DEVELOPMENT OF A WEATHER YIELD INDEX (WYX) FOR MAIZE CROP INSURANCE IN MALAWI

Adams Chavula
Agrometeorologist, FAO-SDRN Consultant[1]

and

René Gommes
Senior Agrometeorologist, FAO-SDRN

March 2006

---

[1] Adams Chavula was funded by the **WB – FAO** Partnership fund.

1. Introduction

The Commodity Risk Management Group (CRMG) of the World Bank has developed a prototype Malawi Maize Production Index (MMPI), constructed from weather data recorded by the Malawi Meteorological Office weather stations throughout the country, to capture maize production levels in Malawi. The aim is to find a simple and objective indicator that can be used as a proxy measure of the countrywide exposure of Malawi maize production to drought and hence serve as a nation-wide food security indicator on which an insurance agreement could be written for the Government of Malawi (GoM).

However, this indicator turned out to be relatively poorly linked with crop yields. FAO proposed to use the tools included in the FAO AgroMetShell (AMS) software[2] to derive an effective weather-based maize yield index (WYX, Weather Yield indeX) that could be used for crop insurance purposes in Malawi. AMS computes a crop specific water balance to derive value-added crop-weather variables that can be combined with other data (e.g. remote sensing inputs, farm inputs such as fertilizer use) and statistically related with crop yield using standard multiple regression techniques[3]. "Value-added crop-weather variables" are variables such as actual evapotranspiration that are know to be more meaningful than raw meteorological variables; see, for instance, the documents listed in footnote 3 for theory and references .

The technical cooperation programme between World Bank and FAO also included a visit by an Agro-meteorologist from the Malawi Department of Meteorological Services (DoMS) to Rome (Italy) for a period of three months under the WB-FAO Partnership Programme to work on the procedure to be implemented.

Based on the work of Hess and Syroka (2004)[4] and French (2004)[5], as well as the discussions the second author held in Blantyre and Lilongwe (in particular with DoMS), an index used for crop insurance should have the following characteristics:

- Tamper-resistance: potential beneficiaries of the insurance should not be in a position to directly or indirectly manipulate the index.
- Objectivity: once the methodology has been defined in precise enough terms, the index value should be independent on who carries out the calculations.
- "Good" correlation with crop yield: A "poor year" is defined as a year in which conditions are bad enough to trigger the payment of claims to insurance subscribers. A "poor year" can be defined based on three approaches at least: (1) absolute yield levels (possibly the most appropriate choice for food security), (2) a percentage of the average local yield (a "fair" choice as expectations are different in high potential and low potential areas and (3) probability of exceeding of a specific yield (this usually gives "good" results in

2 ftp://ext-ftp.fao.org/sd/sdr/Agromet/AgroMetShell/

3 General methodological information can be found under **ftp://ext-ftp.fao.org/sd/sdr/Agromet/Documents/**, especially documents agro005a.pdf and agro003a.pdf.

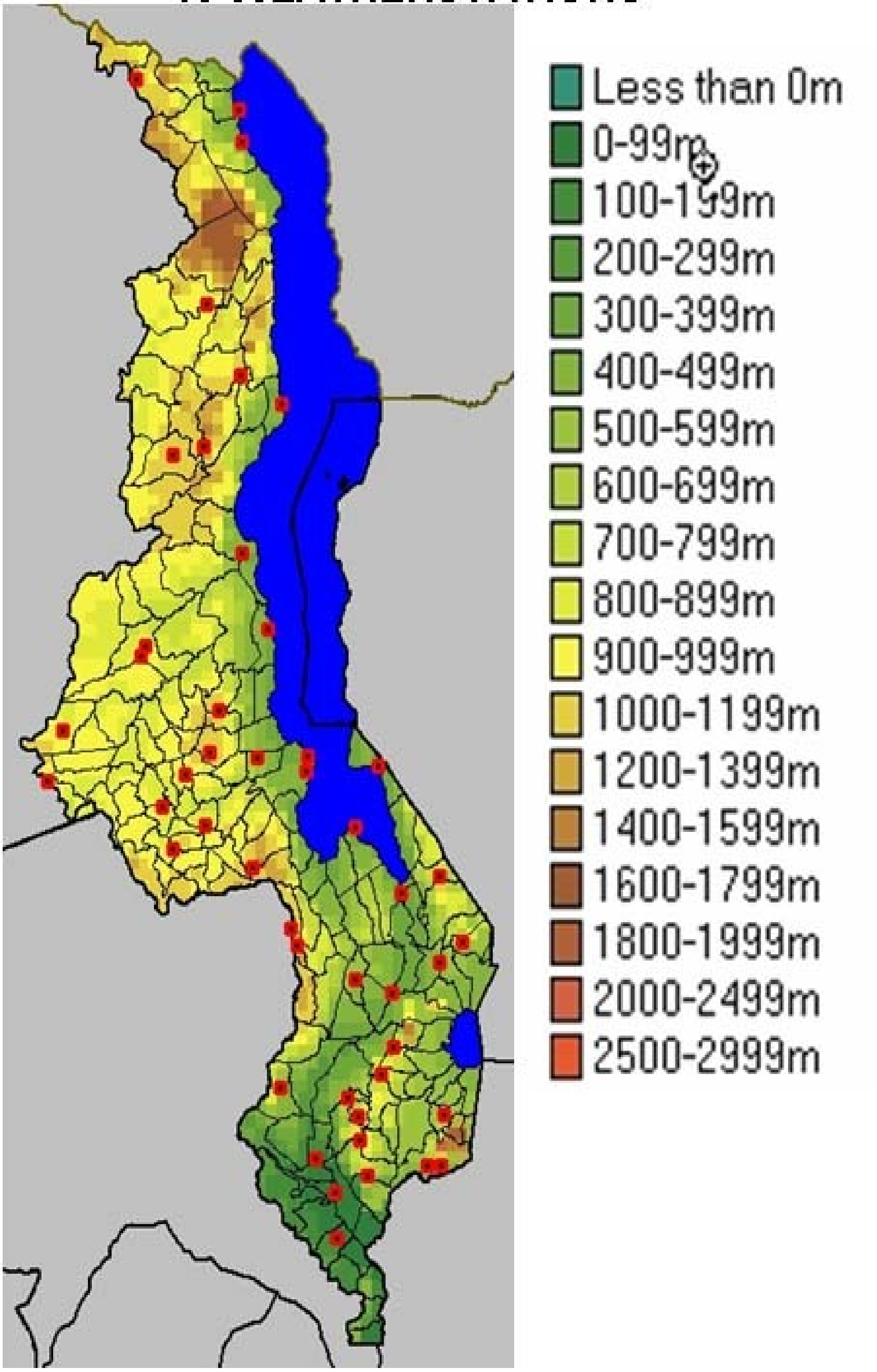
4 **Weather Based Insurance against covariate shocks in Southern Africa, Food Security and Weather Risk Management in SADC, Focus on Malawi**. Report prepared for SADC Secretariat by Ulrich Hess and Joanna Syroka, CRMG/ARD (2004).

5 **Pre-Feasibility Analysis of Index-Based Drought Insurance in Malawi**, unpublished report by Vicky French, available from FAO/ESCG (2004)

terms of statistical significance). Rather than the statistical strength of the correlation between yield and crop weather index, it is the number of false positives (good year assessed to be poor) and false negatives (poor year assessed as good) that constitutes the most important criteria. Some guidance will be needed from ARD/CRMG on: (i) the definition of a poor year and on the criteria to be adopted for test the "goodness" of the methodology; (ii) if the required index has to incorporate yield reductions due to water excess.

- Insensitivity to missing data: the best way to circumvent the occurrence of missing spatial data is to use gridded information that is not too sensitive to individual missing stations, provided sufficient data points are available and the interpolation process takes into account topography and climatic gradients.
- Publicity: the methodology has to be made available to potential subscribers of the insurance. Crop insurance indices will be published regularly in national agrometeorological bulletins and other channels as well as on the website of Malawi Department of Meteorological Services..

FIG.1: MAP SHOWING ELEVATION WITH 46 WEATHERSTATIONS



2. Methodological overview

1) Verify stationarity of time series of meteorological data after 1990: no trend changes occurred during the calibration period (4.2);
2) Derive “pseudo-Penman” equation to be able to compute Potential Evapotranspiration (PET) for all the years of the calibration period (4.1);
3) Define best strategy to determine planting dates based on actual phenological reports (4.6);
4) Determine optimal spatial interpolation technique and compute 556 dekad rainfall grids (46 stations from 1990-2005) and 376 PET (22 stations from1995-2005) grids at a resolution of approximately 5 km (4.4);
5) In the AgroMetShell (AMS) software, create a grid of 3928 points covering the whole country at a resolution of approx 5 km (4.5);
6) Determine altitude of 3928 points, crop cycle lengths and planting dates (based on crop type and elevation) , and read rainfall and PET values (point 4 above) into the AMS database (4.5, 4.6);
7) Compute water balance for the 3928 points from 1994/94 to 2004/2005 (4.6, 4.6, 4.7);
8) Spatialise water balance parameters using Satellite Enhanced Data Interpolation (SEDI) with NOAA NDVI as background images. The water balance parameters mainly include Water Satisfaction Index WSI[6], actual evapotranspiration, water stress and water deficit for phenological phases and the whole crop cycle (4.8);
9) Extract spatial EPA averages of the water balance parameters from the above-mentioned maps (see 8 above) (4.9 and 4.10)
10) Build a calibration matrix of approximately 1365 lines that includes all the spatially averaged water balance parameters (see point 9) and the average EPA yields. The matrix mixes time-series and cross sectional data, i.e. it is made up by 11 years times 154 EPAs. Missing data include 2003 and 2004 plus some others missing;
11) Create spatially continuous maps of average (reference) yields in Malawi (one for “local” maize and one for hybrids (4.11 and 4.12);
12) Identify groups of “independent” variables in calibration matrix through Principal Components Analysis (PCA), and use the variables with the highest correlation with the first components as test explanatory variables;
13) Derive regression equations for “local” and hybrid maize (5.1 and 5.2) and verify statistical validity of equations (5.3).

[6] The FAO WSI is a classical crop monitoring index that expresses the percentage of actual crop water requirements that have actually been met. WSI normally varies from 0 to 100.

FIG. 2: MAPS SHOWING 2005 LOCAL MAIZE YIELD INDEX AND AVERAGE YIELDS

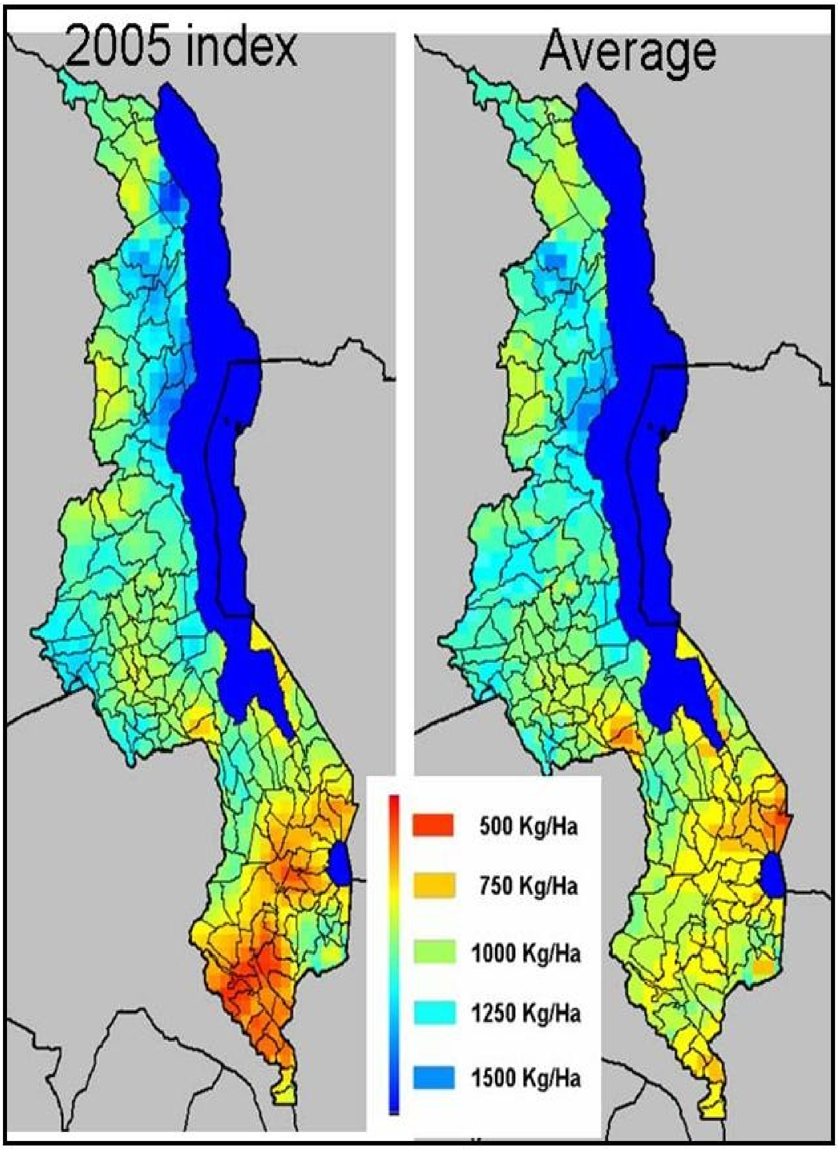
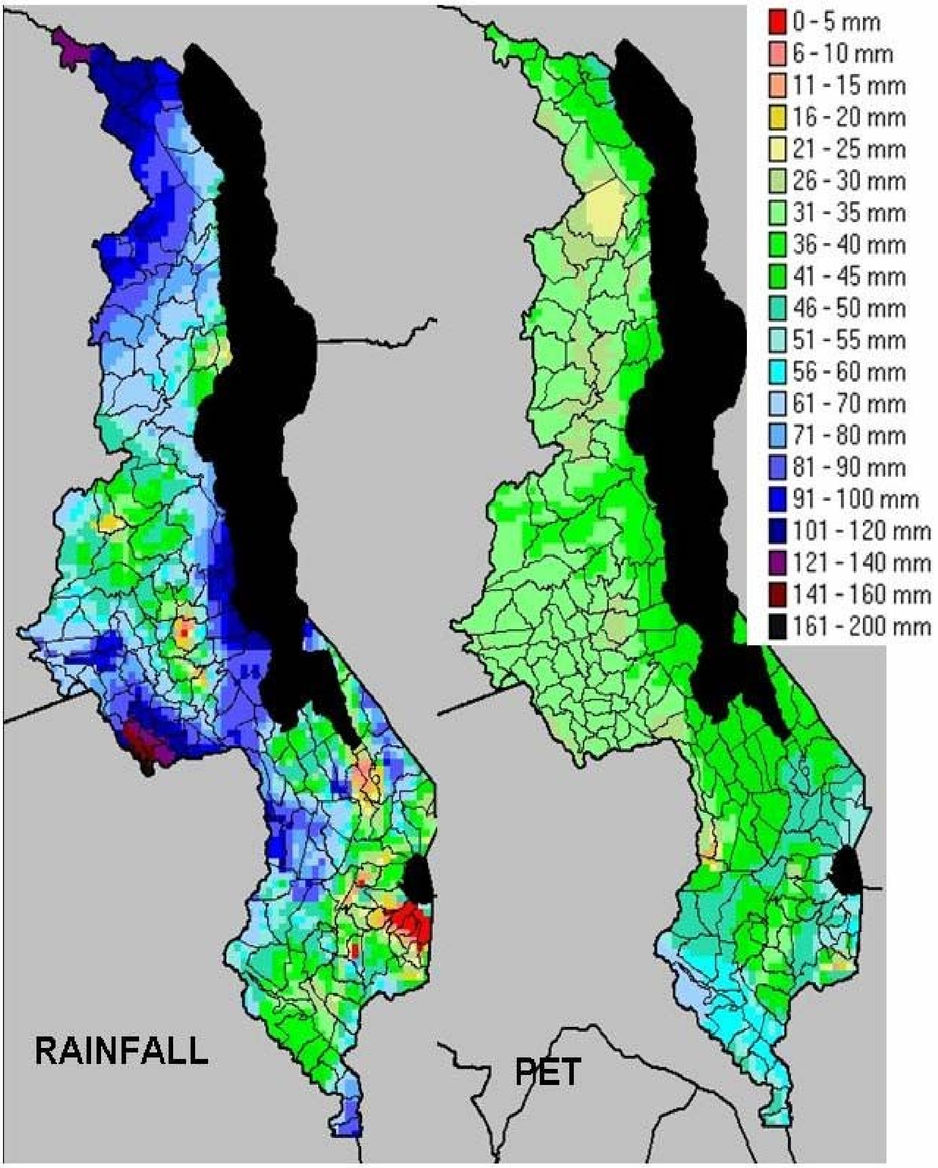


FIG. 3: SAMPLE MAPS SHOWING RAINFALL AND PET GRIDS FOR 01 – 10 JANUARY 2005



3. input data

The weather data used in developing a weather-based Maize Yield indeX (WYX) for Malawi were collected from the Department of Meteorological Services (DoMS), crop and yield data from the Ministry of Agriculture (MoA) through Malawi National Early Warning Unit (NEWU) and Remote Sensing data obtained from FAO/ARTEMIS and through the Southern Africa Development Community (SADC) Regional Remote Sensing Unit (RRSU) in Gaborone, Botswana.

Specifically, the data collected included the following:

a) Long-term (1961-2005) monthly data on Rainfall, Maximum and Minimum temperatures from 22 synoptic stations to carry out a proper trend analysis. The objective was to ensure that the period used to calibrate the WYX is climatically homogeneous;

b) January to December dekadal (10-day period) data on Rainfall, Maximum and Minimum temperatures and Wind Speed from January1990 – May 2005, Sunshine hours from January 1990 – 2003, Relative Humidity January 1995 – May 2005 from 22 full meteorological stations;

c) January to December dekadal rainfall data from 1990 – May 2005 from 24 rainfall stations;

d) Maize phenology data (planting, flowering, maturity) compiled from the Extension Planning Area (EPA) reports collected using the fortnightly phenological reporting forms designed under the Malawi National Early Warning Unit (NEWU) for 1992/93, 1993/94, 1995/96., 1996/97 and 1998/99 growing seasons. The purpose of the data is to ensure that the planting dates determined from the tools in AMS are realistic;

e) Reference data such as soil Water Holding Capacity (WHC), Crop Phenology and Crop Cycle lengths;

f) Area (Ha) and yield (Kg per Ha) data for Maize by Extension Planning Area (EPA) from 1984 to 2005;

g) National Oceanic Atmospheric Administration (NOAA) Normalized Difference Vegetation Index (NDVI) dekadal images for Malawi country window from 1994 –2005. The images were used for the spatial interpolation of water balance outputs using the Satellite Enhanced Data Interpolation (SEDI) technique. See Section 4.8 for details.

4. Tools and details of methodology

4.1 Potential Evapotranspiration: the pseudo-Penman

Missing data characterize most of the real world. For instance, from 2004 there are no radiation data in Malawi due to unavailability of the charts needed to record sunshine hours with the Campbell-Stokes Sunshine recorder. This affected the computation of actual Potential Evapotranspiration (PET) using the conventional Penman-Monteith method which requires sunshine hours. Therefore, a simplified PET Method that does not utilise sunshine hours had to be developed.

The pseudo-Penman equation that was eventually adopted for Malawi is the following (PE in mm/mth):

$$\text{pseudo-Penman} = 41.79 - 1.38 * RH + 13.24 * wind + 3.97 * Tmax + 2.92 * Tmin$$

This performs significantly better than the Blaney-Criddle and the Hargreaves methods. The coefficient of determination is **0.930373**. This was therefore, used to compute actual dekadal PET values (mm per dekad) for 22 full meteorological stations. Complete data for PET computations were only available for the period 1995 – 2005. Prior to this period relative humidity data were not available at most stations.

4.2 Trends in the time series

Southern Africa has apparently been undergoing some climatic trend changes in the last decades or two, in particular with the 1991/92 El Niño and the associated drought, a very rare occurrence in the region over the previous decades. It is, therefore, important to ensure that the time period used for the calibration of crop forecasting methods is devoid of any significant trends that can invalidate the results[7].

As regards precipitations, one observes little change in the monthly long-term series of the stations of Chileka, Chitedze, Mzimba, Salima and Thyolo that were selected for the analysis of the stationarity[8]. In Chileka one observes a significant increase in the in January rainfall, starting in 1984[9], the average passing from 181,2 mm to 238,3 mm. In Thyolo, one also observes an increase in the rainfall in 1984, the average passing from 203,9 mm to 302,2 mm and a fall of the rain in October 1989, the average passing from 44,2 mm to 18,8 mm.

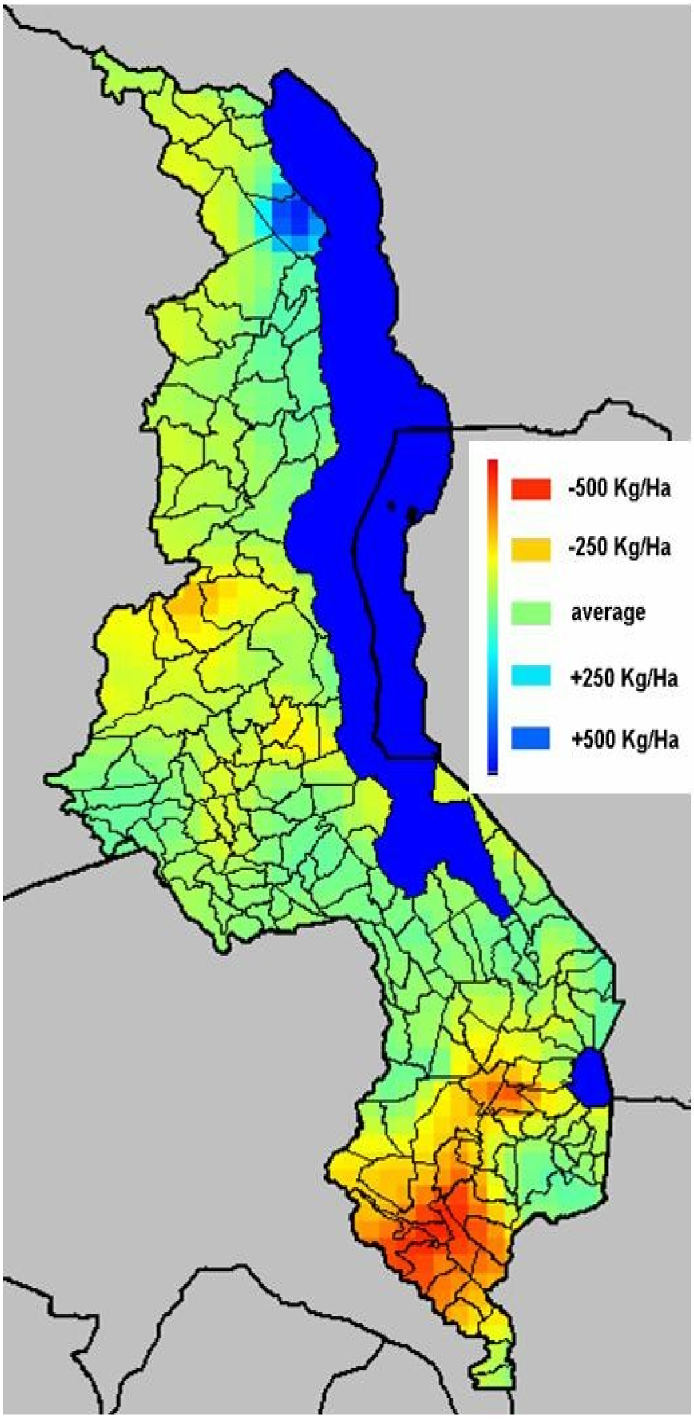
The changes for temperatures are more significant. In general the key-years for the changes of the minimum temperatures are 1982 or between 1986 and 1990. For the maximum ones, the majority of the changes generally take place between 1980 and 1982.

[7] The analyses were kindly carried out by Marc Vandiepenbeeck of the Belgian Meteorological Service.

[8] Stationarity means that there has been no significant change in the average or time trend over the period considered.

[9] 1984 corresponds to the maximum of the drought in the Sahel, confirming the negative correlation between Sourthern and Western African rainfall (http://www.fao.org/sd/eidirect/eian0004.htm)

FIG. 4: MAP SHOWING 2005 LOCAL MAIZE YIELD INDEX DEPARTURE FROM EXPECTED VALUE



This confirms that rainfall and temperatures are indeed stationary for the period used for the calibration of the maize yield index (1995-2005).

4.3 Software tools

In this study FAO – SDRN (Environment and Natural Resources Service) software tools namely, New_LocClim, AgroMetShell (AMS) and WinDisp[10] were extensively used. All of them are easy to use, public domain software, developed and maintained by FAO and its cooperating partners.

New_LocClim was developed to provide an estimate of climatic conditions at locations for which no observations are available. New_LocClim provides nine different interpolation methods that can be compared with respect to pre-given spatial fields. It allows for an extensive investigation of interpolation errors and the influence of different settings on the results. This allows optimising the interpolation with respect to the data analysed. Furthermore, statistical analysis of the interpolated spatial fields is provided and detailed analysis for single geographic points can be drawn.

Windisp is a software package for the display and analysis of satellite images, maps and associated databases. WinDisp was originally developed for the FAO Global Information and Early Warning System.

AgroMetShell is a software tool designed to support crop forecasting. The central part of the software is formed by the Crop Water Balance calculations. Based on rainfall, Evapotranspiration and crop data AgroMetShell can calculate if and when a crop experienced water shortage, eventually leading to reduced crop yields.

4.4 Preparation of climate weather grids

Climate and weather data were collected from 22 full meteorological stations and 24 rainfall stations located throughout Malawi. Figure 1 shows elevation and distribution of the forty six (46) weather stations. The data were systematically gridded (spatialised) before any water balance calculations were conducted[11]. This has the main advantages of (1) providing nation-wide coverage (as opposed to stations with a limited are of spatial representativeness) and (2) significantly reducing the potential negative impact of missing data.

Dekadal spatial grids with a resolution of 0.05 degrees latitude and longitude (approximately 5 km) were created using New_LocClim software. The data were reformatted to fit the format required by New_LocClim. New_LocClim requires the data to be formatted by Longitude, Latitude, Altitude and value. Therefore 36 dekadal spatial grid files were made in each year from January 1990 – December 2004 and 16 files during 2005. Sample maps showing actual dekadal rainfall and PET grids are given in figure 3. For this study, among the interpolation methods available in New LocClim, the Kriging method was selected because it produced minimal mean-

[10] All are available from the ftp site indicated above for AMS.
[11] This represents a very significant amount of work...

square errors between grid estimates and station data. Kriging method provided the best results with the following settings:

Grid Co-ordinates:
- Longitude from: 32.5° to 36°
- Latitude from : -17.5° to -9°
- Grid size: 0.05° (approximately 5Km)
  - Shadow correction performed.
  - Weight: 1
  - Darkness: 1
  - Breadth: 1
  - No negative interpolated values permitted.
  - Maximum Number of Stations used: 5
  - Maximum Distance of Stations used: 300 km, and 400 Km for PET grids only.

PET grids were created for the period 1995 – 2005 only.

The grids were exported from New LocClim to WinDisp images. Image headers were edited for uniform slope and intercept values using ad hoc tools[12] . All dekadal rainfall images were formatted to a slope of 2.37 and an intercept of 0 while the slope for PET images was 0.5 with an intercept of 0. Dekadal Rainfall and actual PET images were later imported into a grid-based station list created in AgroMetShell database (see section 3.5 below).

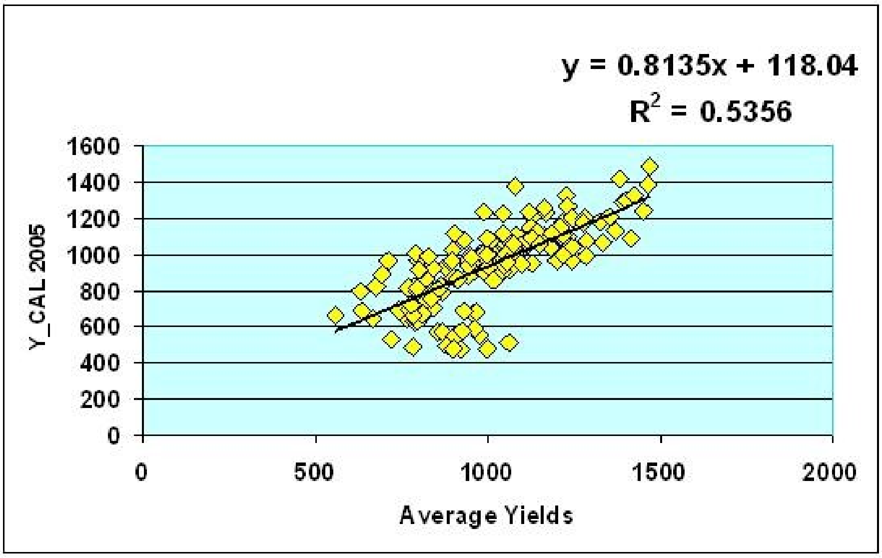
4.5 Creation of station list in AgroMetShell

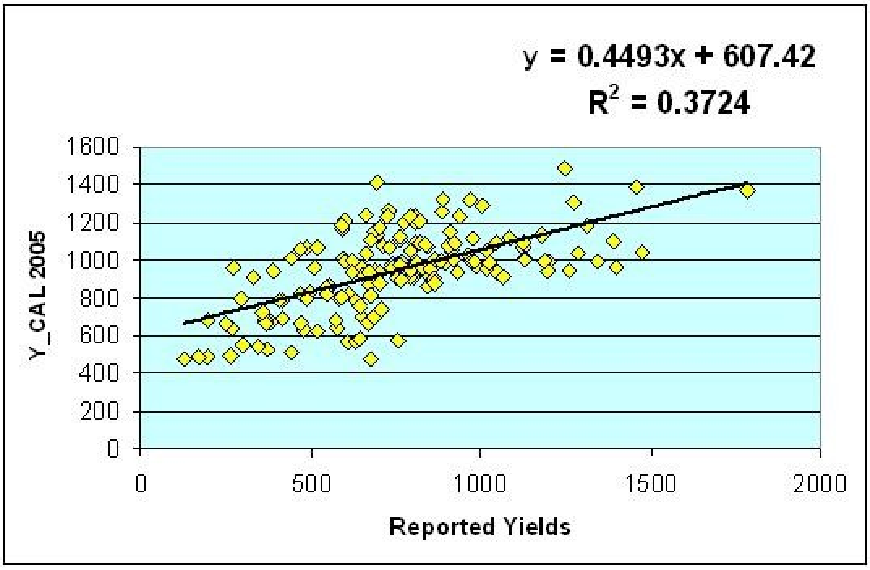
A grid-based station list was created using **"Tools – Create Station list from grid function"** and a boundary map of Malawi (MW_ADM1.BNA) in AgroMetShell. The list was named **"Malawi Grid".** The stations were created at a resolution of 0.05 degrees latitude and longitude (approximately 5 km). The process generated a total of 3928 monitoring points within the borders of Malawi.

Most of weather and climate data needed to analyse the impact of weather on crop yields were imported into a **"Malawi Grid"** in AgroMetShell database from WinDisp images. Dekadal rainfall and PET data for each year were imported from dekadal rainfall and PET images respectively. Similarly normal dekad rainfall and PET data were imported from normal dekadal rainfall and PET images. Station altitude data were imported from Southern Africa Altitude image (Sa_Alt.img). Crop data were imported from crop parameter file which contains data on **soil water holding capacity, effective rainfall, the planting dekad, the cycle length, pre-season crop coefficient and irrigation data** if needed. Crop cycle length data was imported by assigning values to the Southern Africa Altitude image. In Malawi for calculation of FAO Crop Specific Water Balance, the Crop cycle for local and composite maize is taken as 12 dekads or 120 days for elevation of up to 800m and 14 dekads or 140 days for elevation of higher than 800m. Similarly, the crop cycle for hybrid maize is taken as 10 dekads for elevation of up to 800m and 12 dekads for elevation of higher than 800m. Water Holding Capacity (WHC) of the soil was assumed at 50 mm.

[12] See **http://www/hoefsloot/com/agrometshell.htm**, „Two independent implementations of a slope/intercept changer for IDA images"

FIG. 5: CALCULATED 2005 YIELD INDEX BASED ON WEATHER DATA Vs AVERAGE AND YIELDS REPORTED BY MINISTRY OF AGRICULTURE





4.6 Definition of Planting Dekad

Reported average planting dates for maize were summarised by EPA for 1992/93, 1993/94, 1995/96., 1996/97 and 1998/99 seasons. Analysis of the data showed variations in planting periods from one season to another. Based on this information in most EPAs observed average planting dekad for maize was found to occur when the rising rangeland-index (RI) curve crossed the 40% threshold. Therefore, in AgroMetShell the planting dekad for maize was defined based on a rangeland-index of 40% between October in one year and February the following year. The "**rangeland-index**" (RI) is the classic FAO water satisfaction index WSI computed for periods of 5 dekads, with normal Evapotranspiration kept at potential level (KCR=1) and an assumed WHC of 50 mm. Similar to moving averages, the value assigned to a dekad corresponds to the five-dekad period centred about that dekad. Effective rain (Efrain) was assumed to be 100%. Effective rain is a percentage by which actual rain is multiplied to assess actual water supply. Usually Efrain is less than 100% on slopes and Efrain more than 100% in low lying areas. A pre-season crop coefficient (Kcr) of 0.3 was used. The model is run for rain – fed crops only therefore irrigation data were not required. The values of irrigation and the height of the bund for irrigated crops were entered as 0.

4.7 Running FAO Crop Specific Water Balance model

The FAO Water Balance model was prepared to monitor growth and development of two maize "crops" in Malawi namely, "local+composite" (simply referred to as "local" in this report) and "hybrid" for each season from 1994-95 to 2004-05. The Crop cycle for local and composite maize was 12 dekads for elevation of up to 800m and 14 dekads for elevation of higher than 800m. Similarly, the crop cycle for hybrid maize was 10 dekads for elevation of up to 800m and 12 dekads for elevation of higher than 800m.

4.8 Making images from water balance model outputs

Calculations of water balance model in AgroMetShell (AMS) were done on a 3928-point grid with a resolution of 0.05 degrees latitude and longitude, as indicated above. The results of water balance model in AMS are saved in the "summary dekad output" file. The water balance variables were converted into images with maximum and average NOAA NDVI images during each season as background images to guide spatial interpolation[13] (see 4.10). In each season maximum NOAA NDVI image was used to guide interpolation of water excess images over the phenological phases described as **initial, vegetative, flowering, ripening, and whole cycle** (sum of previous values). Similarly, an average NOAA NDVI image was used to guide interpolation of water deficit images described as **initial, vegetative, flowering, ripening, and total for whole cycle** (sum of previous values), actual Evapotranspiration and actual crop Evapotranspiration images over the phenological phases and the value totalled over the growing season.

4.9 Extracting average statistics by Extension Planning Area (EPA) in WinDisp

The FAO Crop Specific Water Balance model produced various water balance parameters for the various stages for growth and development of maize for each

[13] The Satellite Enhanced Data Interpolation (SEDI) function of AMS was used.

season from 1994/95 to 2004/2005. The variables included excess soil water, actual Evapotranspiration, soil water deficit over the initial, vegetative, flowering and ripening stages. Other products included total water requirement and water requirement satisfaction index at the time of monitoring as well as at the end of the growing season. EPA – wide average statistics were extracted from images of the various water balance parameters using the **"Process – statistics function"** and an EPA boundary map in WinDisp. Eventually the average statistics for the various water balance parameters were calibrated against average EPA maize yields for two maize "crops" in Malawi namely, "local+composite" and "hybrid" using Microsoft excel. Relationships between various water balance parameters and EPA maize yields were established. Significant variables were selected for final regression models. In this study "local+composite" is modelled as local maize (LMZ) while "hybrid" as hybrid maize (HMZ). The original selection of the variables was done through a Principal Components Analysis[14] (PCA).

4.10 Creating seasonal Maximum and Average NDVI images in WinDisp

Maximum NOAA NDVI image for each growing season was created using **"Process - Series - Max function"** in WinDisp. Dekadal NOAA NDVI images for Malawi for the period October to May were used in the series.

3.11 Creating EPA Average Co-ordinates in WinDisp

Maize yields from 1984 to 2005 were presented in Malawi old Extension Planning Area (EPA) format. EPA average maize yields for the series were calculated. The coordinates for the centre of gravity of each EPA polygon were computed. A file with EPA central point latitudes and longitudes and average yields was created.

3.12 Creating Average Yield Image Map in New LocClim

The average yield image map was created in by using EPA central point coordinates and average yields and a grid with a resolution of 0.05 degrees latitude and longitude. The average yield (Yavg) image map for each crop was used in the final maize yield index forecasting equation.

[14] When the results of the current work were presented to a group of statisticians and economists in Rome, they suggested that the first PCs could be used instead of the agrometeorological variables. While recognising that the suggestion is a very sound one from a statistical point of view (independence of explanatory variables), the authors prefer to use the variables that display the highest correlation with the first Pcs. The main reason for that is to retain some control over the „meaning" of the variables.

5. Results of yield calibration

5.1 Local Maize

Average EPA maize yields were regressed against selected significant variables, mixing cross-sectional and time series data. This results in about 1350 observations corresponding to the number of EPAS times the number of annual yield values (per EPA).

The regression coefficients and other related statistical data are shown in Table 1. The table also shows the number of observations, number of degrees of freedom, standard error of average maize yield estimates (Yavg) and standard errors of the regression variables.

Table 1: The regression coefficients used to forecast local maize yield (Ycal(local)) with relation to average yields, total water deficit and excess, Evapotranspiration at vegetative stage and water deficit at vegetative stage.

SUMMARY OUTPUT						
Regression Statistics						
Multiple R[15]	0.967					
R Square	0.935					
Adjusted R Square	0.934					
Standard Error	276.038					
Observations	1365.000					
ANOVA						
	df	**SS**	**MS**	**F**	**Significance of F**	
Regression	5.0	1481055742.4	296211148.5	3887.4	0.00000	
Residual	1360.0	103628035.1	76197.1			
Total	1365.0	1584683777.5				
	Coefficients	**Standard Error**	**t Stat**	**P-value**	**Lower 95%**	**Upper 95%**
Intercept	0.00000	#N/A	#N/A	#N/A	#N/A	#N/A
Yavg	0.93242	0.031	30.515	0.00000	0.8725	0.9924
DEFtot	1.81576	0.141	12.907	0.00000	1.5398	2.0917
WEXtot	-0.17464	0.038	-4.551	0.00001	-0.2499	-0.0994
ETAveg	2.35743	0.397	5.938	0.00000	1.5786	3.1362
DEFveg	-26.50667	3.942	-6.725	0.00000	-34.2389	-18.7744

Therefore, the final Malawi local maize forecasting equation is, for each pixel,

$$\text{Yield}=0.93242*\text{Yavg}+1.81576*\text{DEFtot}-0.17464*\text{WEXtot}+2.35743*\text{ETAveg}-26.50667*\text{DEFveg}$$

where **Yield** is the weather – based yield index for local maize, **Yavg** is average yield for local maize, **DEFtot** is total water deficit, **WEXtot** is total water excess, **ETAveg** is Evapotranspiration at vegetative stage and **DEFveg** is water deficit at vegetative stage.

The required average local maize yield image was created in New LocClim and converted into a WinDisp image. The size and projection parameters of this image were different from others that were created in AMS. Therefore, the yield image was rescaled using one of the images that were created in AMS as a reference image. Pixel-based maize yield index maps covering the whole country with a resolution of 0.05 degrees latitude and longitude (approximately 5 km) were produced for each

[15] For regressions through the origin, R and Rsquare are meaningless. Actual values are those given in figure 5.

season from 1994/95 to 2004/2005 using **"Tools – Image Calculations with formula"** in AgroMetShell. In this case the final Malawi local maize forecasting equation was used in the calculations. EPA average local maize yield indices were extracted using WinDisp. The results were compared against actual average and EPA yields reported by Ministry of Agriculture. Summary of the results are presented in figures 5 and 6. Maize yield index maps will be produced and updated every ten days.

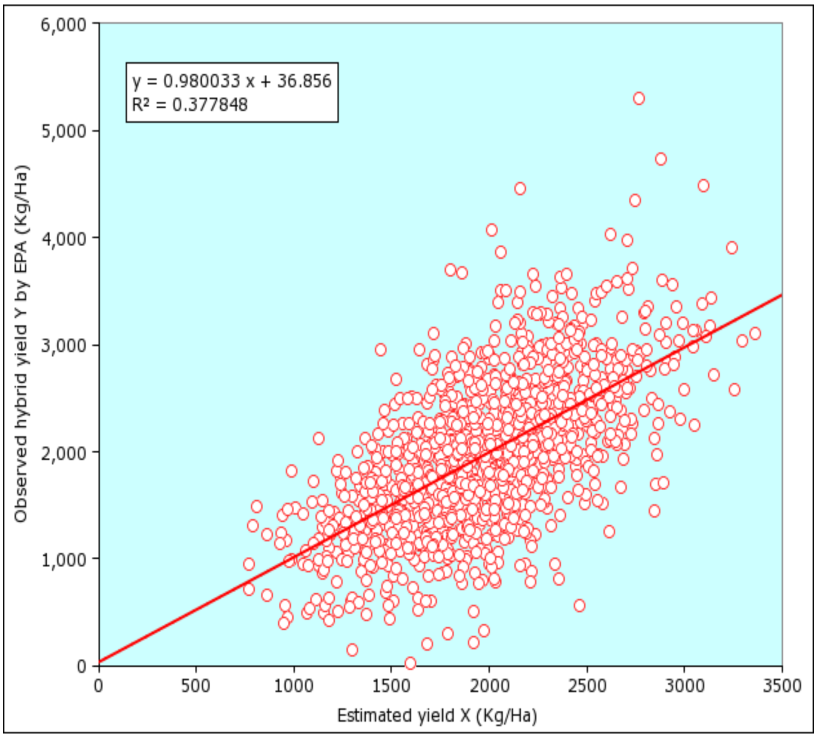
5.2 Hybrid maize

For hybrid maize, the following equation was derived

$$Yield/Yavg=0.03233Year-2E\text{-}05Year^2+0.01004WRSIfin-0.008ETAtot+1.2E\text{-}05\ ETAtot^2$$

where **Yield** is the weather – based yield index for hybrid maize, **Yavg** is average yield for hybrid maize, **Year** is the year for which the calculations are done, **WRSIfin** is the Water satisfaction index computed for the end of the crop cycle and **ETAtot** is the total ETA over the cycle.

Figure 6: hybrid maize yields by EPA Vs estimated yields



Compared with local maize, the main differences include the following: both **Year** and **ETAtot** appear with a quadratic term. This non-linearity corresponds to a slight downward trend of hybrid maize yields, while such a trend was absent from local maize. Regarding ETA, the effect of this variable on yields is negative, with a minimum (maximum negative effect) that occurs around 350 mm, corresponding to the average of EPA observations. Altogether, the main and dominant variable is the water satisfaction index.

The table below shows the regression statistics

Table 2: regression coefficients used to forecast hybrid maize yield with relation to average yields, time trend, ETATtot and final water satisfaction index WRSIfin.

	Coefficients	Standard Error	t Stat	P-value
YEAR	0.0323309380677901	0.00255793887897	12.639449024224	1.0548166419691E−34
Year2	−1.58291127762E−05	1.2758615224E−06	−12.4066072200787	1.4643498754567E−33
WRSIfin	0.0100372456321416	0.00105366099872	9.52606734459727	7.1961250677491E−21
ETAtot	−0.007994982189596	0.00218636598184	−3.65674468776174	0.00026518906534861
ETAt2	1.217215811356E−05	3.1260238495E−06	3.89381485856269	0.00010347055910971

5.3 Validation

Very detailed validation was done for local maize. No statistically significant values of the regression coefficients were found for the following situations:

- coefficients are computed with the whole data set of 1365 observations
- coefficients are computed with early years and late years (data sorted according to time, and split in the "middle" OF THE DATA SET)
- the data set is randomly split in two halves

This indicates the stability over time and across EPAs of the derived equations.

6. Conclusions

The methodology has demonstrated the possibility of producing weather based maize yield index for Crop Insurance for any point in Malawi every ten days starting from planting time. Real-time pixel based maize yield indices covering the whole country with a resolution of 0.05 degrees latitude and longitude can be objectively produced. The methodology is repeatable by anybody who has access to basic weather data. The methodology uses gridded information that is not too sensitive to individual missing stations, provided sufficient data points are available and the interpolation process takes into account topography and climatic gradients. The methodology is temper-resistant, potential beneficiaries of the insurance are not in a position to directly or indirectly manipulate the yield index.

Therefore, the developed maize yield index satisfies all the desirable criteria for maize crop insurance in Malawi. First estimates of Index can be provided at planting time and updated in real time throughout the season. However, the index needs to be refined using criteria to be provided by insurance experts. More specific products for crop insurance can be prepared, and the methodology can easily be extended to other crops.