Monthly Prediction of the Vegetation Condition: A case for Ethiopia

Getachew Berhan
Addis Ababa University, Ethiopia
getachewbl@yahoo.com

Tsegaye Tadesse
National Drought Mitigation Center
University of Nebraska-Lincoln, USA
ttadesse2@unl.edu
Introduction

- Drought is one of the most important challenges facing the planet.
- Nearly 50% of the world’s most populated areas are highly vulnerable to drought (UNEP, 2006).
- Frequent and severe droughts have become one of the most natural disasters in sub-Saharan Africa resulting in:
  - Serious economic crisis,
  - Social crisis,
  - Environmental crisis.
- About 180 million people in Africa live in drought-prone areas, and 50 million people are threatened with starvation in case of rain failure.
We need data to construct information

Plenty of data sources these days

When data increase—may hide pattern and to take action (hide relevant information)

Information is the start-up to take action

We need relevant information to take knowledge based decision.

What is knowledge?
Knowledge: a justified true belief, and it is created and organized by the flow of information (Nonaka, 1994).

This research is on explicit knowledge

- With the patterns observed from data,
- Where this pattern can easily be understood by humans and validated by test data with some degree of certainty (Han and Kamber, 2006).
Objective

The objective of this research is to develop a monthly drought monitoring approach using data mining and knowledge discovery from the database approach.
Materials and Methods

Study Area and Sample Size

Covering the whole country Ethiopia, a total of 2812 sample points
Research Framework

Materials and Methods...
Relevant Attributes Selection

Possible sources of qualitative evidence:

- Archival records
- Direct observation

Three criteria were used for selecting the attributes

1. Relevance for agricultural drought
2. Availability of the data for modeling (cost)
3. Statistical criteria
<table>
<thead>
<tr>
<th>Selected Attribute (Acronyms)</th>
<th>Format</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDNDVI</td>
<td>Raster</td>
<td>NOAA AVHRR</td>
</tr>
<tr>
<td>DEM</td>
<td>Raster</td>
<td>USGS</td>
</tr>
<tr>
<td>WHC</td>
<td>Raster</td>
<td>USGS</td>
</tr>
<tr>
<td>veg_Ethiopia</td>
<td>Vector</td>
<td>Ecodiv.org</td>
</tr>
<tr>
<td>Landcover</td>
<td>Raster</td>
<td>ESA</td>
</tr>
<tr>
<td>SPI_3month</td>
<td>Raster</td>
<td>IRI</td>
</tr>
<tr>
<td>PDO (Pacific Decadal Oscillation)</td>
<td>Point data</td>
<td>NOAA</td>
</tr>
<tr>
<td>AMO (Atlantic Multi-decadal Oscillation Index)</td>
<td>Point data</td>
<td>NOAA</td>
</tr>
<tr>
<td>NAO (North Atlantic Oscillation)</td>
<td>Point data</td>
<td>NOAA</td>
</tr>
<tr>
<td>PNA (Pacific North American Index)</td>
<td>Point data</td>
<td>NOAA</td>
</tr>
<tr>
<td>MEI (Multivariate ENSO Index)</td>
<td>Point data</td>
<td>NOAA</td>
</tr>
</tbody>
</table>
Modeling
Materials and Methods...
Assumption = possible to model drought in space and time dimensions using its attributes

Dependent variable here is SDNDVI

\[ NDVI = \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}} \]
Materials and Methods...
A 24 years of data (1983 – 2006) of the 11 key attributes were used for drought modeling experiment.

The years 1983 – 2006 were used because complete dataset for all the 11 attributes were obtained in these time periods for the whole modeling exercise for developing the knowledge base.

During the modeling experiment 10 – 30 rules were produced.
From the experimental datasets, 80% were used for training and 20% for testing the models (Gopal et al., 1999).

Using the 24 years historical datasets of the 11 attributes, a total of 10 models were developed for predicting drought in one to four month time lags for the growing season of June-October.
Regression Tree models

Rule 1: [124 cases, mean -59.0, range -251 to 163, est. err 62.0]
if
3 Month SPI <= -20
SDNDVI <= 0
Land cover in 14 (rain fed croplands), 20 (mosaic cropland (50-70%) / vegetation (grassland/shrub land/forest)
then
DroughtObject = 57.6 + 0.83 3 Month SPI - 0.023 DEM + 0.15 SDNDVI
### Results and Discussions

#### Summary of validation results for June-October regression tree models

<table>
<thead>
<tr>
<th>Months</th>
<th>Evaluation on test data</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Average error</td>
<td>Relative error</td>
</tr>
<tr>
<td>June one month</td>
<td>37.5</td>
<td>0.49</td>
<td>0.85</td>
</tr>
<tr>
<td>June two month</td>
<td>46.488</td>
<td>0.61</td>
<td>0.77</td>
</tr>
<tr>
<td>June three month</td>
<td>51.748</td>
<td>0.67</td>
<td>0.71</td>
</tr>
<tr>
<td>June four month</td>
<td>179.753</td>
<td>0.57</td>
<td>0.77</td>
</tr>
<tr>
<td>July one month</td>
<td>27.255</td>
<td>0.36</td>
<td>0.92</td>
</tr>
<tr>
<td>July two month</td>
<td>39.691</td>
<td>0.51</td>
<td>0.84</td>
</tr>
<tr>
<td>July three month</td>
<td>189.925</td>
<td>0.59</td>
<td>0.75</td>
</tr>
<tr>
<td>August one month</td>
<td>22.184</td>
<td>0.29</td>
<td>0.95</td>
</tr>
<tr>
<td>August two month</td>
<td>186.404</td>
<td>0.58</td>
<td>0.75</td>
</tr>
<tr>
<td>September one month</td>
<td>180.224</td>
<td>0.57</td>
<td>0.77</td>
</tr>
</tbody>
</table>
August one month

- Observed Values
- Model Predicted Value
Model Implementations

- The pattern observed corresponds to the accuracy levels obtained.
- August one month prediction was the highest accuracy and also with best pattern.
- As the prediction month length increase the pattern observed decreased.
Results and Discussions...
August One Month Drought Prediction Map for 1984

Legend
- Zonal District Boundary
- Extreme Drought
- Severe Drought
- Moderate Drought
- Near Normal
- Moist
- Very Moist
- Extreme Moist
Results and Discussions...
The errors in the models are attributed to:

- low spatial and temporal resolution of input datasets
- Heterogeneous ecosystems of Ethiopia
Model Evaluation

Method

- The model evaluation here was done in the context of “fitness for purpose”.
- The evaluation of the drought model product was done using the “Meher” season (long rainy season) yield data.
- 41 “Meher” growing Zones were used for this evaluation.
- Two criteria were used for selecting Zones:
  - Availability of data from 2000 - 2006 Meher season
  - Zones are in “Meher” crop growing districts

Results and Discussions...
Results and Discussions...

Figure 41: Drought Object August one month prediction maps for the year: (a) 2000, (b) 2001, (c) 2002, (d) 2003, (e) 2004, (f) 2005, and (g) 2006.
Results and Discussions.

- Zonal Map of Ethiopia
- DroughtObject Image
- Zonal Output Attributes From DroughtObject Image for each Zone
Results and Discussions...
The model evaluation showed the reliability of model products.

The errors in correlating model output with yield data is mainly attributed to coarse administrative yield data.
6. Conclusions

- The research has shown that drought can be effectively identified and monitored using the developed.

- The new system enables drought prediction better, faster and cheaper.

- Compared to previous work, ten more weeks can be predicted in advance with the new approach (i.e. a total of 4 months in advance drought prediction is possible).

- The problem of meteorological point data collection and analysis at coarse resolutions level were solved with this new approach and the coverage and time delay issues are addressed.

- Decision makers at different levels can use the new system to plan proactively for potential food insecurity.

- Drought planning can be made at risk-based drought management approach rather than a crisis-based approach.
Contributions

In terms of Metrics

- **Effectiveness**: Model evaluation $R^2$ up to 0.91
  - Up to 4 months in advance prediction

- **Efficiency**: Low cost attributes used

- **Applicability**: Proactive planning is possible
Future Research

Three key future researches proposed

1. Design and develop *droughtOutlook* system

2. Develop database and Data warehouse system

3. Develop an Empirical Drought Early Warning and Agile Insurance System
Future Research – having the working system in place
Future Research...
The Big Picture

- Support Ground Stations (SGSs)
- Satellite Application Facilities (SAFs)
- Back-up & Ranging Ground Station (BRGS)
- Primary Ground Station (PGS)
- Data from other meteorological satellites
- Processed images and meteorological products
- Satellite control (back-up)
- Raw images & DCP data
- Processed images & meteorological products
- Satellite control (primary)
- High Rate User Stations (HRUS)
- Low Rate User Stations (LRUS)
- Aircraft Data Collection
- Ship Data Collection
- Weather Balloon Data Collection
- Data Collection Platforms (DCP)

Control & Processing Centre
Darmstadt, Germany
THANK YOU