A Change Point Analysis of Drought Frequency in East Africa

Chao Han (with Dong-Yun Kim)
Department of Statistics, Virginia Tech

April 15, 2010

Outline

- Drought in East Africa
- Literature Review
- Standardized Precipitation Index
- Poisson Process and Inter-arrival Time
- Change Point Test
- Application to 42 sites in East Africa
- Application to CRU data
- Drought Frequency Change and ENSO

Drought in East Africa

- Drought is a recurrent phenomenon of tremendous importance in East Africa. Food security and socio-economic well-being of African nations is highly influenced by the large rainfall variations. Particularly, the knowledge of drought frequency change is of significant value to decision makers and farmers.

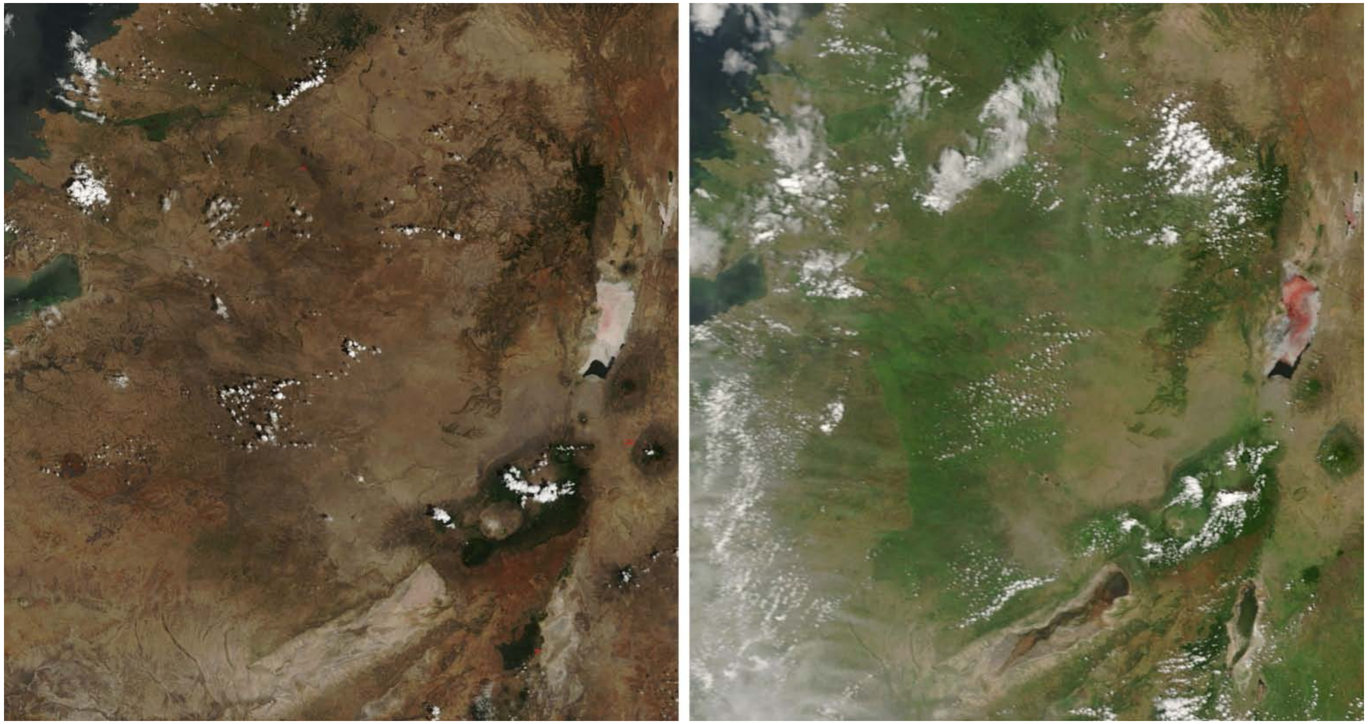


Fig1: Image from NASA satellite for Serengeti Plain on Jan 2006 (left) and Jan 2005 (right)

Literature Review

- Some African climate change studies have been conducted (Hulme et al., 2001; Nicholson, 2001). Long term or year-to-year African rainfall changes have been studied by Nicholson, 1989; Janowiak, 1988; Hulme, 1992; Nicholson, 1993.
  - Those works mainly visually determine changes based on the graphing of inter-annual rainfall over time and no objective statistical test of the presence of change was used.

- Ogallo (1979) used Spearman rank correlation test to investigate trends in the annual rainfall data.
  - This test can not detect the change point, i.e., it can not answer the question about when did the change happen.

- Although numerous works talked about the rainfall changes in Africa, few studies have dealt directly with drought frequency change problems.

Virginia Tech

Standardized Precipitation Index (SPI)

- There are many definitions for drought and indices. We use SPI by McKee et al. (1993)
- Since its introduction, has been widely used for monitoring drought in U.S. and elsewhere; NOAA (National Oceanic and Atmospheric Administration) routinely publishes monthly SPI in its website
- Easy to calculate: SPI can be computed solely based on precipitation.
- Easy to interpret: SPI value indicates deviation from long-term precipitation mean; negative SPI for drought, and positive for wet conditions
- Tsakiris and H. Vangelis (2004) define an event a drought when the SPI is continuously negative and reaches an intensity where the SPI is −1.0 or less. The event ends when the SPI becomes positive.

Virginia Tech

Drought Classification by SPI

Table1: drought classification according to SPI values and time scales

SPI value	Drought Classes
0 to -0.99	Near normal
-1.0 to -1.49	Moderate drought
-1.5 to -1.99	Severe drought
-2.0 or less	Extreme drought
SPI time scale	**Associated Drought Type**
6-month	Agricultural drought
12-month	Hydrological drought

Depends on the purposes for most of the drought studies, we choose the time scale to be 6-month.
Information from a 6-month SPI can be associated with agricultural drought.
We also limit our concern on the moderate drought where SPI falls between -1 and -1.5.

Poisson Process and Inter-arrival Time

- The occurrence of many extreme climate events can be modeled by Poisson process N(t).

- Let N(t) denotes the number of drought up to current time t. N(t)=k if and only if exactly k events happened up to time t. Let W_k be the time of occurrence of the k^{th} event. $X_k=W_k - W_{k-1}$ denotes the time between the $k\text{-}1^{th}$ and k^{th} events, and those durations between successive events X_1, X_2,… are the inter-arrival times.

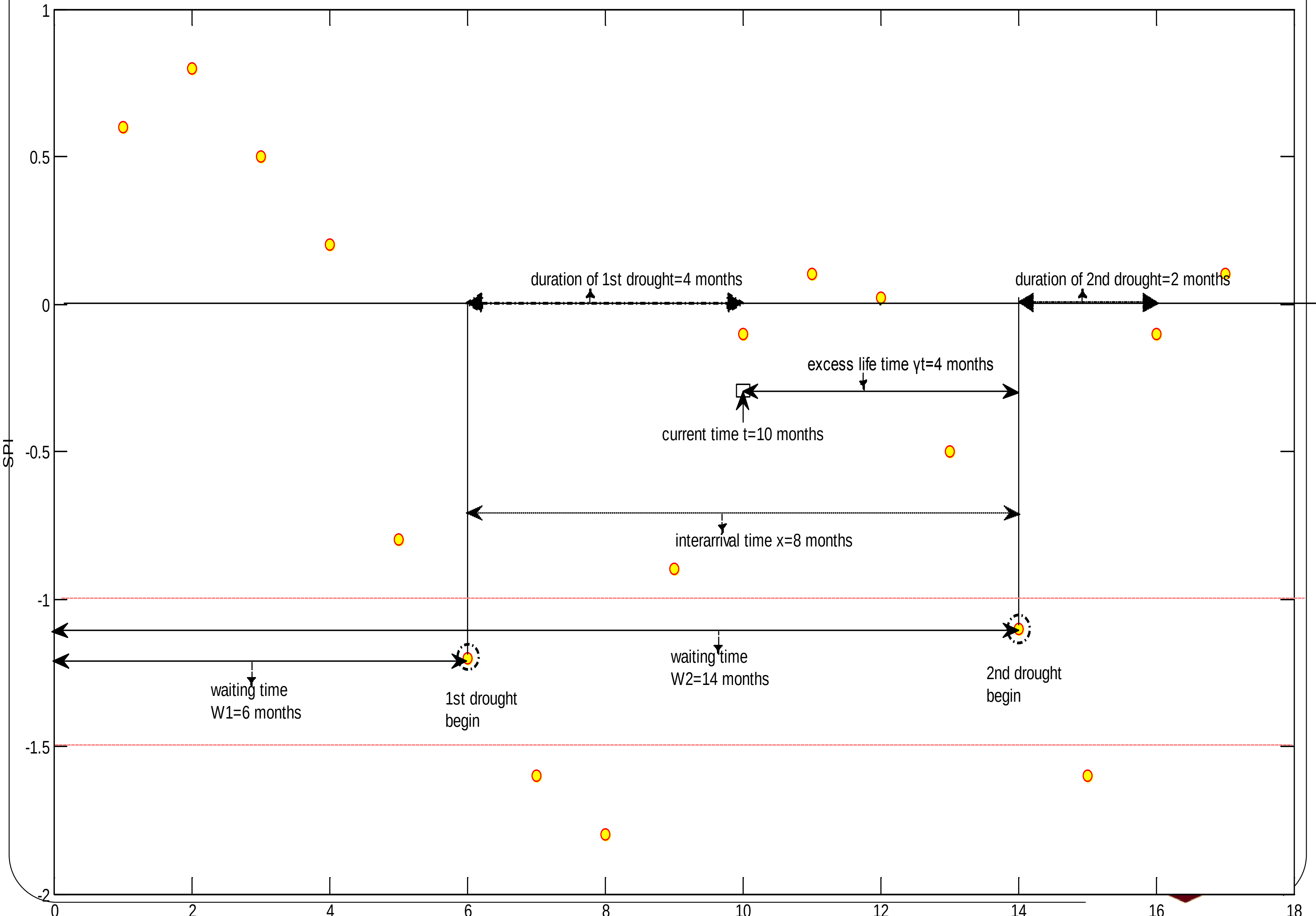
- N(t) is a Poisson process if the inter-arrival times are independently and identically distributed with exponential distribution, which is defined by its probability density function:

$$f(x \mid \theta) = 1/\theta e^{-x/\theta}, x \geq 0, \theta \geq 0 \qquad (1)$$

where x is the inter-arrival time and the mean of x : E(x)=θ.

Virginia Tech

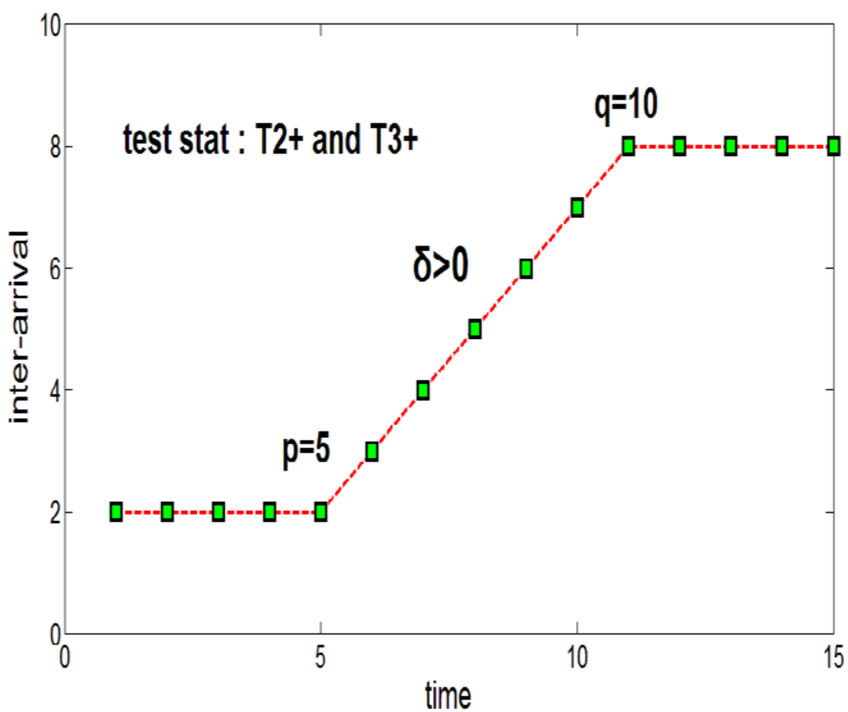
Fig2: Inter-arrival Time

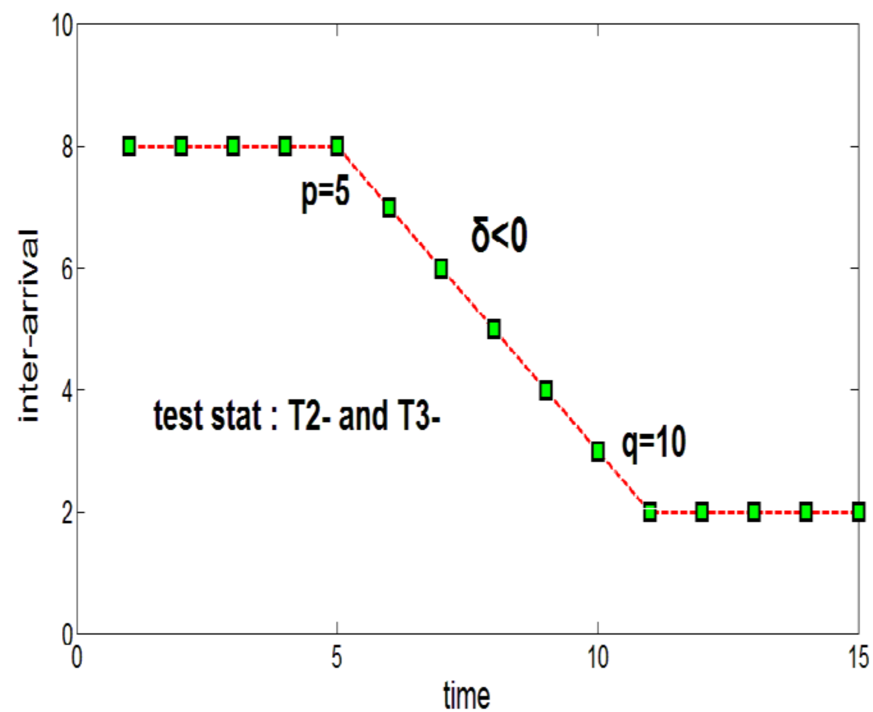


Change Point Test

- The objective of this study is to detect whether drought frequency has changed during a long period of time, and if there is a change, when the change happened or where is the "change point" in the time series.

- The points on either side of the change point are on average bigger or smaller than the other points.





Change Point Test (cont'd)

- The test statistics come from Gupta and Ramanayake (2001)

- Mathematically, suppose X1, …, Xn (inter-arrival times here) is a sequence of independent exponential random variables with distributions: $f(x \mid \theta) = 1/\theta e^{-x/\theta}$ i= 1,…,n. Our test statistics are able to test which of the following hypotheses Ho or Ha is true:

$$H_0 : \theta_i = \theta, vs. H_a : \theta_i = \begin{cases} \theta, & i \le p, \\ \theta + \delta \dfrac{(i-p)}{(q-p)}, & p < i \le q, \\ \theta + \delta, & q < i \le n. \end{cases} \qquad (2)$$

where p and q are the unknown change points such that $1 \le p < q \le n$. Actually, p denotes the beginning of the change period and q denotes the end of the change. The means of the sequence data before p and after q should be different if change occurs.

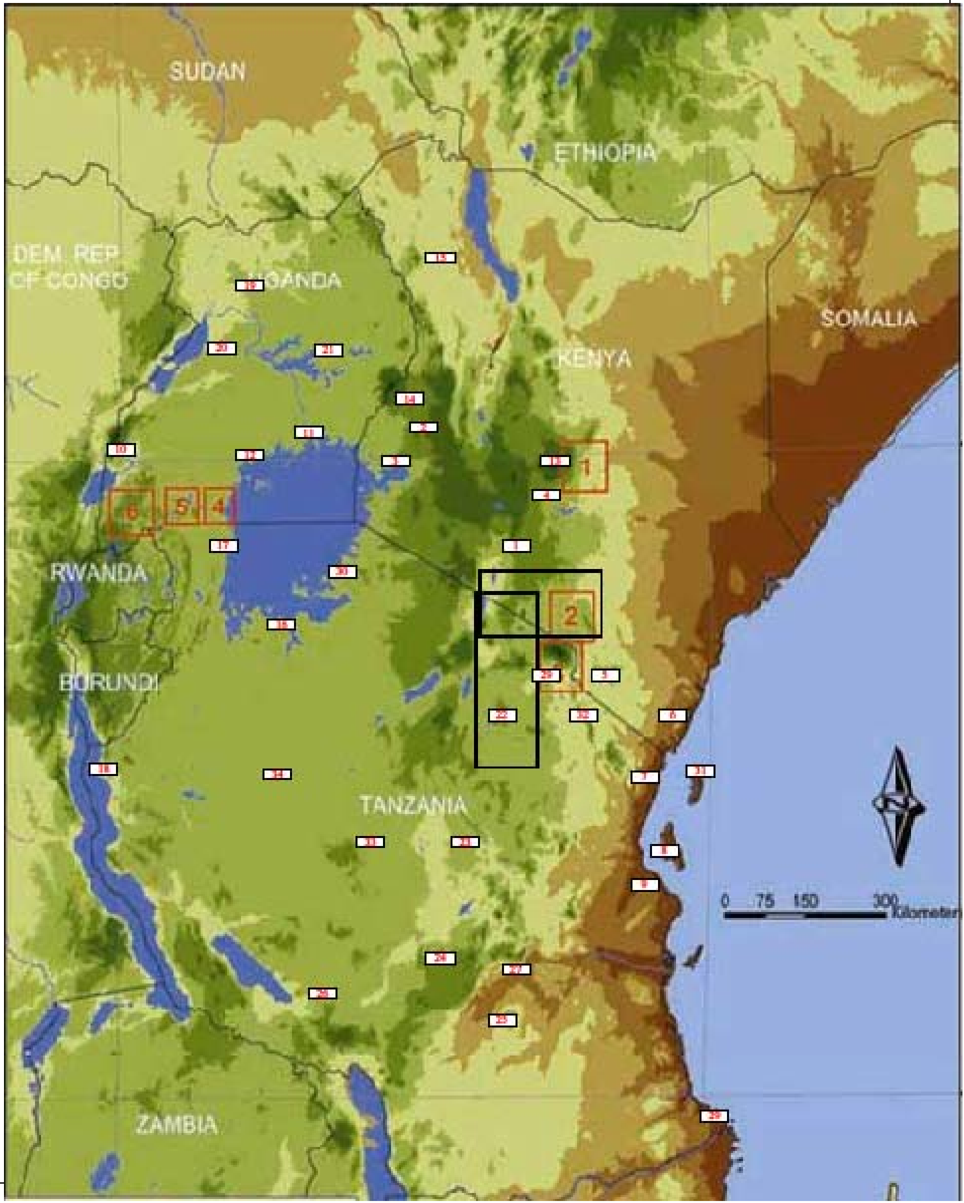
Virginia Tech

Change Point Test (cont'd)

- We use the test statistics T2 and T3 in Gupta's paper.
  - T2 is good at testing the change occurred in the middle of the sequence
  - T3 has a better power if change point fall at the end of the sequence.
- when we want to test Ho against Ha: $\delta>0$, i.e. the mean for inter-arrival time becomes longer after then change, name the statistics T2+ and T3+, for HA: $\delta<0$ situation, the shorter inter-arrival time case, name the statistics T2- and T3-.
- T_2+ or T_3+ signal->there is an upward mean change of the inter-arrival time, or the droughts occurred less frequently. T_2- or T_3- signal->the droughts occurred more frequently.

Data used for analysis

Fig3: locations of the 42 regions



The data we use in our study comes from 42 regions in East Africa (mostly Tanzania) time period mainly from 1950~2006.

After the assumption examination, we can safely treat our data as a sequence of indep exponential random variables and apply Gupta's statistics to solve our problem.

Test Results

- After computing T_2+, T_2-, T_3+ and T_3- for each region's moderate drought inter-arrival times, we found that 7 of the 42 regions have at least one of the four test statistics signaled.

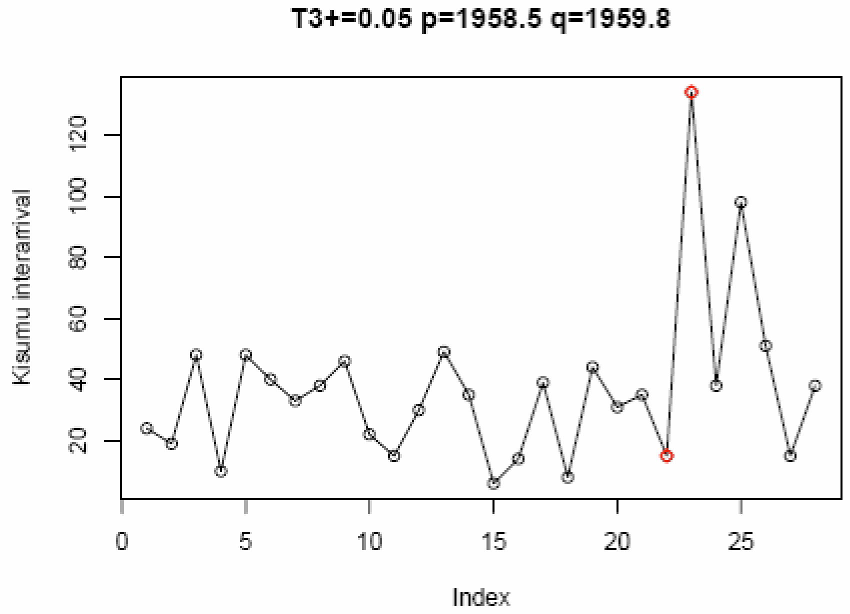
Table2: Significance level of the test statistics and estimations of p, q (yr.month)

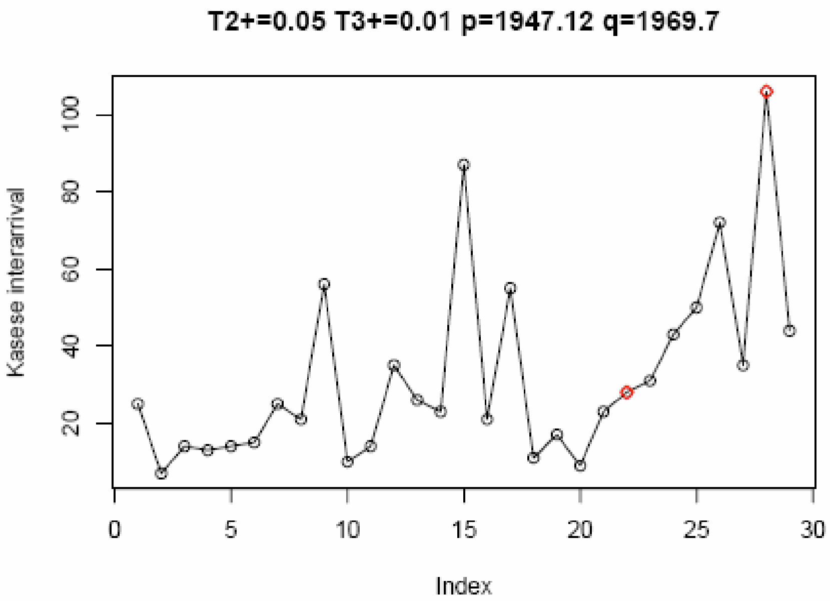
	test for δ>0, or drought freq decrease				test for δ<0, or drought freq increase			
site	T2+	T3+	p	q	T2-	T3-	p	q
Amani	0.05	0.00	1981.07	1983.04	0.00	0.00	NA	NA
Kasese	0.05	0.01	1947.12	1969.07	0.00	0.00	NA	NA
Kia	0.00	0.00	NA	NA	0.05	0.00	1983.05	1990.10
Kisumu	0.00	0.05	1958.05	1959.08	0.00	0.00	NA	NA
Mahenge	0.05	0.05	1981.11	1982.04	0.00	0.00	NA	NA
Musoma	0.05	0.00	1993.12	1996.04	0.00	0.00	NA	NA
Soroti	0.05	0.01	1946.02	1949.12	0.00	0.00	NA	NA

Test Results (cont'd)

Fig4: Time series plots of inter-arrival times for signaled regions

In the plot for Kisumu, there is an abrupt change before p and after q, which means the drought frequency becomes shorter.

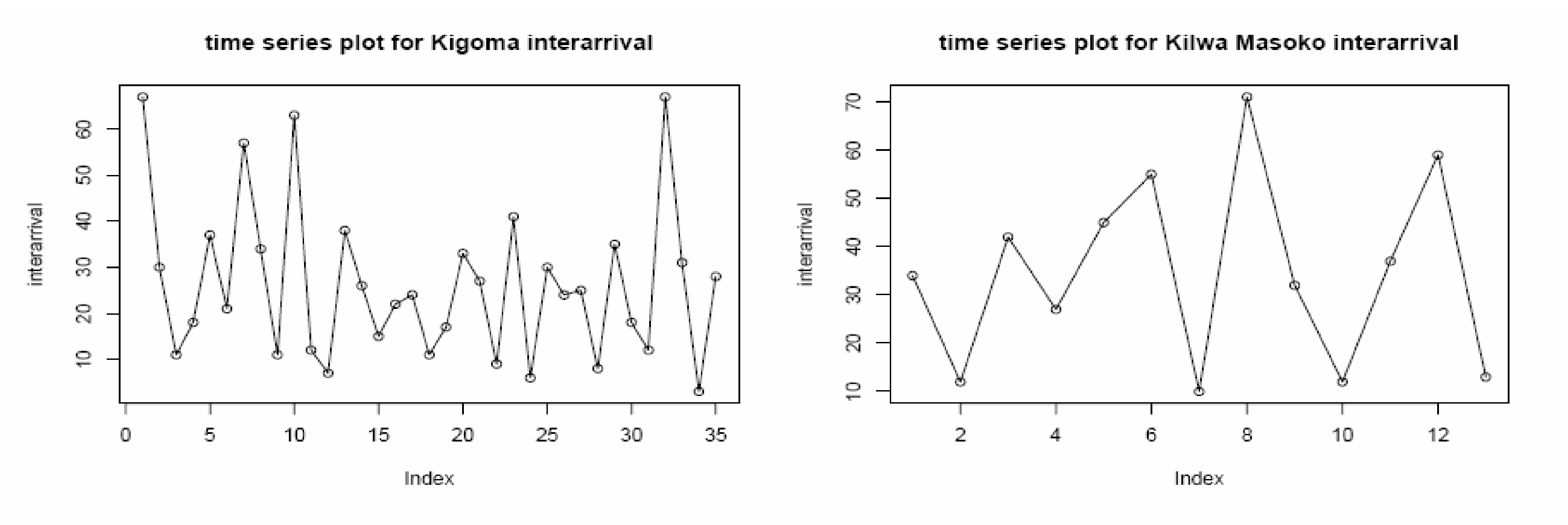




In the plot for Kasese, there is a linear change to upward gradually, so several points lie between p and q.

Test Results (cont'd)

Fig5: Time series plots for 2 non-signal regions



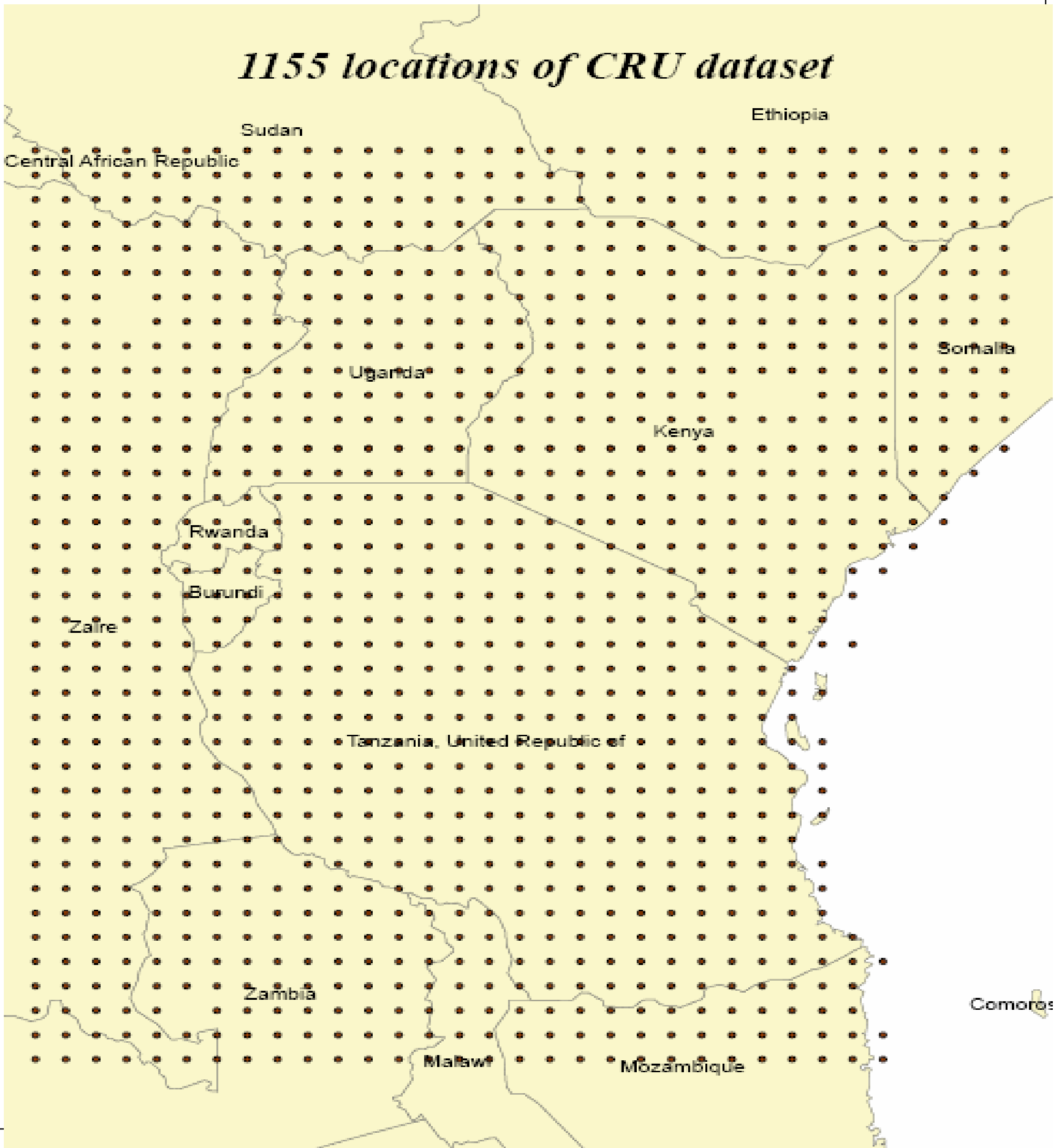
We could not tell any kind of change or trend from these two plots. Points randomly scattered around a center line which also consistent with the result of out test statistics: none of the four test statistics signaled.

Test statistics works quite well.

Climate Research Unit (CRU) data

The CRU data consisted of 1155 locations covering the area: longitude 26.8E~42.8E latitude 12.8S~5.8N and time period from 1950 to 2000.

Fig6: Locations of the 1155 CRU dataset

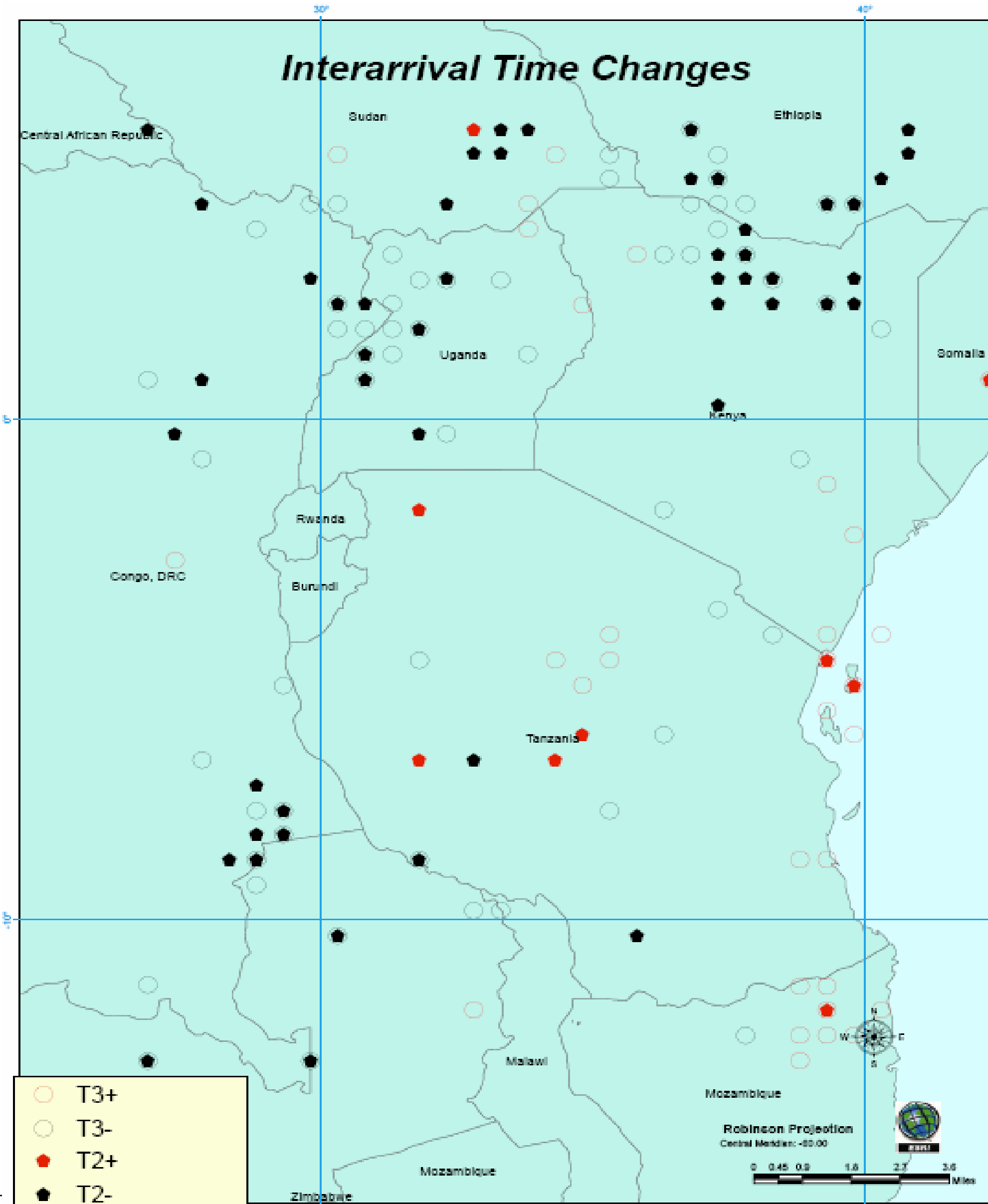


Test Results for CRU data

We calculate T2+, T3+, T2- and T3- for each of the 1155 regions, 128 regions' statistics signaled.

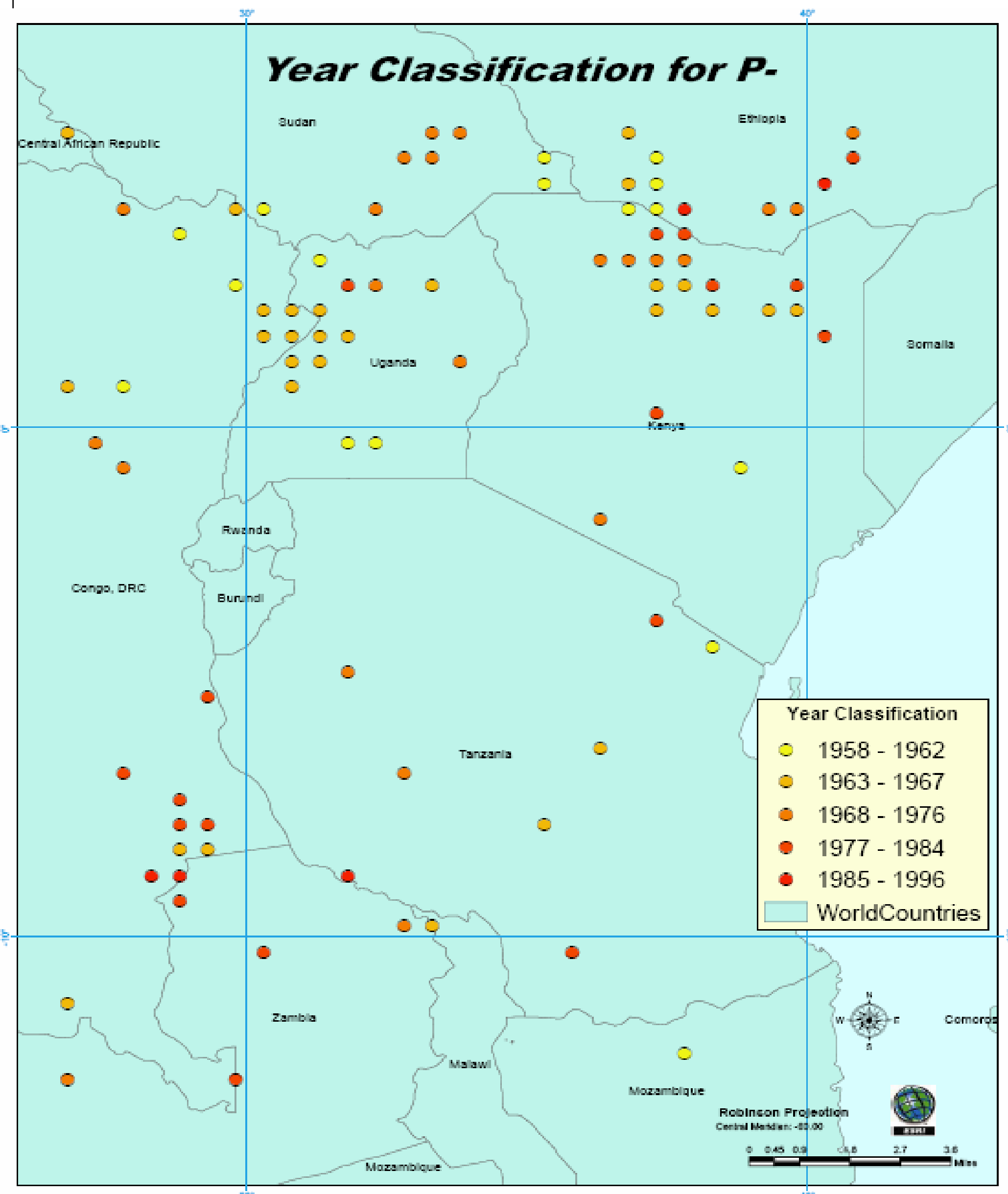
Red color indicates inter-arrival times changed to larger value, or equivalently, drought frequency became less. Black is equivalent to a more frequent drought change. Regions where has experienced a more frequent drought change are located: North Kenya, Northwest Uganda, South Ethiopia, North and Southeast Congo. Droughts become less frequent in the center Tanzania and along the coastal area.

Fig7: Locations of the 128 signaled regions



Test Results for CRU data (cont'd)

Fig8: Year classification of p-



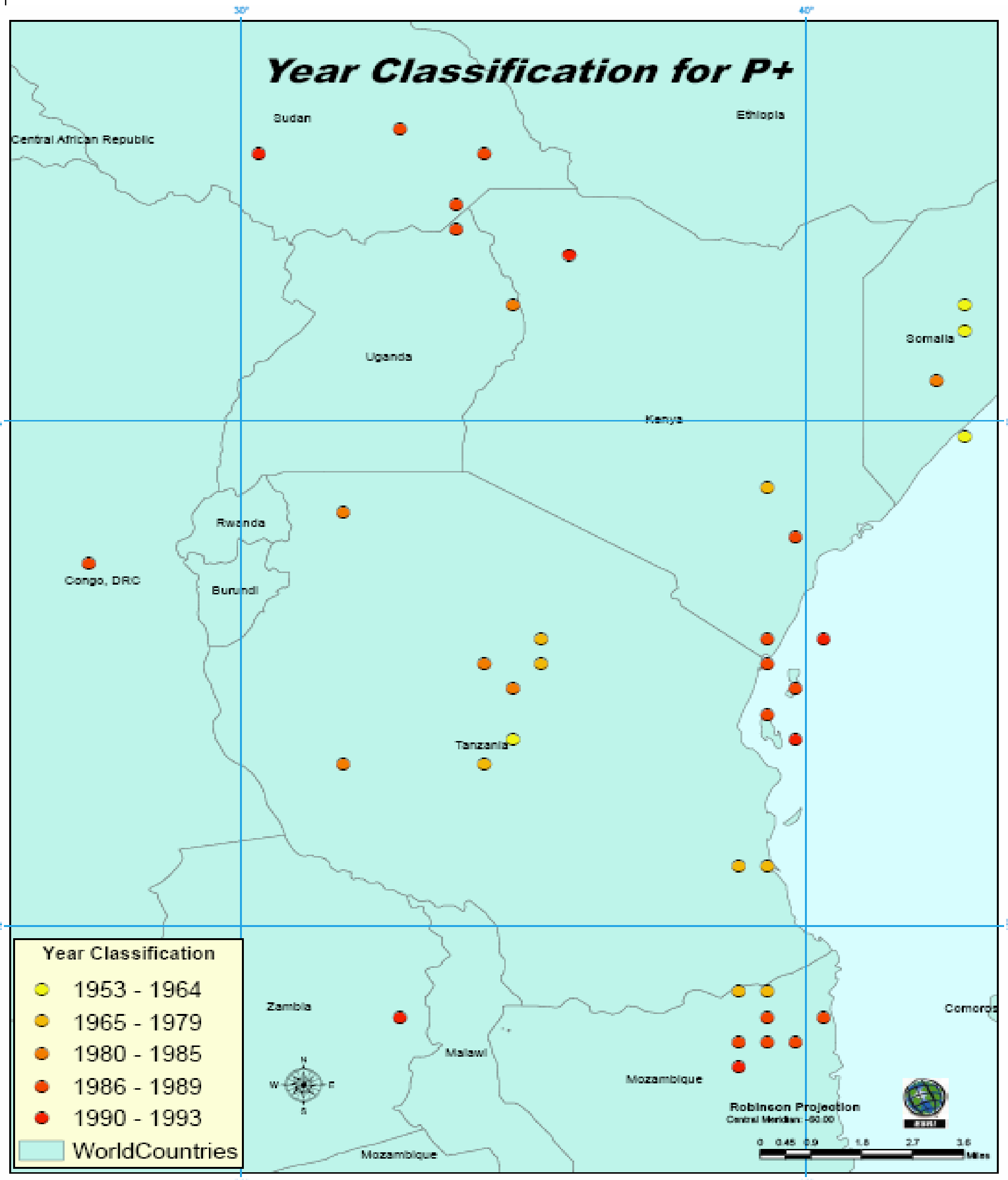
For the p's associated with T2+ or T3+, we call them p+, the change time of the less drought frequency. Then p- is the change time of the more drought frequency.

Droughts become less frequent in the center Tanzania (back to the 1950's) and along the coastal area (since 1980's or 1990's).

Virginia Tech

Test Results for CRU data (cont'd)

Fig9: Year classification of p+



North Kenya and Southeast Congo (happened around 1970's or 1980's) Northwest Uganda and Northeast Congo (around early 1960's or late 1950's) South Ethiopia (change time varied), all experienced a more frequent drought change.

Virginia Tech

Drought Frequency Change and ENSO

- Many studies have found some association between East African rainfall anomalies and ENSO. (e.g. Lindesay et al., 1986; Nicholson and Entekabi, 1986; Ogallo, 1988; van Heerden et al., 1988; Ropeleweski and Halpert, 1996; Nicholson and Kim, 1997; Mutai et al., 1998).

- Particularly, Ward (1998) suggested that ENSO's influence is mainly on higher rainfall frequency fluctuations.

- We use the formal statistics procedure to test whether ENSO and drought change associated with each other.

Drought Freq Change and ENSO (cont'd)

- An interesting phenomenon we discovered is that many of the p+ or p- change years are happen to be ENSO years.

Table3: Major El Nin˜o/Southern Oscillation Events Occurring within 1950–1996 (Ogall, 1979)

El Nin~o	La Nin~a
1951, 1953, 1957, 1963, 1965, 1969, 1972, 1976, 1982, 1986, 1991	1950,1955, 1964, 1970, 1973,1975, 1988

Virginia Tech

Drought Freq Change and ENSO (cont'd)

Table4: Consistency between p+ or p- change year and ENSO year

	P+ year	Also El Nin~o	Also La Nin~a	Also ENSO
Total Number	39	13	14	13+14=27
Percentage	100%	13/39=33%	14/39=36%	27/39=69%

	P- year	Also El Nin~o	Also La Nin~a	Also ENSO
Total Number	89	45	11	45+11=56
Percentage	100%	45/89=51%	11/89=12%	56/89=63%

To see whether the change years are really related to ENSO, we use the Goodness of Fit test to test whether the dist of the change years is influenced by the three categories: El Nin~o year, La Nin~a year and normal year. We have 99.5% confidence to say that the P+ and P- change years are influenced by ENSO.

Conclusions

- the drought frequency change point problem be answered by accurate statistical estimation. Also the change time can be estimated.
- the test statistics worked really well in the real example.
- regional drought change characters were analyzed for the larger East Africa area.
- showed a certain degree of association between the drought frequency change and ENSO based on the statistical testing procedure.

References

- Gupta A.K. and Ramanayake A. (2001): Change points with linear trend for the exponential distribution. J. Statist. Plann. Inference, 93, 181–195.
- Hulme M. (1992): Rainfall changes in Africa 1931-1960 to 1961-1990. Int. J. Climatol., 12, 685–699.
- Hulme, M., Doherty, R., Ngara, T., New, M., and Lister, D. (2001): African climate change: 1900–2100. Climate Res., 17, 145–168.
- Keim, B.D., Cruise, J.F. (1998): A technique to measure trends in the frequency of discrete random events. J. Climate, 11, 848– 855.
- Lindesay, J. A., Harrison, M. S. J. and Haffner, M. P. (1986): The Southern Oscillation and South African rainfall, S. Afr. J. Sci., 82, 196–198.
- Nicholson, S.E. and Entekhabi, D. (1986): The quasi-periodic behavior of rainfall variability in Africa and its relationship to the Southern Oscillation. Archives for Meteorology, Geophysics and Bioclimatology, Ser. A., 34, 311-348
- Nicholson, S. E. (1993): An overview of African rainfall fluctuations of the last decade. J. Climate, 6, 1463–1466.
- Ogallo, L. (1979): Rainfall variability in Africa. Monthly Weather Rev., 107, 1133–1139.

Virginia Tech

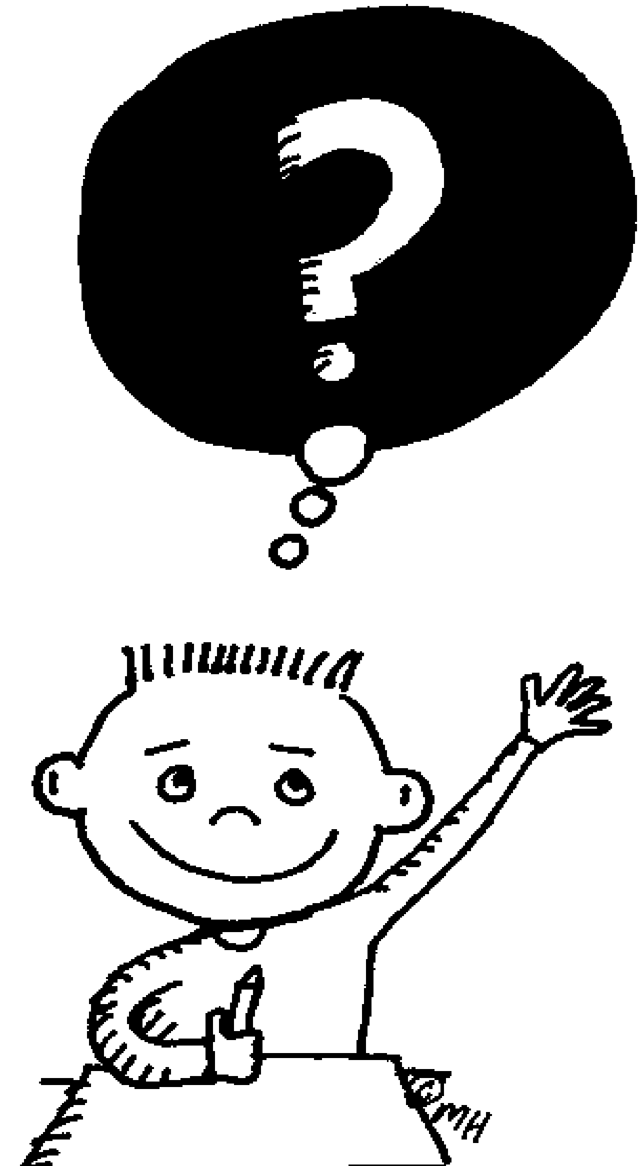
Acknowledgements

- Funding for this project was provided by NSF-CNH (award # NSF-0709671)

- Project title: Dynamic Interactions among people, livestock and savanna ecosystems under climate change (EACLIPSE)

- Precipitation data from Tanzania and Kenya Meteorological Department and Office

Virginia Tech


MH
QUESTIONS


THANK
YOU