary objective of this study was to examine the correspondence between WRSI values and reported sorghum yield. Four years of "woreda" administrative level (with an average district unit of about 1,600 km<sup>2</sup>) sorghum production data, collected by the Ministry of Agriculture, Ethiopia, were used for this study.

## METHODS

### Model Description

The water requirement satisfaction index (WRSI) is an indicator of crop performance based on the availability of water to the crop during a growing season. Studies by Food and Agricultural Organization (FAO) (Doorenbos and Pruitt, 1977)

## ON THE NEED FOR DYNAMIC SCHEDULING OF IMAGING SATELLITES

have shown that WRSI can be related to crop production using a linear-yield reduction function specific to a crop. More recently, Verdin and Klaver (2002) demonstrated a regional implementation of WRSI in a grid-cell based modeling environment on southern Africa.

WRSI for a growing season is calculated as the ratio of seasonal actual evapotranspiration (AET) to the seasonal crop water requirement (WR). The water requirement of the crop at a given time of the growing season is calculated by multiplying the reference (Potential) evapotranspiration with a crop coefficient, whose values are published by FAO (FAO, 1998). AET represents the actual amount of water withdrawn from the soil water reservoir ("bucket") where shortfall relative to potential evapotranspiration (PET) is calculated by a function that takes into consideration the amount of soil water in the "bucket".

The soil water content is obtained through a simple mass balance equation where the amount of soil water is monitored in a volume defined by the water holding capacity (WHC) of the soil and the crop root depth.

That is,

$$SW_i = SW_{i-1} + PPT_i - AET_i$$

Where, SW is soil water, PPT is precipitation, and i is the time step index. Currently, the model is run in dekadal time step; however, the model can also run in a daily time step.

The most important inputs to the model are precipitation and potential evapotranspiration (PET). FEWS NET at the USGS calculates daily PET values for Africa at 1.0-degree resolution from 6-hourly numerical meteorological model output using the Penman-Monteith equation (Shuttleworth, 1992; Verdin and Klaver, 2002). Blended satellite-gauge rainfall estimate (RFE) images for the African continent are obtained from NOAA at 0.1-degree (~10 km) spatial resolution. Rainfall images are produced using an interpolation method that combines data from Meteosat cold cloud duration (CCD), the Special Sensor Microwave/Imager (SSM/I) of the Defense Meteorological Satellite Program, the Advanced Microwave Sounding Unit (AMSU) on board the NOAA-15 polar orbiter, and reporting rain gauge data from Global Telecommunication System (GTS) (Xie and Arkin, 1997). In addition, the WRSI model uses relevant soil information from the FAO (1988) digital soils map and topographical parameters from Digital Elevation Model (DEM) derived data (HYDRO-1K, Gesch et al. (1999)).

WRSI calculation requires a start-of-season time (SOS) and end-of-season time (EOS) for each modeling grid-cell. Maps of these two variables are particularly useful in defining the spatial variation of the timing of the growing season and, consequently, the crop coefficient function, which defines the crop water use pattern of crops. The model determines the SOS using onset-of-rains based on simple precipitation accounting. The onset-of-rains is determined using a threshold amount and distribution of rainfall received in three consecutive dekads. In this study, SOS was established when there was at least 25 mm of rainfall in one dekad followed by a total of at least 20 mm of rainfall in the next two consecutive dekads. The length of growing period (LGP) for each pixel was determined by the persistence, on average, above a threshold value of a climatological ratio between rainfall and potential evapotranspiration. Thus, EOS was obtained by adding LGP to the SOS dekad for each grid cell.

At the end of the crop growth cycle, or up to a certain dekad in the cycle, the sum of total AET and total WR are used to calculate WRSI in a Geographic Information System (GIS) environment at 0.1 degree (about 10 km) spatial resolution. A case of "no deficit" will result in a WRSI value of 100, which corresponds to the absence of yield reduction related to water deficit. A seasonal WRSI value less than 50 is regarded as a crop failure condition (Smith, 1992).

### **Crop Data**

Crop production data were obtained from the Ministry of Agriculture of Ethiopia through the FEWS NET Office, Addis Ababa, Ethiopia. The production data included planted area in hectares and production in tons at the Woreda Administrative Unit. Sorghum production data were available from 1996 through 1999. Only the "meher" season data were used; the meher season represents the main growing season in Ethiopia, encompassing the period between May and October. The meher season accounts for 90-95% of the annual crop production of the country (FEWS NET, 2001).

In order to perform spatial analysis in GIS, the production data were related with district boundary map files using the woreda administrative name in ArcView (ESRI, 2000). Various database utilities from Microsoft Access (Microsoft, 2000) were used to sort the data by year, season, and crop type. Some of the districts did not have a matching district-name between the map boundary file and production table, in which case the production data for those districts could not be used for comparison with the WRSI images.

Microsoft Excel\* (Microsoft, 2000) was used for statistical and graphical data analysis. Yield (tons/ha) per district was calculated as the ratio between production in tons (after converting reported quintals to tons) and planted area (ha).

## **WRSI**

The crop model output (hereafter referred simply as WRSI) was run for the 4-year time period for which there were corresponding crop production data. The model was run to simulate the water requirements of sorghum for the meher season from 1996 through 1999. Typically, the start of season varied from May to July and the corresponding end of season was determined by adding a 3- to 6-month growing period, depending on the agro-climatic region coinciding with the pixel location in the country. WRSI images that were generated as of November 30 of each year were used to characterize the outcome of each season as a whole.

## **Analysis Procedure**

District data were grouped by average WRSI values. Since the WRSI data are in a raster data format, the average value of the pixels located inside each district boundary was calculated to represent the district. Districts were grouped using two criteria: the first grouping is entirely based on 4-year average WRSI values. Districts with an average value of 100 were labeled as "optimum" and were treated differently than districts that had an average WRSI less than 100 ("drought-prone").

There were a 44 districts in the drought-prone category and 23 in the optimum category. The 67 (44 + 23) districts were selected from a grand total of 172 districts that had a Meher (June-September) growing season. The selection was based on the fact that these districts had sorghum yield data reported in all four years. Also, these 67 districts had an average yield less than 2.5 tons/ha. This threshold value was used to screen out possible outliers and to confine analysis to those districts with similar production potential. The mean sorghum yield for all 172 districts was 1.15 t/ha with a standard deviation of 0.84 t/ha. The minimum and maximum yields were 0.12 and 5.0 t/ha.

Based on the second criterion, districts were further grouped based on the percentage of sorghum planted area within a district (percent area). A grouping by percent area was made to test whether the correspondence between WRSI and reported yield depends on the percent and/or size of the planted area in the district. The larger the percent area, the better the expected spatial correspondence between WRSI and actual planted area.

For each year, the mean of WRSI and average yield were calculated for each of the nine groups of districts defined by intervals of percent of area planted. On the other hand, the mean of all 44 drought-prone districts was analyzed to obtain a country-wide assessment of the yield and the corresponding WRSI variability from year to year.

## **RESULTS AND DISCUSSION**

Figures 1 and 2 show the temporal patterns of country-wide average WRSI and yield data for the optimum and drought-prone district groups, respectively. Although the WRSI values remained at maximum (100) as expected in the optimum districts, the yield values varied from year to year. From the basic formulation of WRSI where it only takes into consideration water supply and demand, the year-to-year yield variation in the optimum districts can be attributed to factors other than water. Particularly, the reported yield in 1997 was much lower than the other years, with a reduction of 22% from 1996. The temporal yield variability in the drought-prone districts can be attributed to a combination of water stress and other non-water stress factors. If we can determine the magnitude of variability caused by non-water factors, the yield variability due to water stress can be estimated as the difference between total variability minus the variability due to non-water factors. In this study, the yield variability in the optimum districts is assumed to be a result of non-water factors since the WRSI value was 100 in all years, i.e., there was no rainfall deficit for the crop.

Figure 2 shows that the average WRSI value for the drought prone districts was the lowest in 1997 while 1996 and 1998 remained close to maximum, with 1999 slightly lower than maximum. The corresponding yield patterns followed similar patterns in that the 1997 yield was the lowest and the 1999 yield was lower than both 1996 and 1998. The change in yield is shown to be more sensitive than the change in WRSI values. For example, the WRSI change from 1996 to 1997 is about 9% reduction; however, the corresponding reduction in yield was 52%. As it was shown in the optimum districts, 22 % of the 52 % yield reduction maybe attributed to non-water stress factors. This holds with the assumption that the

---

\* Mentioning a commercial product does not imply an endorsement by USGS/FEWS NET.

non-water stress factors affect both the optimum and drought-prone districts similarly. Therefore, it appears that the yield reduction from 1996 to 1997 as related to water stress maybe estimated as 30%.

Determining the effect of non-water factors on yield reduction should be established differently for different crops. For example, a comparable study on corn did not show yield reduction in the optimum districts in 1997 (Senay and Verdin, 2000), thus yield variability due to non-water factors should be determined for each crop individually.

Such differences between WRSI and actual yield reduction are generally handled by a yield reduction response factor (Ky), which is crop specific and varies from place to place (Reynolds, 1998). A seasonal Ky value for sorghum is published by FAO to be 0.9 (FAO, 1996). This indicates that for every unit of WRSI reduction from the maximum (100), only a 0.9 unit reduction is expected in actual yield, i.e., a 10% WRSI reduction will translate into only a 9% yield reduction from the maximum yield. However, it has to be pointed out that Ky values were established using field experiments using high-yielding varieties that are well adapted to the environment and, more importantly, grown under high level of crop management.

Despite the correspondence in temporal patterns between WRSI and reported yield, the magnitude of the yield reduction in comparison with the WRSI reduction was much larger (more than 3 times) than what is reported in the literature.

At least for Northern Ethiopia, where most of the districts are located, even a slight reduction in WRSI corresponded with large values of yield reduction. One of the key differences between the published yield reduction function and what is reported from the field is the level of crop management practice. While it is stated (FAO, 1996) that Ky values were developed under a high level of crop management, there is no information on the level of crop management practice in the studied districts. In addition, the issue of scale of observation can be very important. The WRSI model is a coarse scale simulation where a single pixel at 10 km could encompass different agro-ecological zones and may comprise several thousand subsistence farmers.

While the correspondence in the temporal patterns between WRSI and reported yield is promising, a yield reduction function needs to be established for each region instead of using Ky values, which were established and intended to be used at a farm level. This can be achieved by establishing a relationship between WRSI and yield at large spatial scales. Since the accuracy of district yield data is very important for this purpose, a systematic data collection with proper quality assessment and quality control is necessary.

Since the WRSI values for a given crop depend on the length of growing periods, which affects the rainfall distribution pattern, it is recommended that a relative WRSI value instead of the absolute WRSI be used. The relative WRSI is defined as the ratio of one year's WRSI to the mean WRSI value of a representative period of years. By using relative WRSI values, one would reduce the effect of using an incorrect length of growing period for a particular region.

The effect of the degree of spatial correspondence between reported yield data and district wide average WRSI was investigated. Some districts have such an elongated shape that part of the area has high WRSI values while the rest of the district has low WRSI values. Since we do not know exactly where in the district the reported yield comes from, a relationship between district yield and district average WRSI could be misleading. Table 1 shows district average yield and corresponding WRSI values for district groups. The district groups are based on percent area of sorghum planted in the district (source of the yield data) while the WRSI average is for the entire district.

The average percent sorghum area in the district varied from as small as 0.6% to a high of close to 22% (Table 1). The temporal comparison between WRSI and yield over different percent area groups did not show a trend to disprove the null hypothesis that the relationship between WRSI and reported yield does not depend on the amount of sorghum area in the region. Generally, in all nine district groups, the 1997 yield and the corresponding WRSI were the lowest. The trend in other years was comparable from one district group to the other. However, this hypothesis had a weakness in that it only considered sorghum area; a small sorghum area may be offset by a large area occupied by another crop. In this instance, total agricultural area may be more appropriate as it relates to the farming potential of the area. Further investigation is required using total agricultural area.

The district groups were used to make a scatter plot of the relationship between WRSI and yield (Figure 3). Due to the nature of the data collection and reporting in the country (and many other countries), individual districts tend to have large errors (personal communication, Ministry of Agriculture Officer, Ethiopia). Because of this, averaging several districts tends to cancel out some of the reporting errors. The averaging of district yield had a similar effect whether it was based on aggregating the district by province or percent-area.

From the scatter plot (Figure 3), it can be observed that there is a good correlation ( $r = 0.77$ ) between WRSI and reported yield. With all 36 data points (9 district groups over 4 years), 59 % of yield variability can be explained by WRSI. The  $R^2$  increased to 0.75 when only the WRSI values below 98 were considered. This certainly demonstrates both the weakness and strength of WRSI as a yield estimator. The weakness of WRSI is manifested with the high variability of

## **ON THE NEED FOR DYNAMIC SCHEDULING OF IMAGING SATELLITES**

yield when WRSI is close to 100. In good rainfall years or optimum regions, factors such as fertilizer use and other management practices will determine yield variability. However, WRSI is showing its effectiveness in estimating yield reduction for drought prone districts, with the  $R^2$  value increasing from 0.59 to 0.75 when only districts with average WRSI value less than 98 were considered. For the purpose of demonstrating the spatially explicit nature of WRSI maps, Figure 4 shows the WRSI image for the relatively dry year of 1997.

**Table 1: Average Sorghum Yield and WRSI grouped by Percent Area.**

Group*	%Area	Crop Area (ha)	Count	Sorghum Yield (t/ha)				Sorghum WRSI			
				1996	1997	1998	1999	1996	1997	1998	1999
A	0.6	773	6	0.98	0.56	1.28	1.08	98	87	100	99
B	1.6	1,356	4	1.08	0.71	1.20	0.94	99	90	100	100
C	2.5	3,413	11	0.95	0.51	0.89	0.65	95	88	100	99
D	3.5	2,778	6	1.24	0.57	1.31	1.17	99	91	100	100
E	4.3	4,562	3	1.07	0.86	1.15	0.92	99	92	99	97
F	5.6	8,313	4	0.89	0.38	0.86	0.64	97	90	100	99
G	7.1	4,877	4	1.29	0.64	0.91	0.83	99	93	99	95
H	10.8	8,755	3	1.56	0.33	1.48	0.95	99	88	100	98
I	21.9	13,721	3	1.42	0.51	1.27	1.17	100	89	99	97
Average:				1.17	0.56	1.15	0.93	98	90	100	98

\*: Below are the names of the districts associated with each group.

Group A: Endamehoni, Gidan, Bugna, Mulona Sululta, Ofla, Hintalo Wajirat

Group B: Enarj Enawga, Kutaber, Alaje, Tenta

Group C: Dehana, Sayint, Enticho, Mekdela, Sekota, Gulomahda, WadlaI, WadlaII, Zikuala, Enticho, Debresina

Group D: Enbise Sar Midir, Werie LeheI, Enemay, Werie LeheII, Ambasel, Wegde

Group E: Meket, Jama, Shebel Berenta

Group F: Kola Temben, Samre, Guba Lafto, Habru

Group G: Degen, Kola Temben, Dejen, Moretna Jiru

Group H: Kobo, Gerar Jarso, Alamata

Group I: Weremo Wajetuna Mida, Lay Betna Tach Bet, Hidabu Abote

## CONCLUSION

The main objective of this study was to evaluate the correspondence between WRSI and reported sorghum yield data. It was shown that the temporal patterns of WRSI and yield corresponded well, particularly in highlighting the relatively dry year of 1997. The grouping of districts into optimum and drought-prone classes appeared to provide a means to estimate impact of crop stress factors not related to water stress. This was accomplished by assuming the yield variability in the optimum districts (where WRSI was always a maximum at 100) to be caused by factors other than water stress such as inputs, crop management and pest damage. For example, the 1997 sorghum yield reduction (52%) from 1996 in the drought prone districts was evaluated to be caused by both water deficit (30%) and non-water stress factors (22%).

It was shown that the coefficient of determination ( $R^2$ ) between WRSI and reported yield increased from 0.59 to 0.75 when only districts with WRSI value less than 98 was considered. This increase in  $R^2$  values demonstrates the strength of WRSI in estimating relative yield change in drought prone areas. Indeed, it is intuitive that the usefulness of WRSI is marginal if water is not a limiting factor since yield variability is controlled by factors other than water.

There was no indication that the percent area of sorghum in a district influenced the relationship between WRSI and reported yield. This was probably due to a wrong assumption in using only sorghum area instead of total agricultural area in the district. Future studies should reevaluate and address the importance of the size of agricultural area while

### ON THE NEED FOR DYNAMIC SCHEDULING OF IMAGING SATELLITES

comparing district-wide WRSI with yield data collected from a fraction of the district area.

This study also showed that a small change in WRSI can result in a larger change in yield. Published yield reduction functions were generally at a plot level whereas the relationship we evaluated comprised of several thousand farm-plots in a district. We recommend that such yield reduction function be established for various regions instead of using or assuming a yield reduction function that was established at experimental stations.

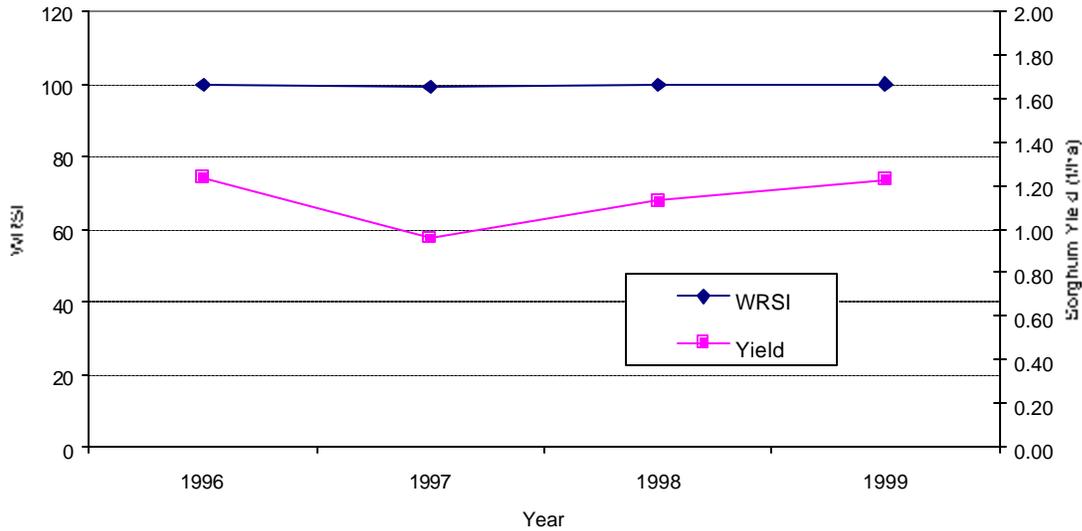


Figure 1: Temporal Patterns of WRSI and Sorghum Yield in 23 "Optimum" Districts

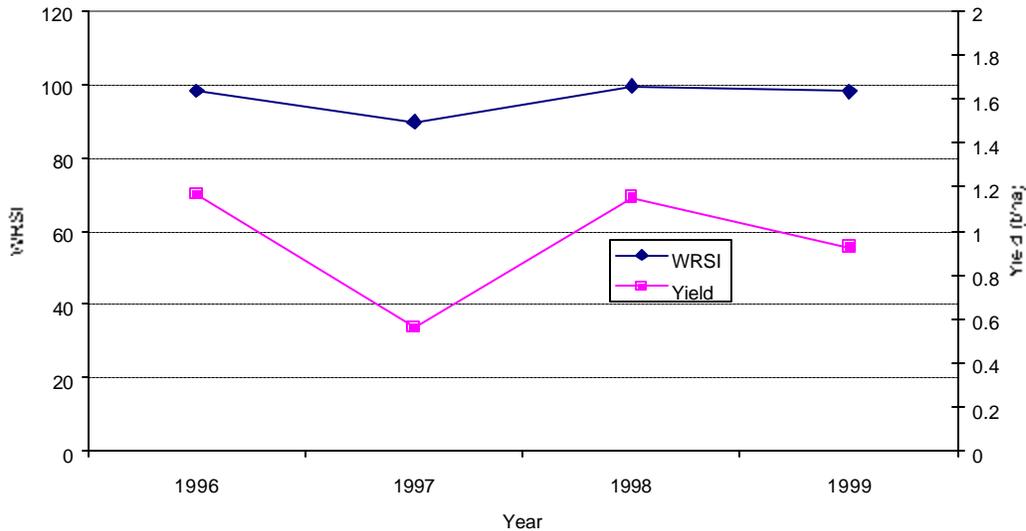


Figure 2: Temporal Patterns of WRSI and Sorghum Yield in 44 "Drought-Prone" Districts.

**ON THE NEED FOR DYNAMIC SCHEDULING OF IMAGING SATELLITES**

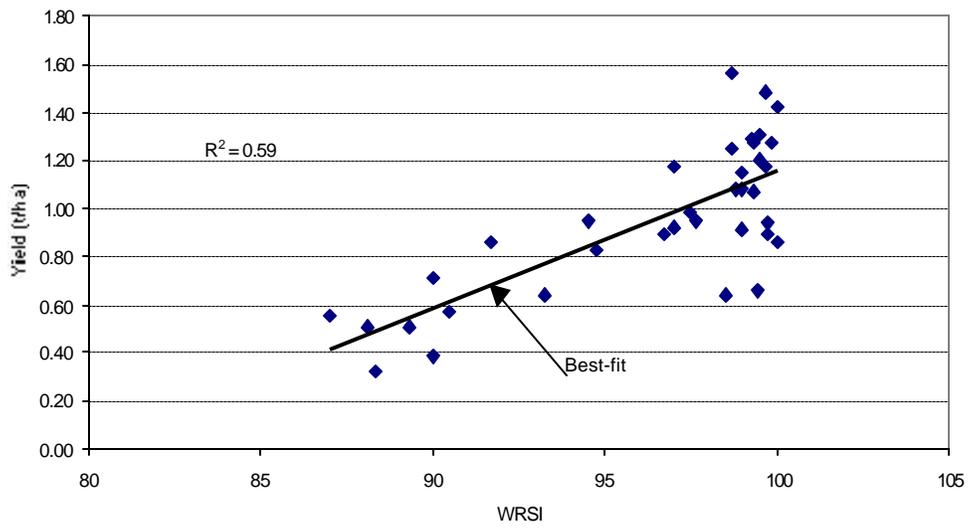


Figure 3: Scatter Plot of Sorghum Yield versus WRSI from 44 “drought-prone” Districts.

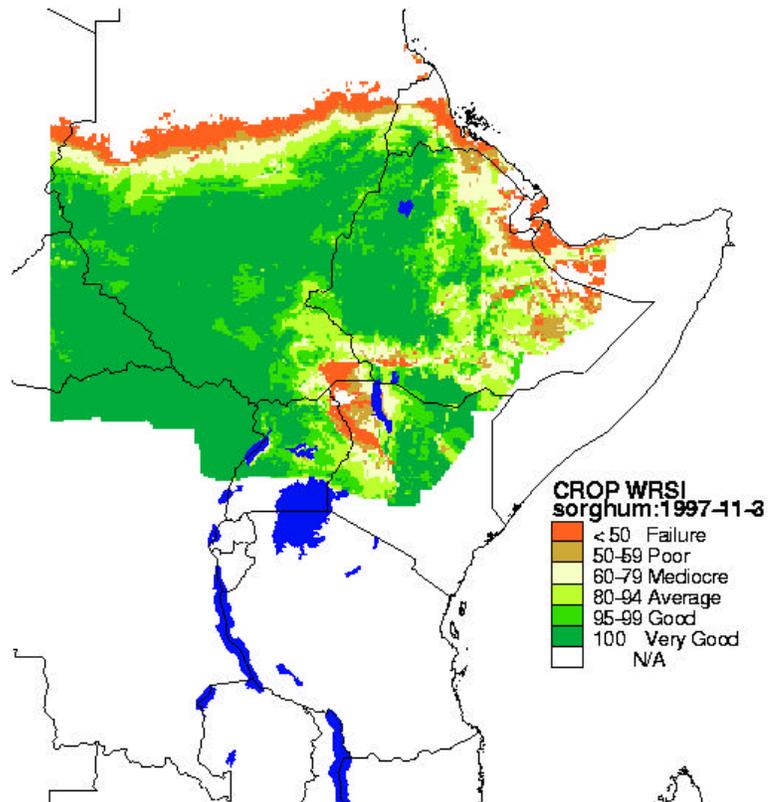


Figure 4: Sorghum WRSI during 1997.

ON THE NEED FOR DYNAMIC SCHEDULING OF IMAGING SATELLITES

## REFERENCES

- Doorenbos J and W.O. Pruitt, W.O. (1977). *Crop Water Requirements*. FAO Irrigation and Drainage Paper No. 24. Food and Agricultural Organization of the United Nations, Rome, Italy.
- ESRI. (1995). Arc/Info Online manual, Redlands, CA.
- ESRI. (2000). Environmental Systems Research Institute, ArcView 3.2a. Redlands, CA.
- FAO. (1988). FAO/UNESCO Soil Map of the World. Revised Legend. World Resources Report 60. FAO, Rome.
- FAO. (1996). Yield Response to Water. FAO Irrigation and Drainage Paper 33. Rome, Italy.
- FAO. (1998). Crop Evapotranspiration. FAO Irrigation and Drainage Paper 56. Rome, Italy.
- FEWS NET. (2001). Ethiopia Network on Food Security. Monthly Report, No. 4/01.
- Gesch, D.B., Verdin, K.L., and Greenlee, S.K. (1999). New land surface digital elevation model covers the Earth. *EOS, Transactions of the American Geophysical Union*, v. 80, n. 6, pp. 69-70.  
<http://edcdaac.usgs.gov/gtopo30/hydro/africa.html>
- Microsoft Corporation. (2000). Microsoft Access.
- Microsoft Corporation. (2000). Microsoft Excel.
- Reynolds, C.A. (1998). Estimating crop yield by integrating the FAO crop specific water balance model with real-time satellite data and ground-based ancillary data. Ph.D. Dissertation, The University of Arizona.
- Senay, G.B. and J. Verdin, 2001. Using a GIS-Based Water Balance Model to Assess Regional Crop Performance. Proceedings of the Fifth International Workshop on Application of Remote Sensing in Hydrology, October 2-5, 2001, Montpellier, France
- Shuttleworth J. (1992). *Evaporation. Handbook of Hydrology*, Maidment, D. (ed.). McGraw-Hill: New York; 4.1-4.53.
- Smith, M. (1992). Expert consultation on revision of FAO methodologies for crop water requirements. FAO, Rome, Publication 73.
- Verdin, J. and Klaver, R. (2002). Grid cell based crop water accounting for the famine early warning system. *Hydrological Processes*, 16: 1617-1630.
- Xie, P. and Arkin, P.A. (1997). A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bulletin of the American Meteorological Society* 78(11): 2539-58.