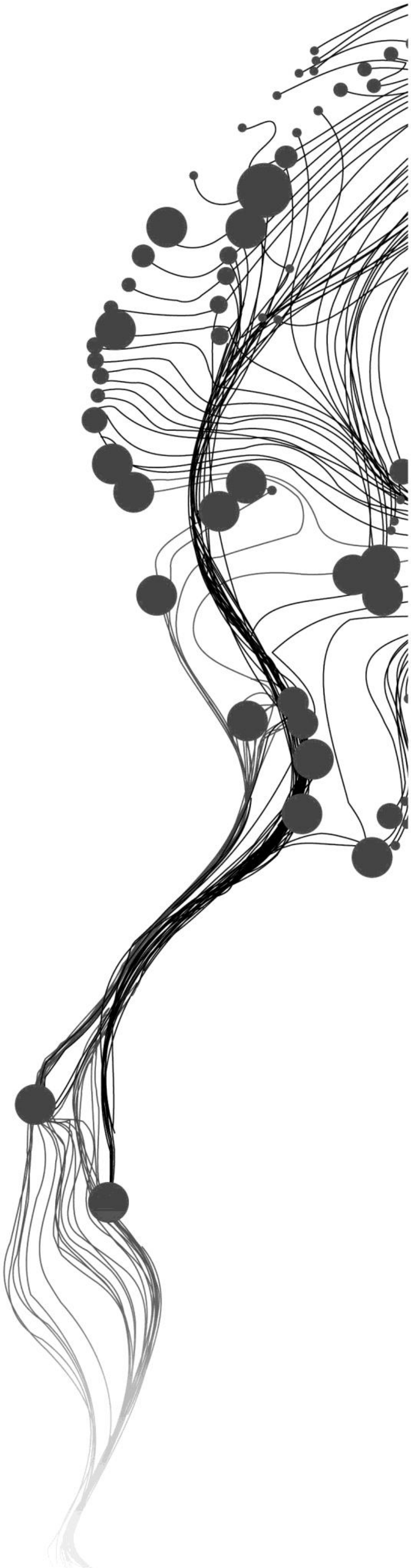
PREDICTING DYNAMICS OF VEGETATIVE DROUGHT CLASSES USING FUZZY MARKOV CHAINS

SIQI DING
March, 2011

SUPERVISORS:
Prof. Dr. A. Stein
Ms. Dr. W. Bijker
Ms. C. Rulinda MSc



PREDICTING DYNAMICS OF VEGETATIVE DROUGHT CLASSES USING FUZZY MARKOV CHAINS

SIQI DING
Enschede, The Netherlands, March, 2011

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Geo-informatics

SUPERVISORS:
Prof. Dr. A. Stein
Ms. Dr. W. Bijker
Ms. C. Rulinda MSc

THESIS ASSESSMENT BOARD:
Prof. Dr. A. Stein (Chair)
Dr. Ir. S. de Bruin (External Examiner)
Ms. Dr. W. Bijker (Member)
Ms. C. Rulinda (Member)

ABSTRACT

Drought is a naturally occurring event, causing temporary imbalance water availability and vegetation damage. It exists when the amount of received precipitation has been significantly below normal recorded levels. To reduce the devastating effects of drought and minimize the losses, early warning system can help decision and policy makers to implement policies timely. Satellite-based Normalized Difference Vegetation Index (NDVI) data are consistently available and continuous in space and time, applied to prediction drought in this research.

As drought has many characteristics, varying region by region and may last for several months or even years, it represents a challenge to fully evaluate the characteristics for its prediction. In this research, the main objective is to predict the changes of vegetative drought classes, also called states. This is done by modelling these changes using Markov chains applied to predefined fuzzy vegetative drought states.

Four regions of different agricultural patterns in Kenya are selected as study areas to apply this method. There is a strong relationship between NDVI and accumulated almost three-month precipitation data. The highest correlation value R can be larger than 0.9. It indicates that NDVI can be an indicator for vegetative drought prediction. The every dekadal NDVI data are acquired from FEWS NET from 2004 to 2008. Fuzzy membership functions are applied in this research as a description of drought classes. The vegetative drought classes are classified by fuzzy classification.

Under the Markovian property tests, the NDVI anomaly data can be modelled in first-order Markov chain, but the time homogeneity is interrupted by the data in February and September. The validation data is the comparison of prediction result and pre-existing reference data. Half of the pixels in study area are well predicted by fuzzy Markov chain. And also, the changing of fuzzy membership function influences the result of prediction. In conclusion, the Markov chain with fuzzy membership function has the potential to be applied in vegetative drought prediction and provide benefit for early warning system.

Key words: NDVI, vegetative drought, fuzzy, Markov chain

ACKNOWLEDGEMENTS

I would like to express my earnest gratitude and thankfulness to my supervisors, Prof. Dr. Ir. A. Stein and Dr. Ir. Wietske Bijker for their constant support, timely guidance, encouragement, valuable advices and suggestions throughout the research work.

My special thanks to my advisor, Ms. C. Rulinda, PhD candidate in the department of Earth Observation Science. Her encouragement, enthusiastic spirit, constant support, valuable guidance and understanding through out the research work which has led to the completion of this thesis. Her deep interest and dedication towards the improvement of my research will always be remembered. We had a remarkable time together to solve obstacles.

I take this opportunity to appreciate my dear cluster-mates: Diego, Namhyun, Aminah. We spent most of our time together in the past 5 months sharing joy and tears, relax and stress. There are lots of fun stories and complaint among us, which remains in my memory. Also I want to give best wishes and appreciation to my best friend Nazani, who always supports me and deserves all the best.

I would like to thank all my dear friends: Bernard, Bormy, Ervin, Fengfan, Hsiang Hsu, Fathi, Ivar, John, Natia, Mesele, Pankaj, Ms. Tina Tian, Haining, Dongdong, Liu Xin, Sheng Tian, Hongjie, Xiaoxiao and Jia for supporting and praying for me. Great thanks to my housemates: Frank, who always introduces nice music to me; Mylene and Janine for making delicious food and girl's talk.

I am extremely grateful to Ms. Weiwei Fu for delicious Chinese food, undoubted support and encouragement through out my research period. She is becoming part of my life during the thesis time, even providing a room for me in her place.

I want to express my thanks to Dr. D. G. Rossiter, my favourite lecturer and American student in ITC. Actually I learned a lot from him during every Chinese class, though I am the Chinese teacher. It says, "A friend in need is a friend indeed". Whenever I feel excited or confused, he is the one first came to my mind that I want to share the story with. Thanks for being there with me.

Last but not least, I want to express my heartfelt gratitude to my dear parents, who are always supporting me, loving me and respecting every decision I made. With the parents' support, I am encouraged to overcome the difficulties and enjoy my life in Netherlands. I cannot express how much you all mean to me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Drought

Drought is a naturally occurring event, causing temporary imbalance water availability and vegetation damage. It exists when the amount of received precipitation has been significantly below normal recorded levels over an extended period of time, such as several weeks, months or years. As most vegetation is dependent on rainfall, drought has negative effects on natural vegetation and agricultural productivity. When drought occurs, it also leads to diminished power generation, disturbed riparian habitats, and suspended recreation activities, which are closely linked to livelihoods of citizens and the national economy (A. K. Mishra & Singh, 2010). Because drought is an integral part of climatic variation, it can occur in all climatic regimes no matter how is the mean annual rainfall. Contrast to aridity, a permanent feature of the climate and restricted to low rainfall areas, drought is a temporary aberration (D. A. Wilhite & Svoboda, 2000).

Generally, drought itself does not kill the most among all the natural hazards. However, because of drought and lack of alternatives to water plants, vegetation in form of pastures for livestock and crops wilt (Speranza, Kiteme, & Wiesmann, 2008). This can cause crop loss, reduced harvest, feeds for livestock, and in severe cases livestock and human deaths. Therefore, a large number of people die from food-shortage and famine, which are the most serious outcomes of drought (Smith, 2009). In North Carolina, U.S.A., the economic loss incurred by the drought of the year 2002 was estimated as $413-148 million for both agriculture and municipalities, while $610 million for agriculture and timber growth loss in South Carolina (Rhee, Im, & Carbone, 2010). The Australian Federal government spent $740 million in aid during the 2002-2005 "Big Dry". The overgrazing and poor cropping methods, deforestation and improper soil conservation techniques may not create drought but they amplify drought-related disaster. During the 1980s and early 1990s, the agricultural drought which affected many countries and people in Africa, were probably some of the most shocking famine emergencies in recent history, causing a 54% reduction in cereal harvest and exposing more than 17 million people in risk of starvation(Rojas, Vrieling, & Rembold, 2011). The entire Horn of Africa has been prone to dry periods over the recent decades. A prolonged drought has crippled agriculture production in rural Kenya from 2009, greatly affecting millions of families who earn their lives on farming, fishing or herding. Nowadays, in Kenya, the drought has forced herder to be away from home in search for water and food supplies for their animals, and as such, leave women, children and elderly to fend for themselves. The droughts cause also threat of violence such as conflict about pastoral territories and education issues of migration children. An estimation of 100,000 cattle have died and the government estimates 10 million people will be affected by food shortages and diseases from drinking dirty water in the present drought.

In theory, the artificial stimulation of rainfall by cloud seeding can reduce the hazard but with the limitation that the clouds should have natural precipitation potential. Unfortunately, such clouds are hardly to be present in large number during drought conditions. Also the additional supply of water is not always the best solution for drought. Drilling of new boreholes in dry areas without proper local management cannot reduce the effect of disaster. Inappropriate irrigation water consuming agricultural productivity will do nothing to alleviate food shortages during drought(Smith, 2009).

To reduce the devastating effects of drought and minimize the losses, preparedness, prediction and early warning system can help decision and policy makers to implement policies timely. With several weeks' lead-time warning, farmers can make decisions of altering agricultural systems, for example reducing high water consumptive crops, to cope with drought. Therefore, to develop practical methods of forecasting drought some weeks or even months ahead should be considerable (Paulo & Pereira, 2007; Pereira, Cordery, & Iacovides, 2009). The challenge of implementing emergency response, such as early warning system, is to know the probability of drought occurrence and measure the phenomena of drought (Rojas, et al., 2011). The phenomenon of drought represents relative instead of absolute variation from norm and is identified on the basis of human impacts rather than physical causes.

Conventional climate-based drought monitoring from weather ground station network is limited by its density and distribution, and difficult to obtain near-real time data. Although the spatial interpolation can provide valuable information, high uncertainties may exist by many factors during interpolation process(Rhee, et al., 2010). Satellite-based data are consistently available and continuous in space and time. As vegetation has direct response to precipitation deficiency, satellite-based observations have proven to be efficient for detecting vegetation dynamic conditions in large coverage and multi-temporal measurements.

1.2. Research Statement

As drought has many characteristics, varying region by region and may last for several months or even years, it represents a challenge to fully evaluate the characteristics for its prediction. In this research, instead of using climate-based dataset, such as precipitation and temperature, Normalized Difference Vegetation Index (NDVI) satellite-based data are used to model vegetation stress, considered as vegetative drought. A Markov chain approach is used to predict the probability of occurrence of vegetative drought.

1.2.1. Research objective

The main objective of the research is to predict the changes of vegetative drought classes, also called states, using the NDVI values. This is done by modelling these changes using Markov chains applied to predefined fuzzy vegetative drought states. The performance of the method is evaluated against pre-existing reference data.

1.2.2. Research questions

The following research questions need to be address in order to achieve the objective:

- What is vegetative drought and how can it be characterised?
- How to determine and model the vegetative drought states?
- How to model the change between vegetative drought states?
- How to evaluate the prediction result?
- How the number and value range of classes can affect the results?

1.2.3. Innovation aim

Commonly, the definition of natural phenomenon for each drought state is crisp in Markov chains. In reality, the transition between vegetative drought classes is a gradual, hence a crisp model is less realistic. In this study, we model these changes using Markov chains applied to fuzzy models of drought states, referred to as *Fuzzy Markov Chains*.

1.3. Overview of research

The overall method used to conduct this study is as follows:

- The research starts with a literature review on drought prediction, followed by a selection of a study area, as drought characteristics can differ from regions.
- Within the study area, study sites are selected to represent the major land use types of interest and a dataset is selected accordingly.
- Vegetative drought classes are determined and modeled in a similar way as Rulinda et al. (2010), accounting for the gradual transition between the classes.
- A transition matrix is built to model the changes between the vegetative drought classes which are then predicted for the next time step.
- To assess the validity of using vegetation stress as a proxy of vegetative drought, a combination of rainfall data and drought reports are used,
- And to assess the accuracy of the prediction, classified vegetative drought maps corresponding to the predicted time step are used as reference data.

1.4. Structure of the thesis

The rest of the thesis is structured as follows: Chapter 2 provides with a literature review of drought prediction using Markov chains. Chapter 3 describes the study area and chapter 4 describes the dataset used in this study. Chapter 5 details the methodology used and in chapter 6 the results are presented and discussed. In the last chapter, a conclusion is given together with some recommendation for further study in this field.

2. LITERATURE REVIEW

2.1. Understanding Drought

As a natural hazard, drought differs from other natural hazards in several ways. First, it is difficult to determine the onset and the end of drought, because its impact accumulate slowly and may linger for years after the termination. Hence, drought is often referred to as a creeping phenomenon. Second, due to the lack of a precise and universally accepted definition of drought, its existence and degree of severity are vague. In reality, drought occur in both dry and humid regions sometimes in the same country (WMO, 2006). Definition of drought must be considered as relative, regional and impact specific as it sometimes varies dramatically from one region to another. The impact of drought is less obvious and spreads over a larger geographical area than those of other natural hazards (D. A. Wilhite & Svoboda, 2000). Take agricultural drought for instance, the impact can be accurately assessed when crop are harvested, which is a few months after the symptoms began to appear (Boken, Cracknell, & Heathcote, 2005). Unlike earthquake or floods constrained to a particular tectonic and topographic setting, drought can extend over regions with potential of causing catastrophe and affect several neighbouring countries (Smith, 2009). These characteristics of drought have increased the difficulties of accurate, reliable and timely estimates of severity and impacts of drought for preparedness plans. However, with improving monitoring, early-warning system and appropriate mitigation actions, the impacts of future drought events can be efficiently reduced (D. Wilhite, 2005).

Drought has three essential features: intensity, duration and spatial coverage. Intensity refers to the severity of impacts caused by precipitation shortfall. It is generally measured by the departure of some climatic index for example Standardized Precipitation Index (SPI), from its long-term mean. Duration is that how long the drought event lasts. Usually drought requires a minimum of a month to become established and then continue for months and years. Another distinguishing characteristic is the spatial coverage. The areas affected by severe drought evolve gradually, and regions of maximum intensity shift from season to season and year to year in the event of a multi-year drought(D. A. Wilhite, 1992).

2.1.1. Types of drought

Drought is generally classified into four categories: meteorological drought, agricultural drought, hydrological drought, and socioeconomic drought. Meteorological drought is usually first detected and defined as a lack of precipitation for a period of time, in comparison to an average level for a specific region. Usually there is a deficiency threshold over a predetermined period time to define this drought, which will vary by location according to local water demands. The Australia Bureau of Meteorology employs a threshold of an exceed of 10 % of normal precipitation in the same period of the year, and the duration period is more than three months (Smith, 2009). Meteorological drought is often the focus of drought forecasting techniques for famine early warning system(Sene, 2010).

Agricultural drought is commonly defined by declining soil moisture and consequent crop failure with below-normal precipitation at the surface over a period of time. The impact of agricultural drought is complex. It depends not only on the magnitude, duration and timing of drought, but also on the responses of the region's soils, plants and animal to water stress(Nagarajan, 2009b). It occurs over a large area and its impact is not accurately assessed until a few months after the symptoms of agricultural drought begin to appear (Boken, et al., 2005). Agriculture is the first economic sector to be affected by drought because many subsistence economies rely mainly on rain-fed agriculture and soil moisture supplies are often quickly depleted. There is no significant relationship between precipitation and infiltration of precipitation into the soil. The soil water-holding capacity, differences between actual evapotranspiration and potential evapotranspiration, and soil water deficits can affect the soil moisture. The soil with low water-holding capacity is more prone to agricultural drought. A definition of agricultural drought should account for the various biological characteristics of crops at different stages of growth, because the plant's demand of water is different(WMO, 2006).

Hydrological drought is associated with a period of inadequate surface or subsurface water supply. Since it takes longer time to show precipitation deficiencies in components of the hydrological system, hydrological drought is generally out of phase or lags behind meteorological and agricultural drought. The recovery of a hydrological system after drought is slow because of the long recharge periods of the system's water supplies. Like agricultural drought, it is no directly related to precipitation amount and the status of hydrological system supplies in lakes, reservoirs, aquifers and streams, because water stored in the system often has multiple and competing purposes, such as irrigation, recreation, flood control and power generation (WMO, 2006). As a result, impacts are difficult to quantify. Stream-flow data have been widely used for hydrological drought analysis and geology is found to be one of the main influencing factors(A. K. Mishra & Singh, 2010).

Socioeconomic drought occurs when the demand for an economic good exceeds supply as a result of a weather-related shortfall in water supply(A. K. Mishra & Singh, 2010). It significantly differs from the other types of drought because it reflects the relationship between supply and demand, and emerges when meteorological, agricultural and hydrological drought adversely affects demand and supplies of economic goods. Demand also increases and is often associated with a positive trend as a result of increasing population, development and other factors(WMO, 2006).

The sequence and impacts between meteorological, agricultural and hydrological drought is illustrated in Figure 2-1. Agricultural, hydrological and socioeconomic drought occurs less frequently than meteorological drought, because these sectors are not directly related to precipitation deficiency. It usually takes several weeks to produce soil water deficiency leading to stress on crops, pastures and rangeland when precipitation deficiency begins. Continued dry conditions for several months bring reduced stream-flow, inflow to reservoirs, lakes, and ponds reduced wetlands. When drought conditions persist for a long period of time, economic, social and environmental impacts come out.

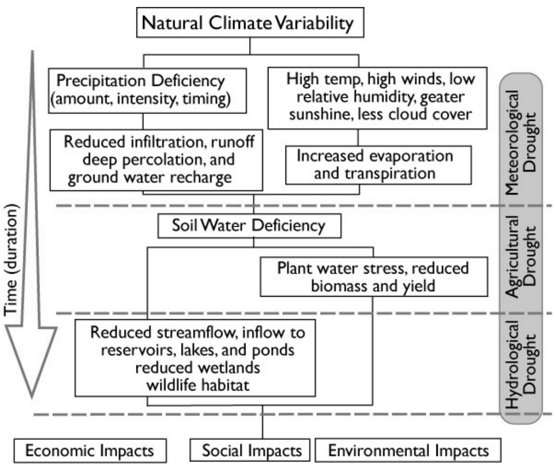


Figure 2-1 Sequence and impacts between Meteorological drought, Agricultural drought and Hydrological drought(NDMC, 2006)

In the past, the primary sources for drought information are climate and meteorological data. With the development of technology, satellite observations have recently proved to be a valuable source of timely and spatially continuous data for monitoring vegetation dynamics over large areas. One of the most commonly used vegetation indices, extracted from satellite images, is NDVI. It has the advantages of minimizing soil and other background effects, reducing data dimensionality, providing a degree of standardization for comparison, and enhancing the vegetation signal (Tadesse, Brown, & Hayes, 2005). In this study, the focus is on the drought impact on vegetation, observed from satellite images. This research does not address desertification and aridity, which are different from drought.

2.1.2. Drought indices

Rainfall, temperature, evaporation, vegetation health, soil moisture, stream flow, and so on are the parameters that are used in drought analysis for scientific evaluation of drought situations (Nagarajan, 2009a). In recent decades, a number of indicators or indices have been developed based on different parameters. Figure 2-2 shows the drought indices used in different examples. Precipitation based indices are often used.

Type	Name	Countries (examples)	Basis of method
Meteorological	Rainfall percentiles	Australia	Percentile values for cumulative rainfall for different durations (e.g. 3 or 6 months), sometimes expressed on a scale of 1–10 (i.e. deciles)
	Standardised Precipitation Index (SPI)	USA, China	Cumulative rainfall for a range of durations (e.g. from 3 to 48 months) compared with values from normalised rainfall probability distributions based on long term records (McKee et al. 1995), expressed in terms of the number of standard deviations from a mean of zero
	Palmer Drought Severity Index (and Palmer Hydrological Drought Severity Index)	USA	Conceptual water balance approach which compares the available water to the climatological normal, considering rainfall, evapotranspiration, soil recharge and runoff (Palmer 1965)
	Water Requirements Satisfaction Index	FEWS NET	The ratio of actual to potential cumulative crop evapotranspiration (Senay and Verdin 2003; see Box 8.3)
Soil moisture	Crop Moisture Index	USA	A component of the Palmer Drought Severity Index calculation focusing on short-term changes in soil moisture (Palmer 1968)
Hydrological	Low flow statistics	Widely used	Flow duration (e.g. Q_{95}), run-sum and other statistical measures of flow in terms of duration, values relative to a threshold etc
	Surface Water Supply Index	Western USA	Weighted sum of snowpack, streamflow, precipitation, and reservoir storage values normalized by non-exceedance frequencies (Shafer and Dezman 1982)
Agricultural	Aridity Index	India	Difference between actual and potential evaporation as a fraction of potential evaporation (World Meteorological Organisation 2009)
	Normalised Difference Vegetation Index (NDVI)	Widely used	Normalised difference between near-infrared and infrared reflectance values, giving a measure of vegetative cover

Figure 2-2 Examples of drought monitoring indicators and indices (Sene, 2010).

2.1.2.1. Drought Indices derived from Hydro-meteorological data

Percent of normal precipitation is one of the simplest index of drought, but it has limitations in defining the appropriate threshold and the distribution should be a normal distribution instead of skewed. Precipitation deciles are commonly used by Australian Bureau of Meteorology and overcome the limitations of percent of normal precipitation. Annual rainfall of a long series of years are sorted in order and classified into 5 equal groups as shown in Table 1. The current precipitation is then assigned to the historical classification (Cancelliere, Mauro, Bonaccorso, & Rossi, 2007). SPI is calculated from a monthly precipitation dataset of ideally a continuous period of at least 30 years. The normal distribution density of historical rainfall data is used to estimate the probability of any precipitation observation, which is with a mean of zero and transformed from a Gamma function(Mckee, Doesken, & Kleist, 1993). The Palmer Drought Severity Index (PDSI) is a soil moisture algorithm using historic records of precipitation and temperature data within a two-layer soil model. The PDSI perhaps is the most widely used index for region drought monitoring and it has been used to illustrate the areal extent and severity of various drought episodes(A. K. Mishra & Singh, 2010). A modified version known as Palmer Hydrological Drought Index (PHDI) is used to monitor long term moisture anomalies that affect water supplies and widely used in USA(Karl, 1986). The relationship between these three indices and drought categories is shown in the Table 2-1.

Deciles classification		SPI classification		PDSI classification	
Deciles	Drought category	SPI value	Drought category	PDSI values	Drought category
Lowest 20%	Much below normal	≤-2.0	Extremely dry	≤-4.00	Extremely dry
Next higher 20%	Below normal	-1.99 to -1.50	Severely dry	-3.99 to -3.00	Severely dry
Middle 20%	Near normal	-1.49 to -1.00	Moderately dry	-2.99 to -2.00	Moderately dry
Next higher 20%	Above normal	-0.99 to +0.99	Near normal	-1.99 to -1.00	Mild drought
Highest 20%	Much above normal	1.0 to 1.49	Moderately wet	-0.99 to -0.50	Incipient drought
		1.5 to 1.99	Very wet	-0.49 to 0.49	Near normal
		≥2.0	Extremely wet		

Table 2-1 Relationship between indices and drought categories

2.1.2.2. Drought indices derived from Remote Sensing data

Utility of remote sensing data especially satellite images have been proven in drought monitoring, mitigating and prediction. Satellite observations provide more spatially and timelier continuous input data sources then ground gauge station observations. Because of the characteristics of the spatial coverage of drought, satellite images enable to understand manifestation of drought in larger area in a less time consuming way than conventional methods. When a drought exists, due to reduction of precipitation, the capacity to carry out the chlorophyllian function on the part of the vegetation is notably reduced (Sharma, 2006). The response of the green vegetation is characterized by a maximum absorption radiation in the red region and large reflection in the neighbouring infrared region. It also has been observed that in unhealthy, ageing or subject to condition of vegetation stress, the reflectance in red region increases while in near-infrared region deceases. The normalized difference vegetation index (NDVI), developed by Tucker in 1979 is the most popular vegetation index used to monitor vegetation at regional to global scales. It is calculated as in equation 2-1.

$NDVI = (\lambda_{NIR} - \lambda_{RED}) / (\lambda_{NIR} + \lambda_{RED})$ (2-1)

Where λ_{NIR} and λ_{RED} are the reflectance in the NIR and Red band, the values vary in [-1,+1]

As NDVI is not sensitive to influences of soil background reflectance at low vegetation cover and lag vegetation response to precipitation deficient, NDVI itself does not reflect drought or non-drought conditions. But the severity of drought may be defined as NDVI anomaly from its long-term. The anomaly NDVI is calculated as the difference between the NDVI at current time step, such as month, and a long-term mean NDVI of the same time step for each pixel. When the $NDVI_{anomaly}$ is negative, it indicates the below-normal vegetation condition and maybe a drought situation. The larger the negative departure, the greater drought severity may be suggested. The formula is equation is in equation 2-2.

$NDVI_{anomaly} = NDVI_i - NDVI_{i,mean}$ (2-2)

where $NDVI_i$ is the value for time step *i*, $NDVI_{i,mean}$ is the long-term mean value of same time step *i*

The Vegetation condition index (VCI) was first introduced by Kogan to measure "health" in vegetation and shows promises in drought detection and tracking (F. N. Kogan, 1995). It is calculated with a long-term maximum and minimum record. The formula is shown is shown in equation 2-3, where *i* is the index of current month or week, $NDVI_{i,min}$ and $NDVI_{i,max}$ are the minimum and maximum values from record of time *i*. The onset, intensity, duration and impact of drought on vegetation can be detected by VCI. The limitation of VCI is that mainly useful in growing season as during the non-growing season, the vegetation is largely dormant (Heim, 2002). The relationship between VCI values and drought categories is shown in Table 2-2.

$$VCI_i = 100 \times \frac{NDVI_i - NDVI_{i,\min}}{NDVI_{i,\max} - NDVI_{i,\min}} \quad (2\text{-}3)$$

VCI value	Drought Category
[0, 5)	Exceptional
[5, 15)	Extreme
[15, 25)	Severe
[25, 35)	Moderate
[35, 50)	Abnormally dry condition
≥50	Non-drought

Table 2-2 Relationship between VCI and drought categories(STAR, 2010)

2.2. Vegetation drought monitoring

The type of drought this research focuses on is vegetative drought. It is similar to the agricultural drought while involving more vegetation than agricultural plants. Various studies have demonstrated the utility of satellite measurement of drought monitoring and provide analysis of the relationship between climate variables, e.g. precipitation, and satellite-based indices(Ji & Peters, 2003). Satellite image-based NDVI and VCI value are commonly used indices for vegetation health evaluation. They had been tested for the correlation with precipitation in East Africa and Iran, which indicated that these vegetation indices can be applied for drought monitoring. Results illustrate a highest correlation given with three-month of precipitation values. NDVI had better performance than VCI, which made NDVI a better indicator for vegetation changes and drought conditions (Bajgiran, Darvishsefat, Khalili, & Makhdoum, 2008; Davenport & Nicholson, 1993). However, there are some limitation of the correlation between NDVI and rainfall. First, in the period of three month, precipitation is not the only factor influencing vegetation. Second, correlation normally performs better during growing season. VCI is also noted as an indirect value showing vegetation stress from droughts, insects, disease and lack of nutrients(Vicente-Serrano, 2007). It is not highly correlated with meteorological drought and it is sensitively influenced by spatial environmental factors(Quiring & Ganesh, 2010). Carefully estimated must be taken when using the VCI for monitoring drought.

2.3. Markov chains for drought prediction

The Markov chain model is named after the Russian mathematician Andrei Andreevich Markov (1856-1922). He developed the Markov chain as a natural extension of sequences of independent random variables. In 1906, he proved that for a Markov chain with positive transition probabilities, named as regular Markov chain, and numerical states, the average of the outcome converges to the expected value of the limiting distribution, the fixed vector(Grinstead & Snell, 1997). A Markov process usually refers to a first-order process of autoregressive processes. The future development of this process is completely determined by the present state and is independent of the way in which the present state has developed, which explains the "first-order" (Chatfield, 2004). A Markov chain is a discrete-state random process of Markov process and it is a simple linear model (Meyn & Tweedie, 1993; Sinclair, 2005).

A Markov chain is a special case of a Markov process, which itself is a special case of a random or stochastic process. In the most general terms, a random process is a family, or ordered set of related random variables X(t) where t is a time parameter. There are many kinds of random processes. Two of the most important distinguishing characteristics of a random process are its state space, or the set of values that the random variables of the process can have; and the nature of the indexing parameter. We can classify random processes along each dimension.

1. State Space: Continuous-state: X(t) can take on any value over a finite or infinite continuous interval or set of such intervals; Discrete-state: X(t) has only a finite or countable number of possible values $\{s_0,s_1,s_2,\ldots,s_i,\ldots\}$. A discrete-state random process is also often called a chain.
2. Index parameter (time): Discrete-time: permitted times at which changes in value may occur are finite or countable; Continuous-state: changes may occur anywhere within a finite or infinite interval or set of such intervals

The state of a continuous-time random process at a time t is the value of X(t); the state of a discrete-time process at time n is the value of X_n (Sinclair, 2005). In a first-order Markov chain, X_{n+1} depends only in X_n, and not on any X_i, $1 \leq i \leq n$, as shown in equation 2-4(Sinclair, 2005).

$$P(X_{n+1} = s_i \mid X_n, X_{n-1}, \ldots, X_1, X_0) = P(X_{n+1} = s_i \mid X_n = s_j) \qquad (2\text{-}4)$$

Mishra (A. Mishra, Singh, & Desai, 2009) combined Joint probability Density Function and Markov chain process to characterize the drought with SPI. The time duration of mean drought interval changed in SPI series and in higher SPI series have higher probability of persistence for coming in the same states. Banik et al. (2002) modelled weekly rainfall data by Markov chain and the result illustrated a better understanding of drought-proneness and identification of the areas for a long term drought mitigation strategy. Paulo & Pereira (2007) used SPI value of 67 years rainfall data into Markov chain for both homogeneity and non-homogeneity to predict hydrological drought. The results showed that the Markov chain approach applied to time series is proved to be a useful tool to understanding the stochastic characteristics of drought and for early warning system.

Markov chains were also used for non-drought studies, such as Balzter (2000) who used four different locations with different climatic and soil conditions, to examine different grassland species. The results showed that Markov chain models are sensitive to changes and information from models is valuable for conservation, land planning and ecology. Compared to regression variables in crop field forecasting, the Markov chain proved to be an attractive alternative to regression analysis, because the approach is non-parametric, simple to implement with Markov property as the only requirement (Matis, Saito, Grant, Iwig, & Ritchie, 1985).

3. STUDY AREA

3.1. Location and extent of study area

The study is conducted in East Africa, a known drought prone region which experienced severe drought episodes in the last decade. Kenya is selected as the study area as it is one of the most affected countries in the region and has a good record of rainfall data that can be used for drought validation.

Kenya has a land area of 580,000 km² and a population of nearly 39 million residents. As Figure 3-1 shows, with the Indian Ocean to its southeast, Kenya is bordered by Somalia, Ethiopia, Sudan, Uganda and Tanzania. Lake Victoria is situated to the southwest, and is shared with Uganda and Tanzania. From the coast, the altitude changes gradually through the coastal belt and plains, 52 meters above sea level, the dry intermediate low belt to what is known as the Kenya Highlands, over 900 meters above sea level. Settlement is confined to places where water can be found. Wildlife is found on the greater part of the low belt(SouthTravles, 2001).

Figure 3-1 Kenya (sources: NationMaster http://www.nationmaster.com/country/ke-kenya)

3.2. Climate characteristics

The climate of Kenya varies by location. It is hot and humid at the coast, temperate inland and very dry in the north and northeast parts of the country. Although Kenya is centred at the equator, it shares the seasons of the southern hemisphere: the hottest time in Kenya is in February and March and the coldest in July and August with only a few degrees cooler. The coastal region is largely humid and wet. The low plateau area is the driest part of the country. The climate along the coast is tropical; this means rainfall and temperatures are higher throughout the year. The further away from the cost, the more arid the climate becomes. For many areas of Kenya, the daytime temperature usually rises about 12 °C. Elevation is the major factor in temperature levels, with the higher areas, on average, as 11°C cooler. At lower altitudes, the temperature increases throughout the day. There are slight seasonal variations in temperature of 4 °C cooler in the winter months. On the higher mountains, such as Mount Kenya, Mount Elgon and Kilimanjaro, the weather can become bitter cold for most of the year. Some snowfall has occurred on the highest mountains.

Higher elevation areas within the highlands receive much larger amounts of rainfall. The Lake Victoria basin in western Kenya is generally the wettest region in the country, particularly the highland regions to the north. Rainfall occurs seasonally throughout most of Kenya. There are two rainy seasons: the first rains, long rainy season, between March and May and the second rains, short rainy season, between October and December. The average seasonal rainfall amounts range between 120–240 mm (long rainy season) and 220–410 mm (short rainy season). The short rainy season is the major farming season, but rainfall in both seasons is highly variable and unreliable(Speranza, et al., 2008). The coast, eastern plateaus, and Lake Basin experience two rainy seasons and the highlands of western Kenya have a single rainy season.

3.3. Land use and agriculture

Both large and small holder farming is carried out in the highlands. Major cash crops are tea, coffee, pyrethrum, wheat and corn. Livestock farming is also practiced. The Lake Victoria Basin is dominated by Kano plains which are suited for farming through irrigation. The northern part of Kenya is plain and arid and pastoralism is the main land use activity.

Kenya derives their livelihoods mainly from crop production and marketing, livestock keeping and sale, as well as from low-income off-farm and non-farm activities. Although the major perennial rivers have potentials for irrigation, rain-fed agriculture is practised, and together with livestock keeping, are the major sources of livelihoods (Speranza, et al., 2008). At least 72% of the households simultaneously engage in crop production, livestock keeping and off-farm. 84% of the households derive part of their income from crops sales, 83% from livestock sales, while households earn off-farm incomes from various activities such as unskilled casual jobs (37%), business (28%), paid employment and pensions (26%), and remittances (20%). Thus incomes from sale of crops, sale of livestock and from off-farm activities are all comparably important but to different degrees at different times, vary from place to place, and are influenced by resource constraints, rainfall variability and drought (Speranza, et al., 2008). The vegetation and agricultural cash crops is shown in Figure 3-2.

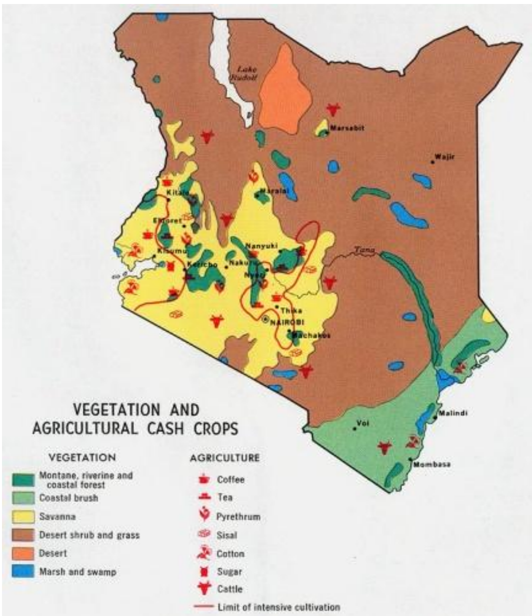


Figure 3-2 Vegetation and agricultural cash crops distribution in Kenya (source: http://www.nationmaster.com/country/ke-kenya)

3.4. Drought in Kenya

Drought is one of the significant hazards in Kenya especially during rainy seasons. When drought occurs, although it is not a terrible natural disaster in directly killing people, it can cause related situation, such as famine, which increases the number of affected people, livestock and deaths. Millions of families who rely on farming, fishing and herding have to leave their home to search water resources and food. Violence will rise because of pastoral territories conflict from immigration families. An estimated 400 people died due to such violence in 2009. As illustrated in Table 3-1, from the 1982 to 2011, drought has the largest number in affecting people. In 1999, the number of affected people reached to 23million, almost 80% of the residents influenced by that disaster. The most recent drought in 2008-2009, had the third number of affected people and the country is still recovering from the effect. Form last 30 years, drought occurred 8 times and totally killed people 196 and affected people 38 700 000, almost the same as the population (Centre for Research on the Epidemiology of Disasters http://www.emdat.be/). Table 3-1 show a list of all the natural disasters of the past 30 years, where it can be seen that drought occurred often.

Nature Disaster	Date(Month/Year)	No. Total Affected people
Drought	12/1999	23,000,000
Epidemic	01/1994	6,500,000
Drought	07/2008	3,800,000
Drought	12/2005	3,500,000
Drought	1991	2,700,000
Drought	07/2004	2,300,000
Drought	01/1997	1,600,000
Drought	03/1994	1,200,000
Flood	09/1997	900,000
Flood	10/2006	723,000

Table 3-1 Top 10 of number of natural disasters in Kenya from 1982 to 2011 (source: The OFDA/CRED International Disaster Database http://www.emdat.be)

4. DATA ACQUISITION

In Kenya, early drought detection, tracking mapping and severity assessment has been considerably constrained by incomplete meteorological data. The rain gauge network density and communication infrastructure have been declining over years, which makes the monitoring relying solely from precipitation data, not always reliable (Felix N. Kogan, 1997). When NDVI data are used to model vegetation stress, rainfall data can be used as a reference parameter for drought prediction. Generally, countries near equator do not have significant change of temperature from month to month or season to season. The dynamics of temperature in the Voi city in Kenya for instance, pixel 4 (one of the study areas, with more details in Section 4.1) in 2007 is shown as Figure 4-1. The highest temperature is 31°C in February, lowest 20°C in August and average around 25°C for the whole year. As such, temperature data are not considered and only satellite-derived NDVI data and rainfall data are considered. The following sections detail the characteristics of the data and the collection processes.

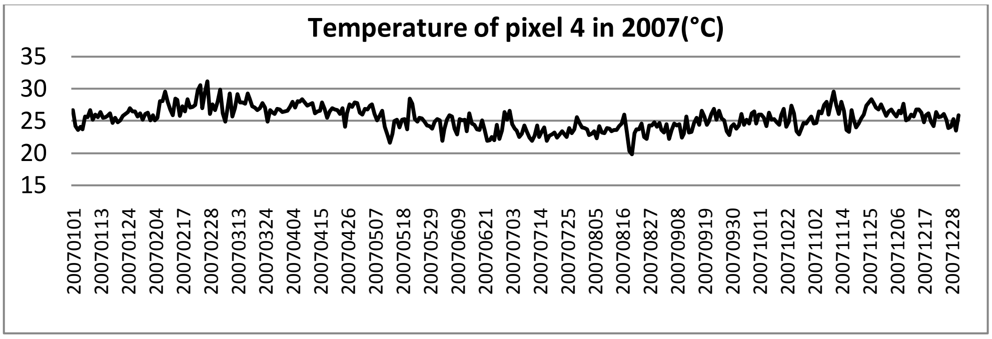


Figure 4-1 Temperature changes of study area pixel 4 in 2007(°C)

4.1. Satellite data acquisition of NDVI

The dataset used in this study is from the GIMMS (Global Inventory Modelling and Mapping Studies), and is composed of 10-day NDVI data from the NOAA-AVHRR (National Oceanic and Atmospheric Administration - Advanced Very High Resolution Radiometer). NDVI is calculated from two channels of the AVHRR sensor, the near-infrared (NIR) and visible (VIS) wavelengths as shown in equation 4-1.

$$NDVI = (NIR - VIS)/ (NIR + VIS) \quad (4\text{-}1)$$

Values of NDVI for vegetated land generally range from about 0.1 to 0.8, with values greater than 0.5 indicating dense vegetation and smaller than 0.3 indicating less vegetation. In the GIMMS 10-day NDVI dataset, the pixel values are stretched from [0, 1] to [0, 250], water values are assigned to 255, erroneous and missing values to 253 and 254. These 10-day composite images are constructed at regular intervals by selecting pixels with the maximum NDVI, in order to construct cloud-free views of the Earth. A 10-day time step is generally selected as the minimum period since the NOAA orbit repeats at that frequency. The data are given in an 8 km Albers Equal Area Conic projection, Clark 1866 ellipsoid and in geographic coordinates, WGS84 datum at 0.07272727 degree resolution per pixel. The data is available from July 1982 to present. The long-term mean data is calculated from 1982 to 2008 and short-term mean data is calculated from 2004 to 2008. This dataset is available for free at FEWS NET (Famine Early Warning System Network http://earlywarning.usgs.gov/fews/africa/index.php). The dataset is inter-calibrated with SPOT Vegetation NDVI, and uses NOAA-17 data since January 2004. The NOAA-17 NDVI data have also been inter-calibrated with NOAA-16 and previous NDVI products. More information can be

found in GIMMS document
(ftp://landval.gsfc.nasa.gov/Documentation/GIMMS_NDVI_8km_doc.pdf).

For this study, a subset of 180 NDVI images of East Africa is selected covering the period from the first dekadal of January to the last dekadal of December from 2004 to 2008. The long term mean values also selected for calculating NDVI anomaly. Spatial subsets of locations of interest are extracted in Kenya and each pixel covers 64 km^2 areas on the ground.

Point 1 (3.117N, 35.617E) is selected in the Rift Valley, as shown in Figure 4-2. It is located west of the Lake Turkana and east of the city of Lodwar, which is the largest town in north-western Kenya. The river Turkwel goes through the area and the Loima Hills lie to its west. This area is arid and hot. The mean annual rainfall is less than 250 mm and the occurrence of rainfall is very erratic and unpredictable, though rainfall occurs mostly during "long rain" season. Hence the agriculture type of this area is pastoral farming. Point 2 (1.75N, 40.067E) is selected near the city of Wajir, which is in the North Eastern Province the headquarters of Wajir District. The satellite image from Google Earth of Point 2 is shown in Figure 4-3. This area is located is arid and prone to drought. In spring 2006, there was a severe famine caused by drought. The annual precipitation is 240 mm and there are two rainy seasons in this area. The agriculture type is pastoral farming in low plateau. Point 3 (0.5S, 37.45E) is selected north-east of the city of Embu, which serves as the headquarters of Eastern Province and Embu District. The satellite image of this area is shown in Figure 4-4. The average annual precipitation is 870 mm, with the highest rainfall in April and November. The annual average temperature is 19.2°C and it varies slightly from month to month. The agriculture type of this area is cultivation. Point 4 (3.4S 38.567E) is selected south of city of Voi, which is a market town in the Coast Province. The satellite image of this area is shown in Figure 4-5. The average temperature is 25.4°C and annual precipitation is 570 mm. The agricultural type of this area is cultivation.

Figure 4-2 Point 1 from Google Earth© images

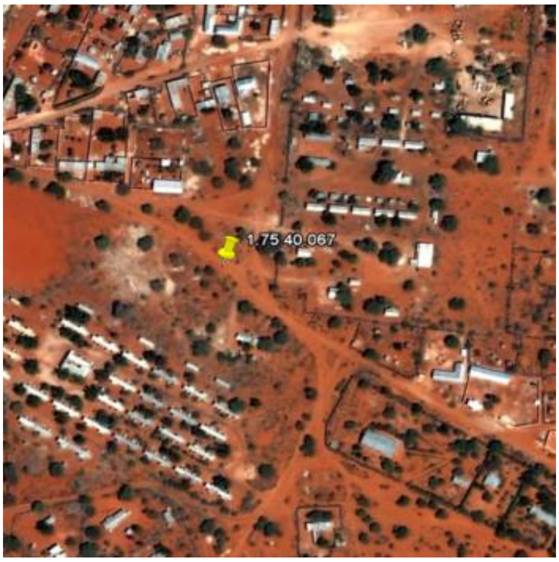


Figure 4-3 Point 2 from Google Earth© images

Figure 4-4 Point 3 from Google Earth© images

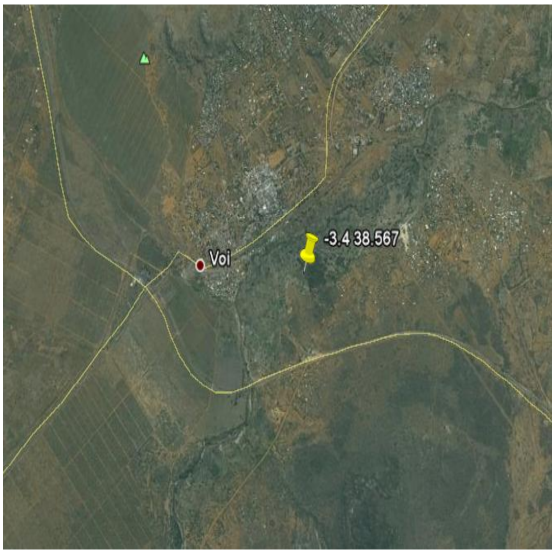


Figure 4-5 Point 4 from Google Earth© images

These points are chosen based on the station position on the National Climate Data Centre (NCDC, http://gis.ncdc.noaa.gov/website/ims-cdo/gsod/viewer.htm), for a later validation of drought using rainfall data collected at those stations. In order to reduce random noise, 4 surrounding points are selected around the each of the five originally selected points. In the context, Region1 represents point 1 and its 4 surrounding pixels, so as for region2, region3 and region4. In Figure 4-6, the light blue dots are shown the locations of the four originally selected pixels in Kenya, on a satellite image. Number all the pixels from 1 to 20. Region 1 includes pixel 1 to pixel 5 and pixel 3 cover the area of point 1. Region 2 contains pixel 6 to pixel 10 and pixel 8 is the original point2. Pixels 11 to pixel 15 are in region 3 and the rest pixels belong to region4. Pixel 13 cover the area of point 3, so does pixel 18 to point 4.

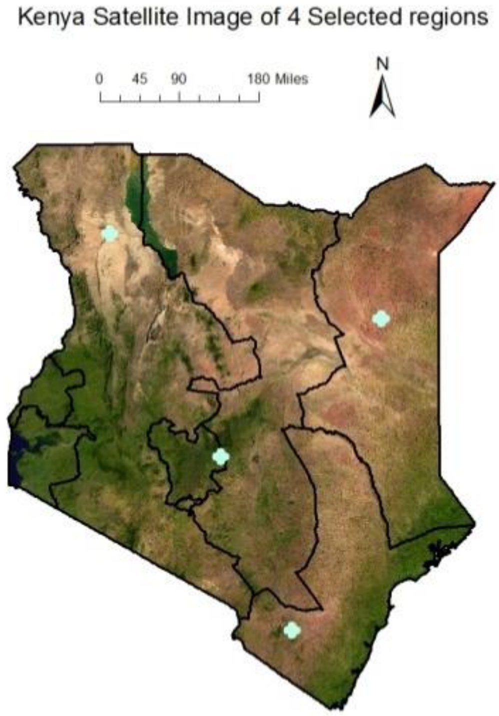


Figure 4-6 Kenya satellite images of 4 selected pixel locations (Source from: http://www.maplibrary.org/stacks/Africa/Kenya/index.php)

4.2. Station data

The Station data are provided by the World Meteorological Organization (WMO) World Weather Watch Program. The input data sources come from the Integrated Surface Data (ISD), which includes global data obtained from the USAF Climatology Centre, located in the Federal Climate Complex with NCDC. The latest daily summary data are normally available 1-2 days after the date-time of the observations used in the daily summaries. The online data files begin with the year 1929, and are now at the Version 7 software level. Over 9000 stations' data are available. This dataset contain a lot of gauge information such as daily mean temperature, maximum and minimum temperature, precipitation and wind. However, considering the incompleteness and low density distribution of station data on in Kenya, the satellite-based Rainfall Estimated (RFE) dataset are used, after comparing the correlation with the gauge information. The correlation of this two dataset is shown in Section 4.4.

4.3. Rainfall estimated data

RFE version 2.0 has been implemented by NOAA's Climate Prediction Centre. Input data used for operational rainfall estimates are from 4 sources:

1) Rainfall amounts from up to 1,000 rain gauges across Africa, quality controlled using maximum/minimum range checks and plausibility checks against the satellite estimates
2) Precipitation estimates from the Special Sensor Microwave/Imagery (SSM/I) microwave sensors on the Defense Meteorological Satellite Program polar orbiting satellites, which are available up to 4 times per day at a resolution of approximately 25 km
3) Precipitation estimates from the Advanced Microwave Sounding Unit (AMSU-B) aboard NOAA-N series polar orbiting satellites, at a similar spatial and temporal resolution to SSM/I
4) Precipitation estimates based on Meteosat 5 and 7 infrared cloud top temperature measurements using the GOES Precipitation Index with a threshold of < 235 K at half hourly intervals and a resolution of 4 km(Sene, 2010)

RFE 2.0 obtains the final daily rainfall estimation using a two-step merging process. First, the three satellite estimates are combined linearly using predetermined weighting coefficients. This step is to reduce the random error of the satellite precipitation. The second step of the merging process compares the satellite-estimated precipitation in step one with GTS rain gauge data to remove as much bias as possible, then sums daily totals to produce dekadal estimates. The data are given in an 8km Albers Equal Area Conic projection, Clark 1866 ellipsoid and in geographic coordinates, WGS84 datum at 0.07272727 degree resolution per pixel. The data are available from 1995 to present. The RFE Long Term Mean was derived from interpolated rain gauge data for the period 1920 to 1980. The Short Term Mean was derived from satellite rainfall estimates for the period 2005 to 2009. In this research, the RFE Long Term Mean value is derived from 1995 to 2008.

4.4. Comparison between station data and RFE data

The comparison is evaluated for every 10-day rainfall data of the 4 originally selected pixels during the whole year of 2008 using the Pearson correlation (r). There should be 36 data in comparison from station data and RFE data, but there are 34 data available from station dataset in pixel 4. The result is shown in Table 4-1. According to the document of Station data from NCDC, it is still possible that for some daily record, the precipitation occurred but was not report. Hence for small precipitation occurrence, RFE has larger value than Station value, which may influence the correlation coefficient. The correlation value in pixel 4 is quite lower than other pixels. It could be that the number of available datasets influenced the result. Also, RFE data is based on the four kinds of data sources, which could have generated errors during data mining.

Pixel label	Available station data	Available RFE data	Correlation coefficient
Pixel 1	36	36	0.99
Pixel 2	36	36	0.87
Pixel 3	36	36	0.92
Pixel 4	34	36	0.75

Table 4-1 Correlation of station data and RFE, the second and third columns show the available numbers of station data and RFE data for comparison

5. METHODS

5.1. Drought characterisation

5.1.1. NDVI anomaly

Drought is a natural hazard of water deficiency from normal demands. It should be considered relative to some long-term average condition rather than absolute condition (D. A. Wilhite & Svoboda, 2000). Traditional methods of drought assessment and monitoring heavily depend on the density of rainfall data network. Reliable satellite imagery coving large regions over long periods of time has efficiently reduced the influence of density limitation (Gouveia, Trigo, & DaCamara, 2009). NDVI is a widely used vegetation index of environmental studies, but it does not reflect drought or non-drought, low value of NDVI, from 0 to 0.3, may represent less density of vegetation or non-growing season. NDVI anomaly calculated as in Equation 2-2 is used for this study as the long-term mean values are provided. When the $NDVI_{anomaly}$ is negative, it indicates a below-normal vegetation condition and hence a prevailing drought situation. The larger the negative departure, the greater drought severity may be suggested.

5.1.2. Correlation with rainfall

This study assuming that NDVI anomaly is caused by drought and can be used as a vegetation drought indicator. If the NDVI anomaly demonstrates a sensitive and consistent response to an inter-dekadal fluctuation of rainfall, it can provide an indicator of drought conditions (Davenport & Nicholson, 1993). In order to study the statistical relationship between various time lag periods of precipitation and NDVI anomaly, Pearson(r) correlation analysis is applied and correlation coefficients between the value of vegetation index and precipitation data are determined.

A Pearson correlation r is a number between -1 and +1 that measures the degree of association between two dataset, in this study, NDVI anomaly and precipitation. A positive value for the correlation implies a positive association. A negative value for the correlation implies a negative or inverse association. Zero implies that there is no linear correlation between the variables.

5.2. Data processing methodology

The main data processing adopted in this study, as shown in Figure 5-1, involve to:

1. Get input data of four study regions of NOAA-AVHRR image from 2004 to 2008.
2. Calculate deviation NDVI anomaly as NDVI deviation from the long-term mean NDVI values.
3. Define drought classes for the study area and use the classes as input states to make the transition matrix of Markov chains. Membership function is used to classify drought.
4. Use both transition matrix and membership function to predict drought
5. Validate the predicted results using satellite image data.

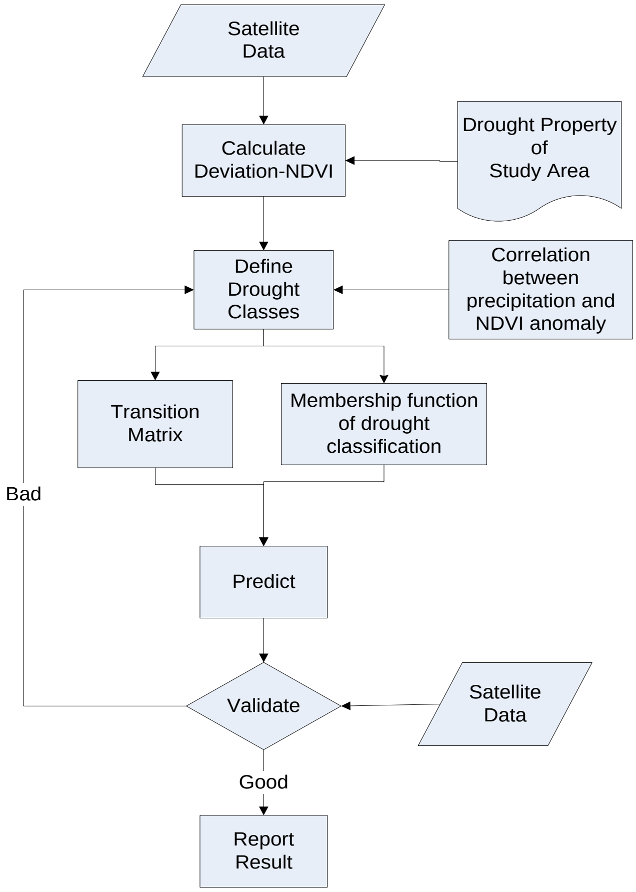


Figure 5-1 Flowchart of the data processing

5.2.1. Drought modelling from NDVI anomaly using fuzzy set theory

Fuzzy sets are sets whose elements have degrees of membership. Fuzzy sets were first introduced by Lotfi A. Zadeh (1965) as an extension of the classical set theory. In classical set theory, the membership of elements in a set is assessed in binary terms according to a bivalent condition with only two values, 0 and 1. In fuzzy set theory, classical bivalent sets are usually called crisp sets. By contrast, fuzzy set theory permits the gradual assessment of the membership of elements in a set, as shown in Figure 5-1. Each fuzzy set is bounded by transition zones. Fuzzy membership function $\mu_A(x)$ has often been regarded as a compatibility function, which denotes the degree in which the proposition "x is A" is true(Drakopoulos, 1995).

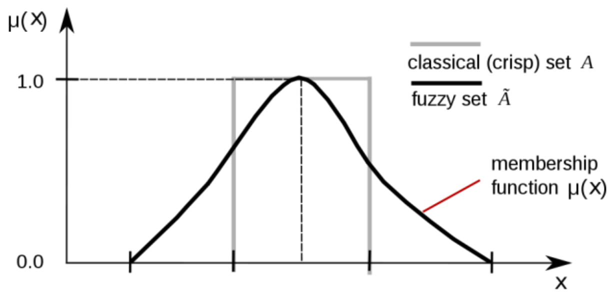


Figure 5-1 Membership function of a fuzzy set

To define a fuzzy set, values of threshold, dispersion and selection of appropriate membership function is required. Make a fuzzy set A in R^n (an Euclidean n-space) is determined by a membership function $\mu_A : R^n$

→ [0, 1], which assigns a membership grade with [0, 1] to element $\mu_A(x)$ and R^n is an Euclidean n-space. The nearer the value of $\mu_A(x)$ to unity, the higher the grade of the membership of x is in A. The probability of a fuzzy event x is defined by the Lebesgue-Stieltjes integral, as shown in equation 5-1(Zadeh, 1968). The probability of a fuzzy event is the expectation of its membership function. The existence of the integral in equation 5-1 is assumed by the assumption that μ_A Borel measurable.

$$P(A) = \int_{R^n} \mu_A(x)dP = E(\mu_A) \quad (5\text{-}1)$$

There are different types of membership functions. Common examples are Boolean, bell-shaped, triangular, Gaussian, trapezoidal, or function with a central core region and upper and lower transition zones with different width. The triangular shape type of membership function has a single core value and linear dispersion slopes, while bell-shaped has curved dispersions. The trapezoidal shape type of membership function has an interval of value on core zone. The parameter of these membership functions can be selected with statistical methods or based on field knowledge(Musaningabe, 2007). In this research, the parameter is selected with statistical methods, and trapezoidal shape function is selected to model the drought.

Vegetative drought is not a crisp phenomenon in nature. Dividing vegetation drought into different classes can be more realistic while using a membership function than using crisp sets. Rulinda et al. (2010) modelled vegetative drought in eastern Africa using the deviation of the current NDVI values from long term mean values, and assessed the evolution of drought conditions. While doing so, they accounted for the gradual transition between drought classes, and assumed a linear relationship between vegetation stress and drought.

In the section 2.1.2, there are several ways in drought classification. For different drought indices, the number of classes and the name of drought categories are different. As NDVI is a new indicator of drought compared to meteorological indices, such as SPI and PDSI, there is no standard of drought classification. The classification in this research is based on the way how SPI classified. There are 7 classes in SPI classification, 3 drought classes, 4 non-drought classes. As the aim in this research focus on drought situation in study area, the number of drought classes increases to 4 based on the SPI and non-drought classes reduces to 2. Hence, number of drought classes is arbitrarily decided to 6 and the drought categories of these six classes are: extremely drought, severely drought, moderately drought, dryness, wet and moderately wet. In the following context, these six classes are numbered from 1 to 6 sequentially. Here, fuzzy classification is introduced as fuzzification. As drought is a vague phenomenon, the drought categories though vague are usually classified into binary classes. The result of fuzzification for each drought classes is a vague partition, consisting of values of each drought class. In Figure 5-2, general membership function of each drought classes is shown. Because the shape and calculation of membership function for middle four drought classes: severely drought, moderately drought, dryness and wet, the similar, only three memberships are shown in the figure. Range $[a_i, b_i]$, $i \in N$, which N being the total number of classes, is the core zone for each membership function of each drought class, where i represents the index of drought classes. Range $[b_{i-1}, a_i]$ is the transition zone. From the Figure 5-2, it clearly shown that the first class and the last drought class are the special cases with just one transition zone.

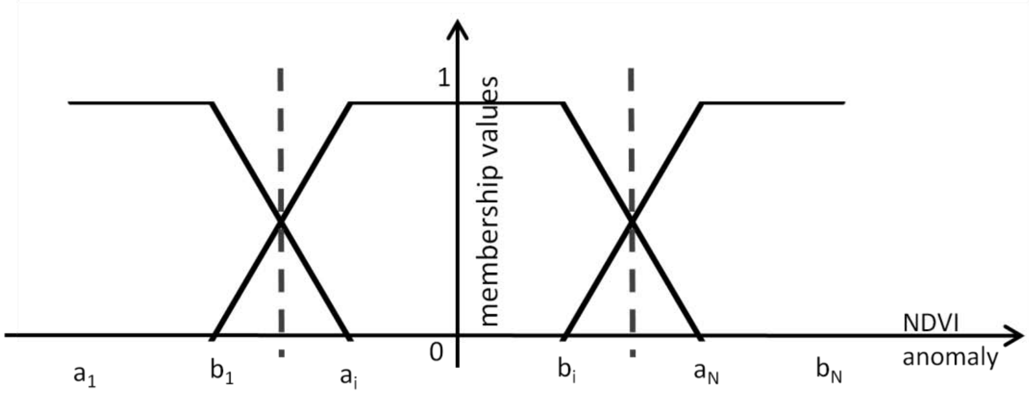


Figure 5-2 Membership functions concept for drought classes, [a_1, b_1] refers to core zone of the first drought class, [a_i, b_i], $i \in \{2, 3, \ldots, N\text{-}1\}$ refers to core zone of middle drought classes, [a_N, b_N] refers to core zone of last drought class The dash lines show the boundaries for crisp classification.

The formula of the six membership functions of drought classes can be represented into three kinds of functions as shown in Figure 5-2. The formulas are shown in equation 5-2, 5-3 and 5-4, where x: NDVI anomaly; $f(x)$ is the membership value, core zone for each drought class is [a_i, b_i], $i \in \{1, 2, \ldots, N\}$ represent drought class.

For first class with only one transition zone on the left:

$$f(x)=\begin{cases}1, x<b_1 \\ \dfrac{x-a_2}{b_1-a_2}, b_1 \le x \le a_2 \\ 0, x>a_2\end{cases} \tag{5-2}$$

The middle functions with two transition zones, where $j \in \{2, 3, .., N\text{-}1\}$:

$$f(x)=\begin{cases}-\dfrac{x-b_{j-1}}{b_{j-1}-a_j}, b_{j-1} \le x \le a_j \\ 1, a_j < x < b_j \\ \dfrac{x-a_{j+1}}{b_j-a_{j+1}}, b_j \le x \le a_{j+1} \\ 0, others\end{cases} \tag{5-3}$$

The last drought class with only one transition zone on the right:

$$f(x)=\begin{cases}0, x<b_{N-1} \\ -\dfrac{x-b_{N-1}}{b_{N-1}-a_N}, b_{N-1} \le x \le a_N \\ 1, x>a_N\end{cases} \tag{5-4}$$

NDVI anomaly is not a widely used indicator for vegetation stress or drought. There is not a standard to classify the drought as SPI or PDSI shown in Section 2.1.2. In this study, original classification is based on the same frequency in the whole dataset. When building transition matrix for Markov chains, equal frequencies can ensure that the number of observations per class is sufficiently high to obtain reasonable estimates(Bickenbach & Bode, 2003). [a_i, b_i] represents 95% observations of each original class and rest 5% observations of the same class are assigned to transition zone. This is the primary step to define ai and bi of each drought class. As slope of membership function are determined by ai and bi, they can be adjusted during prediction until satisfactory.

5.2.2. Drought dynamic modelling using Fuzzy Markov chains

The dynamics of vegetative drought is modelled using Markov chains applied to fuzzy states or classes. As mentioned in Chapter 2, the Markov chains approach has been successfully used to predict the dynamics of drought, and in particular meteorological drought. In those studies, the states of drought are modelled as crisp. In this study we model the states of Markov chains with fuzzy classes which are mentioned above. Dilo et al (2006) use fuzzy Markov chains to model beach erosion with the ratio of average area. Results showed that those transition probabilities can be used to predict the volume of erosion for next year. Miao et al. (2005) predicted the future value of cultivated land demand in land use planning using the traditional time homogeneous Markov chains model with fuzzy probabilities. The result verified that the accuracy of prediction is improved with fuzzy theory. The fist-order Markov chain "is a stochastic process in which the future development depends only on the present state, but on the past history of the process or the manner in which the present state was reached"(Collins, 1975). In a second-order Markov chain, the states of system at time t_2 would also depends on the states on both at time t_0 and t_1. However, even in first-order Markov chain, at time t_1, the state S_1 only depends on state S_0 and not refer to S_{0-1}, S_{0-2}, S_{0-3},…, these past states are included in S_0, which represents the sum of the past history.

As mentioned in Section 2.3, a Markov chain is conditionally independent from X_0, X_1, X_2,…, X_{t-1} given by X_t; the probability that X_{t+1} takes a particular value j depends on the past X_t: where $i, j \in \{1, 2, 3, 4, 5, 6\}$ are the maximum fuzzy class values; $t \in T$, which is the whole time series. The probability formula has a little change from equation 2-4 to equation 5-5.

$$P(X_{t+1} = j \mid X_t, X_{t-1}, ..., X_1, X_0) = P(X_{t+1} = j \mid X_t = i) \quad (5\text{-}5)$$

The transition probability p_{ij} is the element in the probability transition matrix **P**, where the Markov chain is at the next time point in class j, given that it is at the present time point in state i. The transition probability matrix **P** is estimated from the samples, counting the number of times, n_{ij}. The formula is shown in equation 5-6.

$$\mathbf{P} = \left[p_{ij} \right] = P(X_{t+1} = j \mid X_t = i) \quad (5\text{-}6)$$

where, $p_{ij} = \frac{n_{ij}}{\sum_j n_{ij}} \geq 0$, $\sum_j n_{ij} = 1$

A Markov chain is called an regular Markov chain if it is possible to go from every state (drought class) to every state, not necessarily in one step(Grinstead & Snell, 1997), and some power of matrix **P**, there are only positive entries the transition matrix. For a regular Markov chain, there are two important theorems:
Theorem I: Matrix **V** is the power of **P**, in which each row is the same probability vector **v** and all the element in **a** are all positive. The formula is shown in equation 5-7.

$$\mathbf{V} = \lim_{n \to \infty} \mathbf{P}^n \quad (5\text{-}7)$$

Theorem II: If **V** and **v** are as in Theorem I for **P**, then **vP** = **v**. The matrix **V** is defined as the limiting matrix and **v** is called a fixed row vector(Collins, 1975).

Since the chain is an ergodic chain, it is possible to reach drought class j from any other class. Hence, if a Markov chain is started in class i, the expected number of steps of reach class j for the first time is called the mean first passage time from class i to class j. It is denoted by m_{ij}. If $i=j$, in some reference, m_{ij} called mean recurrence time. It is the expected number of steps to return to the same class. The matrix of mean first passage time is denoted by **M**. The formula is shown in equation 5-8.

$$\mathbf{M} = (\mathbf{I} - \mathbf{Z} + \mathbf{E}\mathbf{Z}_{dg})\,\mathbf{D} \qquad (5\text{-}8)$$

Where:

$\mathbf{I}$ is an identity matrix;

$\mathbf{Z} = (\mathbf{I} - (\mathbf{P} - \mathbf{V}))^{-1}$, $\mathbf{P}$ is the transition matrix and $\mathbf{V}$ is the limiting matrix of $\mathbf{P}$, $\mathbf{Z}$ denoted as fundamental matrix

$\mathbf{Z}_{dg}$ results from $\mathbf{Z}$ by setting off diagonal entries to 0

$\mathbf{E}$ is a matrix with all entries 1

$\mathbf{D}$ is the diagonal matrix with j-th entry $1/\mathbf{a}_j$, $\mathbf{a}$ is a fixed row vector of $\mathbf{P}$

The mean first passage time is applied to estimate: the expected time in each class of drought, the recurrence time to a particular drought class and the expected time for NDVI anomaly value change from a particular class to another. It provides useful information for early warning system of drought prediction. Before applying the Markov chain with fuzzification, the Markovian properties should be tested. The following section details the processes.

5.3. Testing the Markovian property of vegetative drought

The reliability of predicting NDVI anomaly, considered as vegetative drought, using Markov chain generally depends on two conditions. First, the data-generating process must meet the Markov property, which involves time homogeneity and time independence. The transition probabilities can be estimated using the Maximum likelihood ratio (*LR*) criteria and Pearson χ^2-tests (Q) under specific null and alternative hypotheses. Although the *LR* and Pearson χ^2 statistic are asymptotically equivalent, in cases of poor asymptotic, they are not equivalent (Anderson & Goodman, 1957). Therefore, both of the statistics will be tested. Second, the estimates have to be based on a number of observations, large enough to be able to rely on the asymptotic properties of the estimators(Bickenbach & Bode, 2003). Otherwise, the accuracy will be rather poor.

5.3.1. Tests of time homogeneity

The test of time homogeneity, i.e. time stationarity, is appropriate for deciding whether the transition probabilities are constant over time. The test is done by dividing the entire sample period into M sub-periods and comparing each transition matrix of sub-periods from the entire sample. Under H_0: $\forall m$: $p_{ij|m} = p_{ij}$ (m = 1, 2, …, M) H_a: $\exists m$: $p_{ij|m} \neq p_{ij}$, $\alpha = 0.05$. The comparison can be implemented using the LR or Pearson χ^2-test as shown in equation 5-10 and equation 5-11(Bickenbach & Bode, 2003).

$$p_{ij|m} = \frac{n_{ij|m}}{n_{i|m}} \qquad (5\text{-}9)$$

where, $n_{ij|m} = \sum_{t \in m} n_{ij|m}(t)$, $n_{i|m} = \sum_{t \in m} n_{i|m}(t-1)$

$$LR^{(M)} = 2\sum_{m=1}^{M}\sum_{i=1}^{N}\sum_{j\in C_{i|j}} n_{ij|m}\ln\frac{p_{ij|m}}{p_{ij}} \sim \chi^2(\sum_{i=1}^{N}(c_i-1)(d_i-1)) \tag{5-10}$$

$$Q^{(M)} = \sum_{m=1}^{M}\sum_{i=1}^{N}\sum_{j\in C_i} n_{i|m}\frac{(p_{ij|m}-p_{ij})^2}{p_{ij}} \sim \chi^2(\sum_{i=1}^{N}(c_i-1)(d_i-1)) \tag{5-11}$$

Where N is the drought classes= {1, 2, 3, 4, 5, 6}; i = {1, 2, …, N} at time t-1; j = {1, 2, …, N} at time t; $p_{ij|m}$ is the transition probability of sub-sample m; p_{ij} is the transition probability of entire sample. $C_{i|m}$ is the set of non-zero transition probabilities in the i^{th} row of the transition matrix estimated from the m^{th} sub-period. c_i is the number of elements in C_i, and d_i is the number of sub-periods for which a positive number of observations is available for the i^{th} row. Both of the tests have an asymptotic χ^2 distribution with degrees of freedom (df) equal to the number of additional independent restrictions imposed by H_0 as compared to H_a. For each sub-period test m, the df =N(N-1), LRprob and Qprob separately represent the probability of LR test and Q test under df (Anderson & Goodman, 1957).

5.3.2. Tests the time independence

The aim of this test is to see whether the chain is the first-order Markov chain. In estimating the order of Markov chain, first, it is required to test order 0 versus order1; second, to test order 1 versus order 2. If the test of order 0 against order 1 is rejected, and the order 1 against order 2 is "fail to reject", then the process is assumed to be the first-order. In this study, the Markov property requires the transition probabilities to be of the first-order. If the order 1 against order 2 is also rejected, then the chain may be of the second-order or higher order and the transition matrix will be misspecified(Bickenbach & Bode, 2003).

To test for the order 0, under H_0: $\forall i$: $p_{ij} = p_j$ (i, j = 1, 2, …, N) H_a: $\exists i$: $p_{ij} \neq p_j$, α = 0.05, where $p_j = n_j/n$, $n_j = \sum_t n_j(t)$. The LR and Pearson χ^2-test read as shown in equation 5-12 and equation 5-13(Bickenbach & Bode, 2003).

$$LR^{(O(0))} = 2\sum_{i=1}^{N}\sum_{j\in C_i} n_{ij}(t)\ln\frac{p_{ij}}{p_j} \sim \chi^2((N-1)^2) \tag{5-12}$$

$$Q^{(O(0))} = \sum_{i=1}^{N}\sum_{j=1}^{N} n_i(t-1)\frac{(p_{ij}-p_j)^2}{p_j} \sim \chi^2((N-1)^2) \tag{5-13}$$

where, $n_i(t-1) = \sum_j n_{ij}(t)$, C_i = {j: p_{ij}>0}

To test the order 1 against order 2, a second-order chain is defined by also considering the vegetation stress classes h (h = 1, 2, 3, …, N) at time t-2 and the pair of successive classes h and i forms a composite state. The probability of moving to state j at time t, given it was in h at time t-2 and in i at t-1, is p_{hij}. The hypothesis H_0: $\forall h$: $p_{hij} = p_{ij}$ (h = 1, 2, …, N) H_a: $\exists h$: $p_{hij} \neq p_{ij}$, α = 0.05. $p_{hij} = n_{hij}/n_{hi}$, $n_{hij} = \sum_{t=2}^{T} n_{hij}(t)$ and $n_{hi} = \sum_{t=2}^{T} n_{hi}(t-1)$. The LR and Pearson χ^2-test read as equation 5-14 and equation 5-15.

$$LR^{(O(1))} = 2\sum_{h=1}^{N}\sum_{i=1}^{N}\sum_{j\in C_{hi}} n_{hij} \ln\frac{p_{hij}}{p_{ij}} \sim \chi^2(\sum_{i=1}^{N}(c_i-1)^2) \tag{5-14}$$

$$Q^{(O(1))} = \sum_{h=1}^{N}\sum_{i=1}^{N}\sum_{j\in C_i} n_{hi} \frac{(p_{hij}-p_{ij})^2}{p_{ij}} \sim \chi^2(\sum_{i=1}^{N}(c_i-1)^2) \tag{5-15}$$

where $C_i = \{j: p_{ij} > 0\}$, c_i is the number of elements in C_i.

5.4. Drought prediction

For each pixel value α at time *t* is applied a fuzzy membership function *f(x)*. *f(a)* =*π(a)* = (π_1, π_2, …, π_N). π_1, π_2, …, π_N refer to the membership value of each drought classes at time *t*. To predict the drought class P_{Class} at time *t+1*, $P_{Class\ t+1}$=π(α)***P**, where **P** is the transition matrix. There are some correlation between NDVI value and RFE data.

The Markov chain is applied on the classified images of the last 10-day of year 2008, π_{2008}, to predict the class probabilities, $P_{Class2009}$, of the first 10-day drought situation in 2009. The formula is shown in Equation 5-16. The drought class with the maximum probability value in $P_{Class2009}$ is the most

$$P_{Class2009} = f_{2008}*\mathbf{P} = \pi_{2008}*\mathbf{P} \tag{5-16}$$

5.5. Validation

The validation data is the same satellite image from GIMMS 10-day NDVI dataset of year 2009. The reference values of 2009 are denoted $V_{Class2009}$. The validation is from comparison of the predicted result $P_{Class2009}$ and $V_{Class2009}$. The compassion is made by the probability distribution in $P_{Class2009}$ and membership value distribution in satellite image in $V_{Class2009}$. The drought class assigned to $P_{Class2009}$ and $V_{Class2009}$ are according to the maximum values. Root Mean Square Error (RMSE) is applied to be the quantity value of validation.

6. RESULT AND ANALYSIS

6.1. Correlation between RFE and NDVI

Table 6-1 and Table 6-2 shows the correlation between RFE data and NDVI value in time series, where t means current time, *t+i* means the current RFE value added to the *i* previous RFE dekadal value, $i = \{1, 2, 3, \ldots\}$. $RFE_{t+i} = \sum_{t-i}^{t} RFE$. In practical, the correlation test is taken until *t+15*, approximately equal to five month rainfall data. Since there are two growing season in Kenya, and each rainfall season lasts about four months and the interval between rainfall seasons are 2-3 months, correlation test taking more than 5 months data makes less sense in Kenya. Also, the correlation values from *t+11* to *t+15* are not shown in the Table 6-1 and Table 6-2, because of space limitation and the values decrease dramatically after *t+10*.

Region	Pixel	*t*	*t+1*	*t+2*	*t+3*	*t+4*	*t+5*	*t+6*	*t+7*	*t+8*	*t+9*	*t+10*
	1	-0.014	0.032	0.114	0.181	0.233	0.278	0.301	0.299	0.300	0.313	0.316
	2	-0.028	0.035	0.143	0.219	0.286	0.338	0.340	0.321	0.297	0.268	0.228
1	3	0.021	0.092	0.229	0.302	0.333	0.371	0.384	0.370	0.368	0.377	0.360
	4	0.028	0.115	0.207	0.276	0.318	0.356	0.370	0.376	0.408	0.438	0.440
	5	0.101	0.180	0.291	0.330	0.377	0.410	0.423	0.424	0.453	**0.480**	0.479
	6	0.014	0.161	0.327	0.472	0.578	0.672	0.736	0.786	**0.804**	0.804	0.788
	7	0.100	0.269	0.444	0.591	0.703	0.773	0.813	**0.832**	0.827	0.806	0.769
2	8	0.079	0.235	0.417	0.559	0.681	0.765	0.819	**0.848**	0.847	0.826	0.789
	9	0.098	0.263	0.436	0.580	0.697	0.783	0.828	**0.852**	0.848	0.823	0.782
	10	0.127	0.298	0.484	0.649	0.761	0.812	0.830	**0.833**	0.817	0.786	0.742
	11	0.067	0.209	0.335	0.433	0.495	0.546	**0.563**	0.560	0.533	0.487	0.441
	12	-0.004	0.120	0.254	0.365	0.468	0.545	0.578	**0.586**	0.564	0.522	0.465
3	13	0.033	0.181	0.330	0.440	0.518	0.572	0.594	**0.594**	0.566	0.531	0.485
	14	0.033	0.175	0.322	0.433	0.515	0.572	0.597	**0.599**	0.579	0.555	0.520
	15	0.132	0.212	0.278	0.303	0.322	0.328	0.310	0.262	0.207	0.141	0.068
	16	0.292	0.459	0.616	0.715	0.765	0.779	**0.781**	0.770	0.754	0.734	0.708
	17	0.279	0.406	0.509	0.588	0.638	0.656	**0.659**	0.652	0.628	0.600	0.569
4	18	0.186	0.343	0.500	0.594	0.645	0.665	**0.679**	0.673	0.645	0.607	0.569
	19	0.312	0.492	0.647	0.742	**0.792**	**0.792**	0.775	0.741	0.712	0.682	0.650
	20	0.198	0.342	0.482	0.561	0.607	0.634	**0.655**	0.652	0.642	0.619	0.584

Table 6-1 Correlation (*r*) between RFE and NDVI for 20 pixels for the period between 2004 to 2008, value in bold is the maximum and significant $\alpha < 0.001$.

From the Table 6-1, study pixels in Region 1(defined in Chapter 1) hardly show signification correlation between NDVI and RFE while region 2 have high correlation value between current NDVI and rainfall accumulation for the past 80 days (column "*t+7*") above 0.8, and both region 1 and region 2 have in the same land use type, pastoral farming. In this region1, the vegetation density is low as shown by the maximum NDVI value being 0.288 for all pixels for 5 years. In region 2, as we can observe from the maximum NDVI value of 0.76 and rainfall amount between 1394mm to 1482mm in 5 years, the vegetation density is higher than in group 1. The rainfall amount is also slightly higher than region 1 of total rainfall amount is 1233mm to 1319mm in a 5-year (2004-2008). Hence, the low correlation observed in group 1 can be caused by the low density of vegetation. In region 3, the highest correlation is for the past 80 days as in the region 2. However, not all of the pixels in region 3 have the high correlation value with rainfall data. The NDVI value in pixel 15 is from 0.792 to 0.1, and the long-term mean value of this pixel area varies from 0.612 to 0.284. The total amount of precipitation in pixel 5 in 2004-2008 is 3100mm, approximately 620mm per year. So the reason why there is a non-significant correlation in this area maybe that the input data has some random noise, as the nearby pixel has high correlation with rainfall data. In region 4, the most highest correlation between current NDVI and RFE accumulation is for the past 70 days, as shown in column "t+6".

Region	Pixel	t+0	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8	t+9	t+10
	1	-0.312	-0.037	-0.048	-0.064	-0.075	-0.056	-0.046	-0.010	0.032	0.091	0.125
	2	-0.137	-0.076	0.004	0.084	0.193	0.291	0.345	0.374	0.394	0.398	0.352
1	3	-0.201	-0.170	-0.029	0.127	0.259	0.393	0.531	0.590	**0.595**	0.571	0.489
	4	-0.065	0.011	0.132	0.247	0.370	0.472	0.573	0.630	0.676	**0.682**	0.640
	5	-0.394	-0.388	-0.335	-0.257	-0.135	-0.026	0.086	0.184	0.240	0.293	0.336
	6	0.321	0.478	0.640	0.762	0.845	0.901	**0.928**	0.923	0.885	0.825	0.750
	7	0.326	0.467	0.629	0.768	0.871	0.935	**0.958**	0.942	0.890	0.822	0.741
2	8	0.363	0.515	0.677	0.806	0.890	0.939	**0.953**	0.931	0.881	0.814	0.735
	9	0.376	0.528	0.684	0.806	0.886	0.931	**0.939**	0.913	0.862	0.795	0.720
	10	0.388	0.540	0.695	0.820	0.903	0.946	**0.954**	0.930	0.881	0.813	0.736
	11	0.176	0.322	0.477	0.622	0.735	0.819	0.870	**0.883**	0.866	0.810	0.733
	12	0.240	0.408	0.575	0.723	0.838	0.920	0.963	**0.968**	0.942	0.885	0.807
3	13	0.293	0.460	0.620	0.754	0.853	0.918	**0.948**	0.941	0.902	0.842	0.767
	14	0.324	0.473	0.618	0.749	0.846	0.911	**0.937**	0.928	0.893	0.829	0.747
	15	0.490	0.583	0.667	0.733	0.769	**0.776**	0.749	0.689	0.606	0.491	0.355
	16	0.562	0.692	0.780	0.847	0.896	0.926	**0.936**	0.928	0.909	0.876	0.831
	17	0.544	0.663	0.753	0.825	0.884	0.926	0.950	**0.953**	0.945	0.926	0.898
4	18	0.605	0.713	0.785	0.835	0.873	0.899	**0.909**	0.903	0.885	0.853	0.814
	19	0.563	0.700	0.797	0.868	0.906	0.926	**0.928**	0.907	0.873	0.831	0.783
	20	0.550	0.676	0.764	0.823	0.861	0.888	**0.901**	0.897	0.875	0.838	0.790

Table 6-2 Correlation (r) between mean RFE and mean NDVI for 20 pixels, mean values are calculated from 1995 to 2008. Value in bold is the maximum and significant $\alpha < 0.001$

From Table 6-2, two pixels in study region 1 has shown a significant correlation between mean NDVI and rainfall accumulation for around last past 90 days. According to the analysis of Table 6-1, the vegetation density of pixel 3 and 4 was higher than in the time period 2004-2008. The mean NDVI value of pixel 3 is in the range [0.092, 0.1], pixel 4 in [0.088, 0.096]. Also the rainfall values of these two pixels do not change a lot every dekadal, which is in the range [0, 14] (mm). But, three other pixels show non-significant correlation between NDVI and rainfall in both Table 6-1 and Table 6-2. The mean NDVI value of these three pixel varied bigger than pixel 3 and 4, thought they have the almost the same rainfall amount in each dekadal. Compared to the analysis of Table 6-1, although the low vegetation area sometimes shows the correlation between NDVI and rainfall, the reliability of these values should be further study analyzed. In region2, the correlation is the highest among all four study regions. All of the pixels show the highest value above 0.9 between NDVI and rainfall accumulation of the last 70 days. Both of NDVI and rainfall value follow the two-season rainfall pattern in the study area. The range of mean NDVI in region 2 is [0.14, 0.364] and the annual amount of rainfall is around 380mm. In region 3 and region 4, the correlation patterns are similar in the two tables. The highest correlation value is with the rainfall accumulated of the almost last three month.

By comparing Table 6-1 and Table6-2, it can be observed that the patterns of maximum correlation value are similar. Both of the tables show that there is a strong relationship between NDVI and almost three previous month's rainfall, indicating a lagged response(Davenport & Nicholson, 1993). It also suggests that NDVI can be an indicator of vegetative drought. However, values in Table 6-2 are much higher than those in Table 6-1.The mean values have much stronger relationship than recent year values. The severe drought reported in Kenya in year 2004, early 2006 and 2008, can explain the low relationship. As after sever drought, the recovery time can take up to several months to be back to normal situation. Another reason could be the artificial change of land cover. All the study regions are selected near the big city in Kenya. During city development, more settles and facilities are build to support the increasing population whom occupied the original place of vegetation. These human behaviours add more uncertainty in vegetation dynamics not only dependant of rainfall amount.

6.2. Drought classification based on NDVI anomaly

The anomaly NDVI data is calculated for every 10-day with the average value from 20 pixels from 2004-2008. The entire observed number of sample is 3600. Figure5-2 shows the NDVI anomaly distribution of the dataset, with the maximum anomaly 0.432 and the minimum -0.376. The NDVI values in GIMMS are converted from [0, 250] to [0, 1].

Figure 6-1shows the histogram of NDVI anomaly for all 20 pixels and the black curve line shows the density distribution of the values. The number of positive anomaly and negative anomaly are almost equivalent. In section 5.2.1, the research applied 6 drought classes in referring to SPI class number: extreme, severe drought, moderate drought, dryness, wet, moderately wet, which denoted by class1, class2, class3, class4, class5 and class6. The sample is divided into 6 vegetative drought classes with equal frequencies to ensure that the number per class is sufficiently high to obtain reasonable estimates. There are 4 drought classes and 2 non-drought classes; also the data distributed is almost normal, equal frequencies here has not the exactly "equal" meaning. In practical, the input data was originally divided into 8 classes with 4 drought classes and 4 non-drought classes, each class has around 450 values. Because the focus should be in drought classes, there is a combination of three non-drought classes into one. Here is the way to get 6 classes. This is however subject to modification according to different needs. The core zone of each fuzzy class is base on the 95% confidence interval, with details in Section 5.2.2.

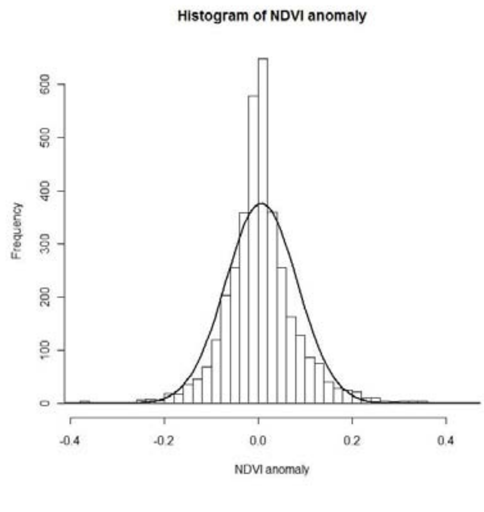


Figure 6-1 Hisogram of NDVI anomaly from the period 2004 to 2008 of all 20 pixels

The core zone of class1 is in the range [min, -0.072], Class 2: [-0.060, -0.036], Class 3: [-0.024, -0.016], Class 4: [-0.004, 0], Class 5: [0.008, 0.012], Class 6: [0.024, max], where min and max represent the minimum and maximum values of the whole dataset. Table 6-3 shows clearly the fuzzy parameters, which is applied to the function as mentioned in Section 5.2.2. Figure 6-2 shows the membership function of drought classes. First a membership function is applied, and the pixel is assigned the class of membership value, as shown in Figure 6-3. In the early 2006, there was reported drought almost all over Kenya. It also can be seen from the figure, in the first month in 2006, all the regions are classified in to drought classes.

Drought category	Class index *i*	a_i	b_i
Extremely drought	1	-0.376	-0.072
Severely drought	2	-0.060	-0.036
Moderately drought	3	-0.024	-0.016
Dryness	4	-0.004	0.000
Wet	5	0.008	0.012
Moderately wet	6	0.024	0.432

Table 6-3 Relation between drought category and drought index, *i* denoted by drought index and [a_i, b_i] represented the core zone of each class

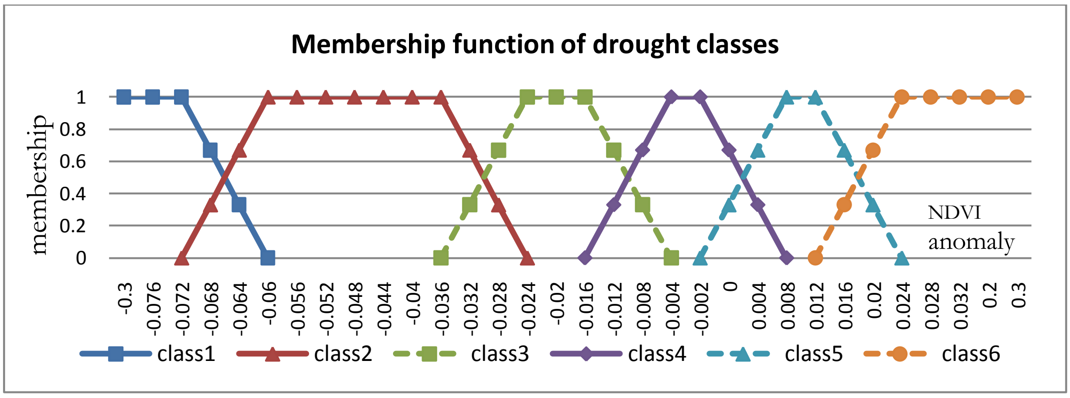


Figure 6-2 Membership function of drought classes

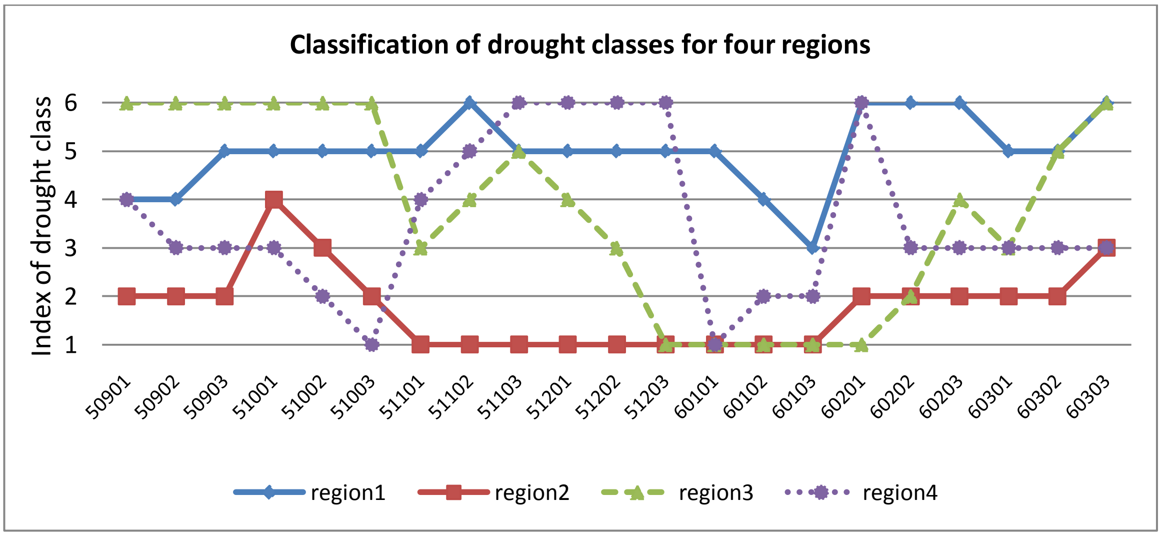


Figure 6-3 Classification of drought classes for four regions of every dekadal from September 2005 to March 2006

6.3. Markovian property tests

6.3.1. MC Transition matrix

Based on "equal" frequency classification, the entire estimated transition matrix is shown in Table 6-4 and in Table 6-5 is shown the estimated transition matrix for 20 pixel calculated for the period from 2004-2008. The process under consideration is a stationary first-order Markov chain.

Initial distribution	
Abs.	Rel.
397	0.11
387	0.11
355	0.10
416	0.12
316	0.09
1709	0.48

Table 6-4 Distribution of entire estimated transition matrix

	Transition Probabilities (*t-1* to *t*)					
classes	1	2	3	4	5	6
1	**0.574**	0.196	0.071	0.035	0.025	0.098
2	0.202	**0.390**	0.189	0.083	0.023	0.114
3	0.101	0.203	**0.304**	0.172	0.034	0.186
4	0.036	0.072	0.161	**0.385**	0.132	0.214
5	0.022	0.019	0.051	0.203	**0.351**	0.354
6	0.019	0.029	0.039	0.050	0.069	**0.795**
	Limiting distribution					
	0.111	0.108	0.100	0.116	0.088	0.477

Table 6-5 Estimated transition matrix for 20 pixel calculated for the period from 2004-2008

6.3.2. Test of time homogeneity

In this study, the sub-samples are divided following two steps. In the first step the samples are divided by the year: M = {2004, 2005, 2006, 2007, 2008}, in total 5 different transition matrices are estimated. The result is shown in Table 6-6, where Q and LR represent the result of Q and LR test, *df* is the degree of freedom and Qprob and LRprob represent the probability of χ^2 distribution, the significant $\alpha < 0.05$. All of the tests are fail to reject that this Markov chain is homogeneity of each year.

Year	Q	LR	df	Qprob	LRprob
2004	12.4	15.8	30	0.998	0.985
2005	31.1	27.2	30	0.411	0.615
2006	34.0	26.1	30	0.282	0.672
2007	35.8	43.3	30	0.215	0.055
2008	16.0	23.3	30	0.983	0.805
Sum	129.2	135.6	120	0.267	0.157

Table 6-6 Pearson and Likelihood ration tests of time homogeneity yearly

In the second step, the samples are divided monthly: M = 12 different transition matrices are estimated. The result is shown in Table 6-7. There are two growing seasons on the study area: March to June and September to December. Monthly division is applied to estimate whether the transition matrix is constant for both growing seasons and non-growing seasons. The result of February and September are less than 0.05, which reject the homogeneity of Markov chain. The rainy seasons in Kenya are from March to June and October to December, with details given in Section 3.2. Since both February and December are the last month before rainy season comes, dynamics of vegetation is in the different growing pattern from rest of the year. It is affected by more factors than during the growing seasons. These effects disturb the homogeneity of Markov chain(Balzter, 2000). This can be one of the reasons why they are not in the homogeneity as other months.

Month	Q	LR	df	Qprob	LRprob
Jan	29.8	34.0	30	0.477	0.281
Feb	46.6	64.6	30	0.027	0.000
Mar	35.2	44.1	30	0.236	0.047
Apr	26.4	34.1	30	0.655	0.277
May	36.0	47.6	30	0.209	0.022
Jun	26.0	28.9	30	0.674	0.521
Jul	29.6	37.2	30	0.489	0.171
Aug	31.4	36.6	30	0.395	0.190
Sep	45.3	58.7	30	0.036	0.001
Oct	37.2	44.9	30	0.171	0.039
Nov	25.4	38.1	30	0.706	0.146
Dec	32.3	38.1	30	0.353	0.147
Sum	401.1	507.0	330	0.004	1.23E-09

Table 6-7 Pearson and Likelihood ration tests of time homogeneity monthly

Remove February and September, re-calculated the estimated transition matrix and re-test time homogeneity monthly, the result is shown in Table 6-8. Both of the Qprob and LRprob of the rest 10 months are failed to reject the homogeneity property, after excluding the February and September from the data sample. The Pearson test statistic is 225.1(Qprob = 0.978, *df* = 270), and the LR statics is 268.3 (LPprob = 0.517). Neither of the tests indicates a statistically significant in homogeneity. It means that without these two months, the estimated transition matrix fits the Markovian property of time homogeneity and can be used for further process. In the next section, this transition matrix is applied for the test of time independence. The estimated transition matrix without February and September is shown as below in Table 6-9.

Month	Q	LR	*df*	Qprob	LRprob
Jan	27.1	25.2	30	0.618	0.713
Mar	18.5	26.3	30	0.950	0.659
Apr	16.5	22.3	30	0.979	0.843
May	17.7	16.9	30	0.964	0.973
Jun	18.3	19.6	30	0.953	0.928
Jul	22.9	29.5	30	0.820	0.492
Aug	26.9	29.6	30	0.630	0.487
Oct	29.0	32.7	30	0.519	0.336
Nov	24.3	36.9	30	0.757	0.179

Dec	24.0	29.3	30	0.772	0.504
Sum	225.1	268.3	270	0.978	0.517

Table 6-8 Pearson and Likelihood ration tests of time homogeneity monthly without February and September

	Initial Distribution		Transition Probabilities (t-*1* to *t*)					
t-1	Abs.	Rel.	1	2	3	4	5	6
1	337	0.09	**0.596**	0.211	0.059	0.027	0.033	0.074
2	299	0.08	0.190	**0.434**	0.175	0.066	0.020	0.115
3	255	0.07	0.108	0.194	**0.324**	0.144	0.072	0.158
4	311	0.09	0.021	0.067	0.163	**0.351**	0.247	0.151
5	230	0.06	0.016	0.025	0.061	0.141	**0.531**	0.227
6	1348	0.38	0.018	0.036	0.032	0.021	0.099	**0.793**
Limiting distribution			0.119	0.106	0.091	0.111	0.082	0.490

Table 6-9 Estimated transition matrix for 20 pixel calculated for the period from 2004-2008 without February and September

6.3.3. Test of time independence

From the time-homogeneous dataset in Section 6.3.2, 2780 observations are available for test of Markovity of order 0 against order 1. As the transition matrix shown in Table 6-9 and applied equation 5-12 and 5-13, the order 0 is tested by comparing each row of estimated transition with the limiting distribution at time *t*. The results of Qprob and LRprob, Q = 800.5569, LR = 693.3868, *df* = 25, are much smaller than significant lever α. The results reject the Markov chain of order 0 and indicate that the process strongly depends on the past. For the test of Markovity of first-order against second-order, there are 2560 observations are available. The result is shown in Table 6-10, applied by the equation 5-14 and 5-15. Six subsamples $h = \{1, 2, 3, 4, 5, 6\}$ are defined, representing the drought class at time *t-2*. Both statistic tests result in Qprob=0.207 and LRprob = 0.200, with Q = 163.9, LR = 164.3 and 150 degree of freedom. The result illustrates that the process is in the first-order as consideration and can be applied in further prediction.

Class in time *t-2*	Q	LR	*df*	Qprob	LRprob
1	30.9	30.2	25	0.193	0.216
2	34.3	29.9	25	0.102	0.229
3	20.2	24.0	25	0.736	0.521
4	27.5	25.3	25	0.331	0.446
5	25.1	25.6	25	0.456	0.431
6	25.9	29.4	25	0.413	0.248
Sum	163.9	164.3	150	0.207	0.200

Table 6-10 Pearson and Likelihood ration tests of first-order against second-order

After all the pre-work of correlation between rainfall and NDVI, and the Markov property test, the results shows that the transition matrix in Table 6-9 is well modelling the dynamic of vegetative drought, which suggests that it could be used in vegetative drought prediction. The fix row vector **v** and mean first passage time **M** is shown in Table 6-11. The values in first mean passage matrix shows how many dekadals it takes to transfer from one class of drought to another class of drought. For example, it will take about

110 days for all the vegetation in class 1 improving to drought class 2. And also it needa more than 80 days for the vegetation original in class 1 first return back to class 1.

Drought class	1	2	3	4	5	6
v	0.113	0.103	0.090	0.110	0.083	0.501
	Mean first passage time (dekadal)					
1	**8.88**	11.31	14.05	14.61	20.00	7.36
2	15.89	**9.70**	12.16	13.39	19.72	6.77
3	19.69	13.58	**11.17**	11.89	18.51	5.82
4	22.68	16.57	13.83	**9.07**	16.26	5.06
5	24.48	18.57	16.74	11.71	**12.05**	3.84
6	24.87	18.85	17.70	14.77	17.21	**1.99**

Table 6-11 fix row vector **a** and first mean passage matrix

6.4. Prediction

In this research, the fuzzy Markov chain is used to predict the first dekadal in January of 2009 from the known data of last dekadal of December in 2008. Table 6-12 shows the membership value of 20 pixels, using the membership functions with parameters shown in Table 6-3 and the prediction value of 2009 is shown in Table 6-13. The maximum membership value in Table 6-12 and probability value in Table 6-13 , shown in bold, represent the highest probability of drought class for each pixel, and then the pixel is assigned to that class. Compared the prediction result to the real satellite image result shown in Table 6-14, 7 out of 20 pixels are assigned to the same class and the average RMSE of 20 pixel = 0.304. All the pixels in region 1 are well predicted. In region 2, 4 out of 5 pixels are totally wrong. In region 3 and region4, most of the pixels are predicted in neighbouring drought class.

Region	Pixel	class1	class2	class3	class4	class5	class6
1	1	0.000	0.000	0.000	0.000	0.000	**1.000**
	2	0.000	0.000	0.000	0.000	0.000	**1.000**
	3	0.000	0.000	0.000	0.000	**0.667**	0.333
	4	0.000	0.000	0.000	0.000	**1.000**	0.000
	5	0.000	0.000	0.000	0.000	**0.667**	0.333
2	6	0.000	0.000	0.000	0.000	0.000	**1.000**
	7	0.000	0.000	**1.000**	0.000	0.000	0.000
	8	0.000	0.000	0.000	**1.000**	0.000	0.000
	9	0.000	0.000	0.000	0.000	0.000	**1.000**
	10	0.000	0.000	0.000	**1.000**	0.000	0.000
3	11	0.333	**0.667**	0.000	0.000	0.000	0.000
	12	**1.000**	0.000	0.000	0.000	0.000	0.000
	13	**1.000**	0.000	0.000	0.000	0.000	0.000
	14	**1.000**	0.000	0.000	0.000	0.000	0.000
	15	**1.000**	0.000	0.000	0.000	0.000	0.000
4	16	0.000	**1.000**	0.000	0.000	0.000	0.000
	17	0.000	**1.000**	0.000	0.000	0.000	0.000
	18	**1.000**	0.000	0.000	0.000	0.000	0.000
	19	0.000	0.000	**0.667**	0.333	0.000	0.000
	20	**1.000**	0.000	0.000	0.000	0.000	0.000

Table 6-12 membership value of 20 pixels of the last dekadal in 2008, maximum value is in bold

Region	Pixel	class1	class2	class3	class4	class5	class6
1	1	0.018	0.036	0.032	0.021	0.099	**0.793**
	2	0.018	0.036	0.032	0.021	0.099	**0.793**
	3	0.017	0.029	0.051	0.101	0.387	**0.416**
	4	0.016	0.025	0.061	0.141	**0.531**	0.227
	5	0.017	0.029	0.051	0.101	0.387	**0.416**
2	6	0.018	0.036	0.032	0.021	0.099	**0.793**
	7	0.108	0.194	**0.324**	0.144	0.072	0.158
	8	0.021	0.067	0.163	**0.351**	0.247	0.151
	9	0.018	0.036	0.032	0.021	0.099	**0.793**
	10	0.021	0.067	0.163	**0.351**	0.247	0.151
3	11	0.325	**0.360**	0.136	0.053	0.024	0.101
	12	**0.596**	0.211	0.059	0.027	0.033	0.074
	13	**0.596**	0.211	0.059	0.027	0.033	0.074
	14	**0.596**	0.211	0.059	0.027	0.033	0.074
	15	**0.596**	0.211	0.059	0.027	0.033	0.074
4	16	0.190	**0.434**	0.175	0.066	0.020	0.115
	17	0.190	**0.434**	0.175	0.066	0.020	0.115
	18	**0.596**	0.211	0.059	0.027	0.033	0.074
	19	0.079	0.152	**0.270**	0.213	0.130	0.156
	20	**0.596**	0.211	0.059	0.027	0.033	0.074

Table 6-13 Predict probability value of 20 pixels of the first dekadal in 2009, maximum value is in bold

Region	Pixel	class1	class2	class3	class4	class5	class6
1	1	0.000	0.000	0.000	0.000	0.000	**1.000**
	2	0.000	0.000	0.000	0.000	0.000	**1.000**
	3	0.000	0.000	0.000	0.000	0.333	**0.667**
	4	0.000	0.000	0.000	0.000	**1.000**	0.000
	5	0.000	0.000	0.000	0.000	0.333	**0.667**
2	6	0.000	0.000	0.000	0.000	0.333	**0.667**
	7	0.000	**1.000**	0.000	0.000	0.000	0.000
	8	0.000	**1.000**	0.000	0.000	0.000	0.000
	9	0.000	**1.000**	0.000	0.000	0.000	0.000
	10	0.000	**0.667**	0.333	0.000	0.000	0.000
3	11	**1.000**	0.000	0.000	0.000	0.000	0.000
	12	0.000	0.000	**1.000**	0.000	0.000	0.000
	13	**1.000**	0.000	0.000	0.000	0.000	0.000
	14	0.000	**1.000**	0.000	0.000	0.000	0.000
	15	0.000	**1.000**	0.000	0.000	0.000	0.000
4	16	**1.000**	0.000	0.000	0.000	0.000	0.000
	17	**1.000**	0.000	0.000	0.000	0.000	0.000
	18	0.000	0.000	0.000	0.000	0.000	**1.000**
	19	0.000	0.333	**0.667**	0.000	0.000	0.000
	20	0.000	0.000	0.000	0.000	0.000	**1.000**

Table 6-14 membership value of 20 pixels of the first dekadal in 2009, maximum value is in bold

Therefore, adjust slope of membership function is under consideration so as to improve the accuracy of perdition. After several trials, because of time limitation, the best slope parameters among all the trails are shown in Table 6-15.

Drought category	Class index i	a_i	b_i
Extremely drought	1	-0.376	-0.300
Severely drought	2	-0.039	-0.032
Moderately drought	3	-0.020	-0.014
Dryness	4	-0.004	0.000
Wet	5	0.012	0.014
Moderately wet	6	0.030	0.432

Table 6-15 Relation between drought category and drought index after adjust, i denoted by drought index and $[a_i, b_i]$ represented the core zone of each class

The prediction value of first dekadal in 2009 is shown in Table 6-16. Compared the prediction result to the real satellite image result shown in Table 6-17, as reference data, 11 out of 20 pixels are assigned to the same class and the average RMSE = 0.260. In region 1 and region 3, more pixels are in good performance. After adjusted by using smaller core zone and bigger transition zone for each class, the prediction shows better result as more pixel are predict into correct drought classes, such as pixel 7. Before adjustment, the maximum probability of prediction of pixel 7 appears in drought class 3, but the validation drought class is class 2. Expanding the transition zone between class 2 and class3, the maximum probability of pixel 7 appears in drought class 2 and the reference data has high membership value in class2. By changing the size of core zones or slope of membership functions, the results change and may improve the accuracy.

Region	Pixel	class1	class2	class3	class4	class5	class6
1	1	0.018	0.036	0.032	0.021	0.099	**0.793**
	2	0.018	0.036	0.032	0.021	0.099	**0.793**
	3	0.016	0.026	0.058	0.127	**0.480**	0.294
	4	0.016	0.025	0.061	0.141	**0.531**	0.227
	5	0.016	0.026	0.058	0.127	**0.480**	0.294
2	6	0.018	0.036	0.032	0.021	0.099	**0.793**
	7	0.135	**0.274**	0.273	0.119	0.055	0.144
	8	0.021	0.067	0.163	**0.351**	0.247	0.151
	9	0.018	0.036	0.032	0.021	0.099	**0.793**
	10	0.021	0.067	0.163	**0.351**	0.247	0.151
3	11	0.236	**0.409**	0.162	0.062	0.021	0.110
	12	0.327	**0.359**	0.136	0.053	0.024	0.101
	13	0.321	**0.362**	0.137	0.053	0.024	0.102
	14	0.346	**0.348**	0.131	0.051	0.025	0.099
	15	**0.413**	0.312	0.111	0.045	0.027	0.092
4	16	0.193	**0.432**	0.174	0.066	0.020	0.115
	17	0.193	**0.432**	0.174	0.066	0.020	0.115
	18	0.260	**0.395**	0.155	0.059	0.022	0.108
	19	0.091	0.169	**0.292**	0.185	0.107	0.157
	20	0.266	**0.392**	0.153	0.059	0.022	0.107

Table 6-16 adjusted predict probability value of 20 pixels of the first dekadal in 2009, maximum value is in bold

Region	Pixel	class1	class2	class3	class4	class5	class6
1	1	0.000	0.000	0.000	0.000	0.000	**1.000**
	2	0.000	0.000	0.000	0.000	0.000	**1.000**
	3	0.000	0.000	0.000	0.000	**0.647**	0.353
	4	0.000	0.000	0.000	0.000	**1.000**	0.000
	5	0.000	0.000	0.000	0.000	**0.647**	0.353
2	6	0.000	0.000	0.000	0.000	**0.647**	0.353
	7	0.038	**0.962**	0.000	0.000	0.000	0.000
	8	0.083	**0.917**	0.000	0.000	0.000	0.000
	9	0.008	**0.992**	0.000	0.000	0.000	0.000
	10	0.000	**1.000**	0.000	0.000	0.000	0.000
3	11	0.173	**0.827**	0.000	0.000	0.000	0.000
	12	0.000	0.000	**1.000**	0.000	0.000	0.000
	13	0.398	**0.602**	0.000	0.000	0.000	0.000
	14	0.068	**0.932**	0.000	0.000	0.000	0.000
	15	0.068	**0.932**	0.000	0.000	0.000	0.000
4	16	0.444	**0.556**	0.000	0.000	0.000	0.000
	17	0.474	**0.526**	0.000	0.000	0.000	0.000
	18	0.000	0.000	0.000	0.000	0.000	**1.000**
	19	0.000	**0.667**	0.333	0.000	0.000	0.000
	20	0.000	0.000	0.000	0.000	0.000	**1.000**

Table 6-17 adjusted membership value of 20 pixels of the first dekadal in 2009, maximum value is in bold

Prediction of the second and third dekadals of January, 2009, are separately shown in the Table 6-18 and Table 6-19, and the maximum probability value are also shown in bold. It suggests that, most of the pixels are in drought situation and only pixel 1, 2 and 5 are not drought in the coming two dekadals. In region 2, all the pixels are in drought class 2, according to Table 6-11, normally, it will more than half year for the vegetation to recover from this drought class. Long-term drought response plan should be applied in to this region. In region 1, the situation getting worse and most of the pixels would go to the drought class 1. The same situation occurs in region 4.

Region	Pixel	class1	class2	class3	class4	class5	class6
1	1	0.018	0.036	0.032	0.021	0.099	**0.793**
	2	0.018	0.036	0.032	0.021	0.099	**0.793**
	3	0.017	0.029	0.051	0.099	0.379	**0.427**
	4	0.016	0.025	0.061	0.141	**0.531**	0.227
	5	0.017	0.029	0.051	0.099	0.379	**0.427**
2	6	0.017	0.029	0.051	0.099	0.379	**0.427**
	7	0.205	**0.426**	0.171	0.065	0.020	0.113
	8	0.224	**0.416**	0.165	0.063	0.021	0.112
	9	0.193	**0.432**	0.174	0.066	0.020	0.115
	10	0.190	**0.434**	0.175	0.066	0.020	0.115
3	11	0.260	**0.395**	0.155	0.059	0.022	0.108
	12	0.108	0.194	**0.324**	0.144	0.072	0.158
	13	0.352	**0.345**	0.129	0.050	0.025	0.099
	14	0.217	**0.419**	0.167	0.063	0.021	0.112

	15	0.217	**0.419**	0.167	0.063	0.021	0.112
4	16	**0.370**	0.335	0.124	0.049	0.026	0.097
	17	**0.382**	0.328	0.120	0.048	0.026	0.096
	18	0.018	0.036	0.032	0.021	0.099	**0.793**
	19	0.163	**0.354**	0.225	0.092	0.037	0.129
	20	0.018	0.036	0.032	0.021	0.099	**0.793**

Table 6-18 adjusted predict probability value of 20 pixels of the second dekadal in 2009, maximum value is in bold

Region	Pixel	class1	class2	class3	class4	class5	class6
1	1	0.018	0.034	0.037	0.042	0.175	**0.693**
	2	0.017	0.031	0.044	0.070	0.277	**0.560**
	3	0.021	0.067	0.163	**0.351**	0.247	0.151
	4	0.021	0.067	0.163	**0.351**	0.247	0.151
	5	0.016	0.025	0.061	0.141	**0.531**	0.227
2	6	0.190	**0.434**	0.175	0.066	0.020	0.115
	7	0.254	**0.399**	0.157	0.060	0.022	0.109
	8	0.291	**0.379**	0.146	0.056	0.023	0.105
	9	0.199	**0.429**	0.172	0.065	0.020	0.114
	10	0.224	**0.416**	0.165	0.063	0.021	0.112
3	11	**0.565**	0.228	0.068	0.030	0.032	0.077
	12	**0.596**	0.211	0.059	0.027	0.033	0.074
	13	**0.596**	0.211	0.059	0.027	0.033	0.074
	14	**0.596**	0.211	0.059	0.027	0.033	0.074
	15	0.340	**0.352**	0.132	0.052	0.025	0.100
4	16	**0.388**	0.325	0.118	0.047	0.026	0.095
	17	**0.419**	0.308	0.110	0.044	0.027	0.092
	18	0.205	**0.426**	0.171	0.065	0.020	0.113
	19	0.279	**0.385**	0.150	0.057	0.023	0.106
	20	0.217	**0.419**	0.167	0.063	0.021	0.112

Table 6-19 adjusted predict probability value of 20 pixels of the third dekadal in 2009, maximum value is in bold

6.5. Validation

The validation compares the predicted result to the reality image result. P_{class} represents the maximum probability value of prediction while V_{class} represents the maximum membership value of validation. Then assigned the drought class of each pixel with maximum value to Pred (prediction class) and Vad.

The Comparison of the predicted result and reality image in the first dekadal of January in 2009, shown in Table 6-17, is shown in Figure 6-4. In the region1, almost all the pixels are correctly classification. In region 2, there is a big different. Checking the class of last dekadal in 2008, there are class jumps from 2008 to 2009. As we know the correlation between rainfall and NDVI, check the accumulated 70 days previous, the rainfall anomaly is around -68 mm. Also, December should be the rainy season in study region, which means the rainfall anomaly affected the result of prediction in the first 10-days in 2009. In this prediction, 11 out of 20 pixels are correctly predicted.

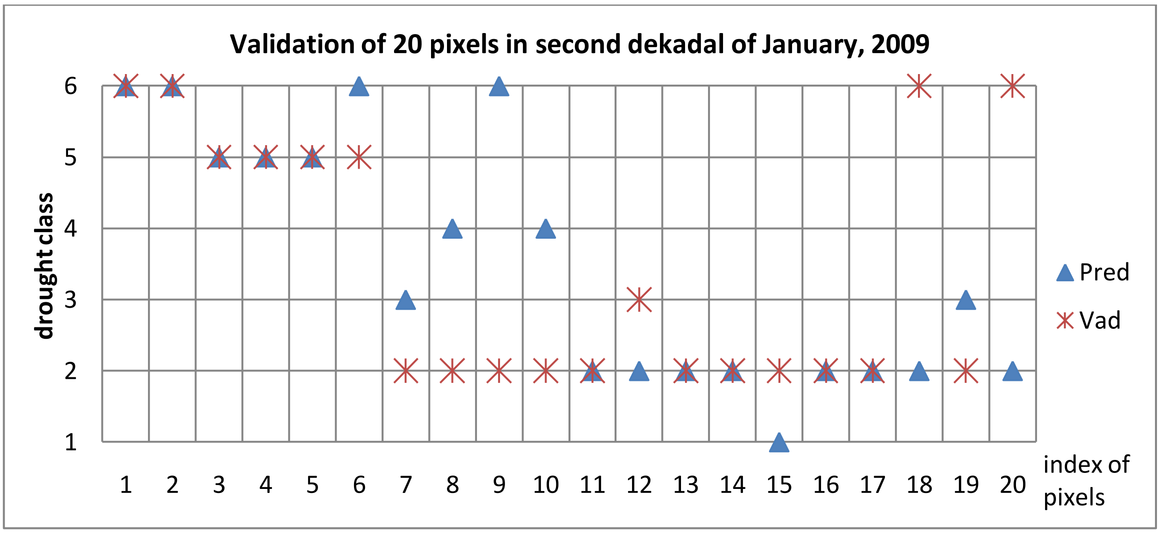


Figure 6-4 Validation drought class of 20 pixels of the first dekadal in 2009

The Comparison of the predicted result and reality image in the second dekadal of January in 2009, shown in Table 6-20, is shown in Figure 6-5. There are 3 pixels miss-classified in to non-drought class which from the validation, they should be in drought. 12/20 pixels well predict the drought, while 10/20 pixels have the right drought class.

Region	Pixel	class1	class2	class3	class4	class5	class6
1	1	0.000	0.000	0.000	0.000	0.176	**0.824**
	2	0.000	0.000	0.000	0.000	0.412	**0.588**
	3	0.000	0.000	0.000	**1.000**	0.000	0.000
	4	0.000	0.000	0.000	**1.000**	0.000	0.000
	5	0.000	0.000	0.000	0.000	1.000	0.000
2	6	0.000	**1.000**	0.000	0.000	0.000	0.000
	7	0.158	**0.842**	0.000	0.000	0.000	0.000
	8	0.248	**0.752**	0.000	0.000	0.000	0.000
	9	0.023	**0.977**	0.000	0.000	0.000	0.000
	10	0.083	**0.917**	0.000	0.000	0.000	0.000
3	11	**0.925**	0.075	0.000	0.000	0.000	0.000
	12	**1.000**	0.000	0.000	0.000	0.000	0.000
	13	**1.000**	0.000	0.000	0.000	0.000	0.000
	14	**1.000**	0.000	0.000	0.000	0.000	0.000
	15	0.368	**0.632**	0.000	0.000	0.000	0.000
4	16	0.489	**0.511**	0.000	0.000	0.000	0.000
	17	**0.564**	0.436	0.000	0.000	0.000	0.000
	18	0.038	**0.962**	0.000	0.000	0.000	0.000
	19	0.218	**0.782**	0.000	0.000	0.000	0.000
	20	0.068	**0.932**	0.000	0.000	0.000	0.000

Table 6-20 adjusted membership value of 20 pixels of the second dekadal in 2009, maximum value is in bold

The prediction has better performance in region 2 than any other regions. In pixel 18 and 20, the prediction class is moderately wet but in validation they should be severe drought. Look back the NDVI anomaly in the January 2009, the values show no continuity as other pixel in the same region. We may assume that either there are errors in NDVI value or suddenly land cover changing.

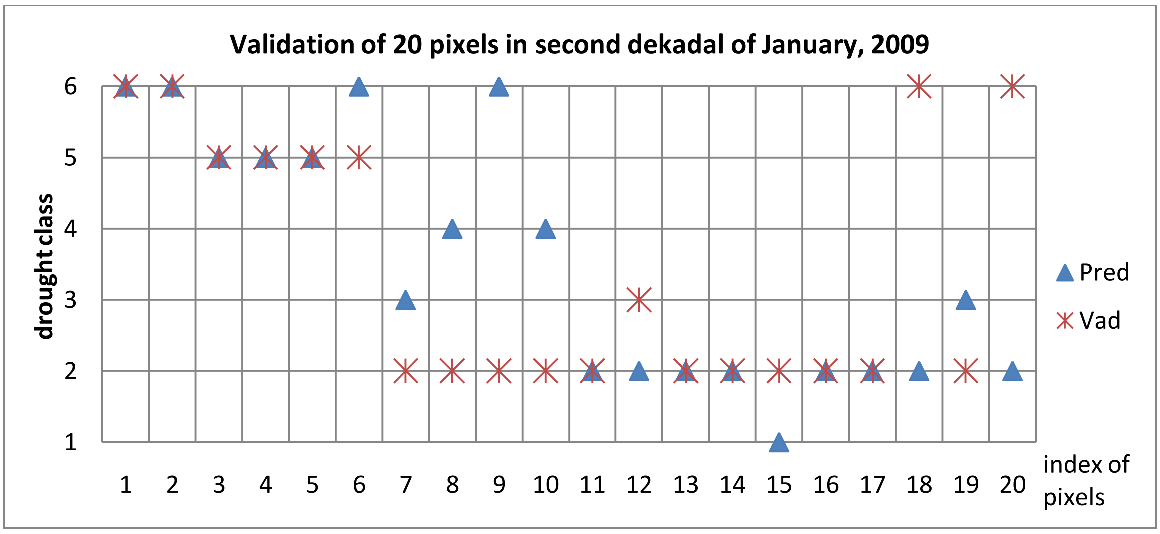


Figure 6-5 Validation drought class of 20 pixels of the second dekadal in 2009

In Figure 6-6, it validates the prediction of third dekadal in January, 2009 to the reference values shown in Table 6-21. There are 11/20 pixels are well predicted in the same drought class as in the validation drought class, 9 pixels miss-predicts to the neighbour drought class and only 1 pixel has 4 step jumps. 15 pixels are well detected the drought situation, thought not all of them are in the correct drought class. Pixel 1 and pixel 15 have the fake warning of drought. Although the drought classes in pixel 17 and 19 are not the same as prediction, the membership values of these two pixels in class 1 and class 2 are almost equal.

Region	Pixel	class1	class2	class3	class4	class5	class6
1	1	0.000	0.000	0.000	0.000	0.000	**1.000**
	2	0.000	0.000	0.000	0.000	0.000	**1.000**
	3	0.000	0.000	0.000	0.333	**0.667**	0.000
	4	0.000	0.000	0.000	**1.000**	0.000	0.000
	5	0.000	0.000	0.000	**0.667**	0.333	0.000
2	6	0.008	**0.992**	0.000	0.000	0.000	0.000
	7	0.173	**0.827**	0.000	0.000	0.000	0.000
	8	0.233	**0.767**	0.000	0.000	0.000	0.000
	9	0.218	**0.782**	0.000	0.000	0.000	0.000
	10	0.158	**0.842**	0.000	0.000	0.000	0.000
3	11	0.158	**0.842**	0.000	0.000	0.000	0.000
	12	0.429	**0.571**	0.000	0.000	0.000	0.000
	13	0.383	**0.617**	0.000	0.000	0.000	0.000
	14	0.383	**0.617**	0.000	0.000	0.000	0.000
	15	0.000	0.000	0.000	0.000	0.000	**1.000**
4	16	**0.609**	0.391	0.000	0.000	0.000	0.000
	17	0.444	**0.556**	0.000	0.000	0.000	0.000
	18	0.444	**0.556**	0.000	0.000	0.000	0.000
	19	**0.579**	0.421	0.000	0.000	0.000	0.000
	20	0.023	**0.977**	0.000	0.000	0.000	0.000

Table 6-21 adjusted membership value of 20 pixels of the third dekadal in 2009, maximum value is in bold

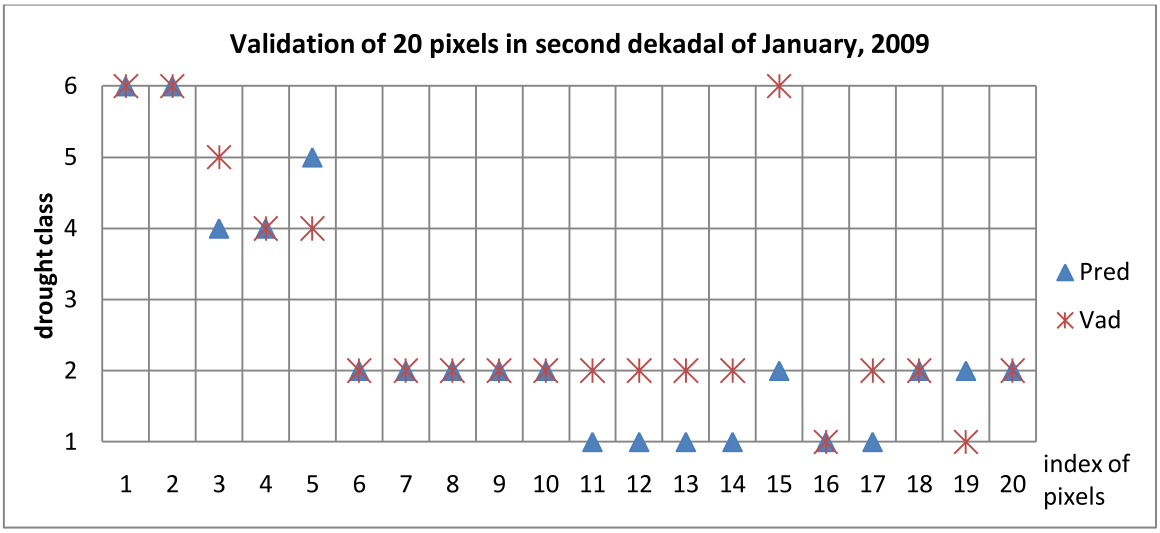


Figure 6-6 Validation drought class of 20 pixels of the third dekadal in 2009

Pixels in the same region as pixel 15 are miss-predicted into neighbour class while pixel 15 jumped from class 2 to class6. Check the original NDVI value of region 3, shown in Table 6-22. In the second dekadal of January, the NDVI values significantly decreased and then, in the next 10 days, the values increased extremely. Although it is not surprise to see decreasing of NDVI as January is not in the rainy season, dramatically flicked should be considered as some error in the satellite images.

pixel	11	12	13	14	15
20080123	0.584	0.516	0.492	0.500	0.404
20090101	0.552	0.544	0.436	0.516	0.484
20090102	0.328	0.092	0.164	0.148	0.384
20090103	0.508	0.376	0.384	0.376	0.620

Table 6-22 original NDVI value of all pixels in region3 from last dekadal of 2008 to the third dekadal of 2009, first column shows the time, in format year/month/index of dekadal

In the Figure 6-7, RMSE values of prediction are shown. Three different shapes represent different dekadals in January. In region 1, most of the values are less than 0.3, which shows pixels in regions 1 can be well predicted with this method. In region 2, pixel 6 to pixel 10, the results in the first dekadal greater than those in the second and third dekadals. As December is the last month of rainy season, we may assume that, in region 2, this method cannot performance well when great change happen. In region 3, all the values are stable in the range [0.20, 0.40], showing the accuracy of this method. As the analysis of Table 6-22, the method can also be applied in region 3. Pixel 18 and pixel 20 show unpredictable in this method in the first and second dekadal. Rest of the values in region 4 are less than 0.3, indicating that the vegetative classes can be well predicted.

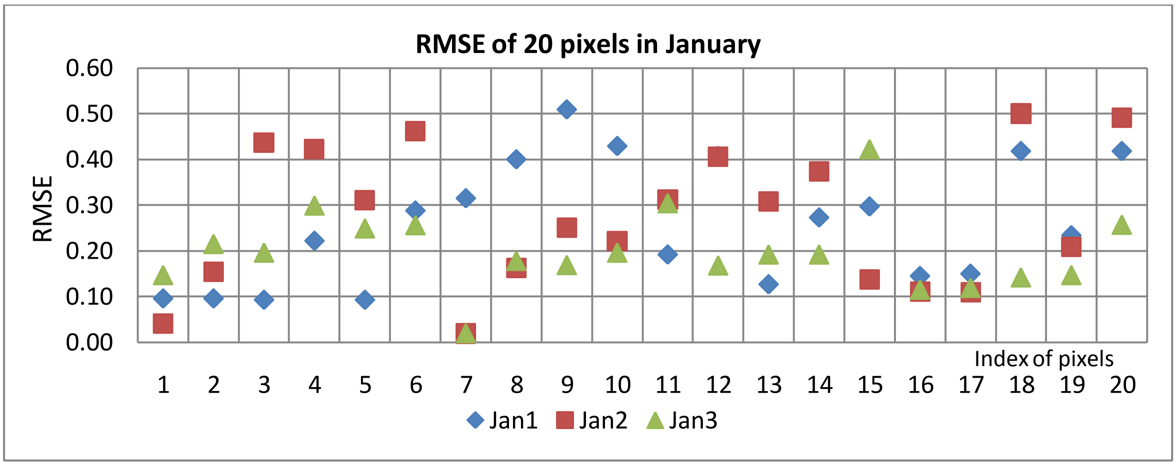


Figure 6-7 RMSE values of 20 pixels in three dekadals in January

7. DICUSSION

7.1. Methodology in this research

In this research, fuzzy Markov chain is applied to predict the dynamics of vegetative drought using NDVI index. As the results shown in the Chapter 6, we may say that this method has a potential in drought prediction. After changing the slope of membership function, it indicates that the accuracy of vegetative drought prediction changes and early warning system can take benefit from it. As time limited, the membership function may not be the optimal in this study area. Hence, we assume that, with trial-and-error, the optimal function could be found and the prediction of vegetative drought can be improved more with fuzzy Markov chain. In this research, the time interval of dataset is 10-day and from year 2004 to 2008. The large amount of data, in total 3600 observations, is suitable to estimate fuzzy Markov chains.

7.1.1. NDVI anomaly

NDVI from satellite images is proved can be used in vegetative drought. For a very long time, drought is measured, modelling, predicted with meteorological data, such as precipitant. However, the density distribution of rainfall gauge station limits the accuracy and continence. Satellite images overcome these limitations. The index, NDVI, has been widely used in every field when it has a relation to vegetation. It can not only give the information about health and density of vegetation, but also contain information direct and indirect to vegetation, such as drought in this research. Normally the meteorological drought is modelled with precipitation data and agricultural drought is modelled with data of soil condition. Since NDVI shows a strong relationship with previous almost three-month precipitation data, it can be used as an indicator of drought. As the vegetation lag precipitation deficiency, it can be a better indicator of soil moisture than of rainfall (Davenport & Nicholson, 1993).

The dataset used in this research is free to obtain on the website. The time period of NDVI is available from 1985 to present. Other aspects of vegetation dynamics can use this dataset for modelling and prediction other than vegetative drought.

7.1.2. Fuzzy classification

Introducing fuzzy set theory into this study makes the drought more reality as a vague object and the result of Markov chain improving. Fuzzy set theory applied as fuzzification to classify the vegetation drought class before applying Markov chain. As NDVI anomaly seldom used in drought study, there is not a standard to classify the drought and related the drought into reality or traditional classification. With the concept of several drought classes and without exact value of each drought class in reality, fuzzy membership functions provide good performance in this situation. Modelling states in Markov chain is the basis, but the crisp states do not fit the property of vegetative drought. Using fuzzification instead of crisp classification makes the states in Markov chain meet the reality and more precise. From the result, we can see that, with the changing of slope of membership functions, the prediction is changing. This fuzzy classification can be further applied in similar area while the study object is vague and crisp classification values are not clear.

7.1.3. Fuzzy Markov chain

Drought is complicated natural disaster compared with others. There are several different factors can affect drought. As it is a relative phenomenon, precipitation deficiency is not the only reason cause drought, the capacity of soil water container, evapo-transpiration which has related to the temperature, and also management of water supplies could cause drought or making the situation worse. Using traditional regression analysis, such as linear regression, many factors should be under consideration. Drought is spatial difference as the needs of water supply, vegetation types vary from region to region. Using linear models, it needs to figure out every affected factor while the model can not be directly applied into other regions where have different vegetation pattern, rainfall pattern or even population density. Compared to it, first-order Markov chain seems to be a simpler approach in modelling and prediction. Applying Markov chain does not need estimate every signal factors in different study region and even not need to find out new factors which may come out in recent years. Modelling Markov chain is based on the time series, which simplifies the effect of each factor and combined them into the time. It saves the time and energy to test various. In this study, Markov chain model the dynamics of vegetative drought in four regions of Kenya. The four study areas have different agricultural types, which all of them proves can be predicted in Markov chains.

7.2. Applications of fuzzy Markov chain

From this study, we can see than fuzzy Markov chain can be applied into vegetative drought prediction. Vegetation drought is most likely to agricultural drought, which we assume that this approach can also applied to agricultural drought. The difference is that agricultural drought most focuses on crops or grass for cattle. Crops have growing seasons and growing pattern, which cannot be modelling through whole year time and should be considered(Banik, et al., 2002). In harvest season, crops can be collected by famers in few days, which do not change gradually as non-agricultural plant does. The fuzzy Markov chain can only be applied during growing season and forecast the yield of crops, because of its property that good performance in vague objects.

Drought is a complex environmental hazard and also a vague object. As well predicting in drought situation provides a potential to let fuzzy Markov chain predict in other environmental issues and geographical phenomena, such as air pollution, climate changing, For example, forest fire modelling need to consider wind speed, wind direction, NDVI of forest and few parameters to make model stable (Umamaheshwaran, Bijker, & Stein, 2007). If we have the enough amounts of data for every minute or even smaller time intervals, the track of fire can be detected. What we need to do is applied the fuzzy level of destroy by fire in small forest area, put the observations in to time series and then using Markov chain to model it. It greatly saving the time of calculates weight for each factor and measure them.

Not only in ecological vague objects, also can the vague objects in other field be considered. In socio-economic area, income convergence has applied to Markov chain theory frequently(Bickenbach & Bode, 2003). As long as large history dataset supported, fuzzy Markov chain can apply as predictive framework.

7.3. Limitations in fuzzy Markov chain

As the fuzzy Markov chain shows a great potential in prediction, the limitation of this research is obvious. First limitation is the time period of NDVI dataset in this research. The NDVI anomaly is the deviation from long-term mean calculation. When the variability in vegetation conditions in a region is very high in any one given year, the mean value can be misinterpreted(Thenkabail, Gamage, & Smakhtin, 2004). In order to improving the accuracy, the longer the mean value can be calculated, the mean is more near the normal value. However, the definition of “long” is tricky. If the land cover has changed dramatically in recent days by human behaviours, the mean value should be calculated just in recent years. If the land cover is changed gradually or just in a regular pattern, the mean should calculated from longer time period.

Furthermore, the fuzzy Markov chain is constructed in 5 year. With expanding or reducing the time period, it could have different trend.

Second, in the study area Kenya, only 20 pixels, grouped into four regions, are considered for the prediction. Although these four regions are in different agricultural pattern and land covers are different, knowing whether they represent all the vegetation growing patterns all over the country needs further study. As in the assumption, vegetation growing is done in a gradually pattern, which can be predicted with historic data using a fuzzy Markov chain. How the different agricultural types fit into the general fuzzy Markov chain should be under consideration. More pixels or regions for the research can solve this limitation. The spatial resolution of a pixel is 8km, covering the area of 64 km^2. It is difficult to validate with ground collected observations, which makes both of the prediction and validation have to solely rely on the satellite images.

Third, the Markov chain is very sensitive to the classification process. Using different approach in classification, equal frequency and equal interval, the property of markovian is changed. For few classification approach used, there was no markovian properties, hence could not be modelled in Markov Chain. That is also why only equal frequency is applied in this study. One of the reason is that the number of observations in each states will affect the result of LR and Q test.

Fourth, although changing the slope of membership function influenced the results, this changing process should have some physical meaning. Although the membership functions are built based on the study data not related to drought report, the results can not be directly applied into early warning system or government response without further analysis of the method on a larger area.

As far as we known, fuzzy Markov chain is good to apply in the gradually class-changing vague objects. The sudden change will not affect Markov model, but affect the prediction. With less required parameters, the model will need to be re-built if the environment suddenly changes dramatically. In drought situation, this can be that people suddenly immigrate which reduce the need of water, new kind of vegetation introduced or global alter by *El* Niño. Slightly environment changes will hardly affect the whole property of fuzzy Markov chain, but need reference data to get accurate prediction result.

8. CONCLUSIONS AND RECOMMENDATION

8.1. Conclusions

Drought is a natural hazard than involves many factors and has complex inter-correlation. Vegetative drought is a term describing the drought affecting by the dynamics of vegetation. This study, NDVI shows a strong relationship with rainfall accumulation for almost three previous months, which reveals that NDVI can be used as vegetative drought indicator. The research applied fuzzy Markov chain in four regions of Kenya. The time period is from 2004 to 2008 and then predicts the vegetative drought in January, 2009. It demonstrates that fuzzy Markov chain can be used in vegetative drought prediction.

- What is vegetative drought and how can it be characterised?
  Vegetative drought is a type of drought similar to the agricultural drought, while involving more vegetation than agricultural plants. It is the vegetation stress caused not by human behaviours but water deficiency, less soil moisture and so on. Under traditional meteorological measurement of drought, it can be characterised with precipitation data, soil parameters, and temperature. Because satellite images have strength in time and spatial continuity, the vegetative drought is characterised by NDVI index. There is a linear relationship between NDVI and precipitation reveals that the NDVI can be an indicator for vegetative drought. Comparing with the NDVI mean value, if the anomaly NDVI value is negative, it could be the appearance of vegetative drought.

- How to determine and model the vegetative drought states?
  The vegetative drought states are determined as drought classes in this study. According to the drought categories in other drought indices and purpose on this study, there are 4 drought states and 2 non-drought states. They are extremely drought, severely drought, moderately drought, dryness, wet and moderately wet. The original determination of the drought states are based on the equal frequency of sample data.

- How to model the change between vegetative drought states?
  As vegetative drought is vague object, crisp drought states do not describe it appropriately. The fuzzy membership functions are applied to account for the gradual transition between states. The higher the value is, the higher probability the vegetative states it belongs to.

- How to evaluate the prediction result?
  With the data from time period 2004-2008, the first-order Markov chain can be modelled for prediction. The validation data is from the same satellite images of 2009. The validation is from comparison the probability value of prediction with the reference data applied in membership functions. The RMSE is applied to be the quantity value of validation.

- How the number and value range of classes can affect the results?
  More classes can describe more detail of drought, but also introduce discontinuity in fuzzy Markov chain modelling. With more classes, the time-independent of first-order is damaged. Less class can fit the Markovian properties, but giving less information in prediction and early warning system. The value range, modelled in membership function, affect the prediction result. In this study, smaller core zone and bigger transition zone in membership functions improve the accuracy of prediction.

8.2. Recommendations

For further studies on predicting drought using this approach, we recommend to first quantify the relationship between drought categories and vegetative index based on the field knowledge in order to ensure the completeness of data model. Also, the acquisition of ground situation is recommended for quantitative validation of prediction result.

The fuzzy membership functions applied in this study is in trapezoidal shape and it effectively predict vegetative drought with Markov chain. For further study, we recommend to select other function types, such as Gaussian to improve the accuracy of result.

Further work is also recommended to upscale and downscale the spatial resolution of sample data. And changing of time interval is also recommended, such as monthly or weekly.

LIST OF REFERENCES

Anderson, T. W., & Goodman, L. A. (1957). Statistical Inference about Markov Chains. *The Annals of Mathematical Statistics, 28*(1), 89-110.

Bajgiran, P. R., Darvishsefat, A. A., Khalili, A., & Makhdoum, M. F. (2008). Using AVHRR-based vegetation indices for drought monitoring in the Northwest of Iran. *Journal of Arid Environments, 72*(6), 1086-1096.

Balzter, H. (2000). Markov chain models for vegetation dynamics. *Ecological Modelling, 126*, 139-154.

Banik, P., Mandal, A., & Rahman, M. S. (2002). Markov chain analysis of weekly rainfall data in determining drought - proneness. *Discrete Dynamics in Nature and Society, 7*, 231-239.

Bickenbach, F., & Bode, E. (2003). Evaluating the Markov Property in Studies of Economic Convergence. *International Regional Science Review, 26*(3), 363-392.

Boken, V. K., Cracknell, A. P., & Heathcote, R. L. (2005). *Monitoring and predicting agricultural drought: a global study*: Oxford University Press.

Cancelliere, A., Mauro, G., Bonaccorso, B., & Rossi, G. (2007). Drought forecasting using the Standardized Precipitation Index. *Water Resources Management, 21*, 801-819.

Chatfield, C. (2004). Autoregressive processes *The analysis of time series: an inroduction* (Sixth edition ed., pp. Ch 3.3). Boca Raton: Chapman & Hall/CRC.

Collins, L. (1975). *introduction to Markov Chain analysis*. Norwich: Geo Books.

Davenport, M. L., & Nicholson, S. E. (1993). On the relation between rainfall and the Normalized Difference Vegetation Index for diverse vegetation types in East Africa. *International Journal of Remote Sensing, 14*(12), 2369 - 2389.

Dilo, A. (2006). *Representation of and reasoning with vagueness in spatial information : a system for handling vague objects.* ITC, Enschede.

Drakopoulos, J. A. (1995). Probabilities, possibilities, and fuzzy sets. *Fuzzy Sets and Systems, 75*(1), 1-15.

Gouveia, C., Trigo, R. M., & DaCamara, C. C. (2009). Drought and vegetation stress monitoring in Portugal using satellite data. *Nat. Hazards Earth Syst. Sci., 9*(1), 185-195.

Grinstead, C. M., & Snell, J. L. (1997). Markov Chains *Introduction to probability*: American Mathematical Society.

Heim, R. R. (2002). A Review of Twentieth-Century Drought Indices Used in the United States. *Bulletin of the American Meteorological Society, 83*(8), 1149-1165.

Ji, L., & Peters, A. J. (2003). Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices. *Remote Sensing of Environment, 87*(1), 85-98.

Karl, T. R. (1986). The Sensitivity of the Palmer Drought Severity Index and Palmer's Z-Index to their Calibration Coefficients Including Potential Evapotranspiration. *Journal of Climate and Applied Meteorology, 25*(1), 77-86.

Kogan, F. N. (1995). Application of vegetation index and brightness temperature for drought detection. *Advances in Space Research, 15*(11), 91-100.

Kogan, F. N. (1997). Global Drought Watch from Space. *Bulletin of the American Meteorological Society, 78*(4), 621-636.

Matis, J. H., Saito, T., Grant, W. E., Iwig, W. C., & Ritchie, J. T. (1985). A Markov chain approach to crop yield forecasting. *Agricultural Systems, 18*, 171-187.

Mckee, T. B., Doesken, N. J., & Kleist, J. (1993). *The relationship of drought frequency and duration of time scals*. Paper presented at the Eighth Conference on Applied Climatology. Retrieved from http://ccc.atmos.colostate.edu/relationshipofdroughtfrequency.pdf

Meyn, S., & Tweedie, R. L. (1993). Markov Models *Markov Chains and Stochastic Stability*: Springer-Verlag.

Miao, Z., Liu, Y., & Wang, H. (2005). Fuzzy-Markov Chain Model with Weights for Prediction of Cultivated Land Demand. *Geomatics and Inforamtion Science of Wuhan University, 30*(4), 305-308.

Mishra, A., Singh, V., & Desai, V. (2009). Drought characterization: a probabilistic approach. *Stochastic Environmental Research and Risk Assessment, 23*, 41-55.

Mishra, A. K., & Singh, V. P. (2010). A review of drought concepts. *Journal of Hydrology, 391*(1-2), 202-216.

Musaningabe, R. C. (2007). *Mining drought from remote sensing images.* ITC, Enschede.

Nagarajan, R. (2009a). Drought indices *Drought Assessment*: Springer.

Nagarajan, R. (2009b). Introduction *Drought Assessment*: Springer.

NDMC. (2006). Understanding and Defining. *What is Drought* Retrieved 15, Aug., 2010, from http://www.drought.unl.edu/whatis/concept.htm

Paulo, A., & Pereira, L. (2007). Prediction of SPI Drought Class Transitions Using Markov Chains. *Water Resources Management, 21*, 1813-1827.

Pereira, L. S., Cordery, I., & Iacovides, I. (2009). Droughts and Desertification *Coping with Water Scarcity* (pp. 47-75): Springer Netherlands.

Quiring, S. M., & Ganesh, S. (2010). Evaluating the utility of the Vegetation Condition Index (VCI) for monitoring meteorological drought in Texas. *Agricultural and Forest Meteorology, 150*(3), 330-339.

Rhee, J., Im, J., & Carbone, G. J. (2010). Monitoring agricultural drought for arid and humid regions using multi-sensor remote sensing data. *Remote Sensing of Environment, 114*(12), 2875-2887.

Rojas, O., Vrieling, A., & Rembold, F. (2011). Assessing drought probability for agricultural areas in Africa with coarse resolution remote sensing imagery. *Remote Sensing of Environment, 115*(2), 343-352.

Rulinda, C. M., Bijker, W., & Stein, A. (2010). Image mining for drought monitoring in eastern Africa using Meteosat SEVIRI data. *International Journal of Applied Earth Observation and Geoinformation, 12*(S), S63-S68.

Sene, K. (2010). Droughts *Hydrometeorology* (pp. 239-267): Springer Netherlands.

Sharma, A. (2006). *Spatial data mining for drought monitoring : an approach using temporal NDVI and rainfall relationship.* ITC, Enschede.

Sinclair, B. (2005, June 9). Markov Chains. Retrieved August, 2010, from http://cnx.org/content/m10861/latest/

Smith, K. (2009). Hydrological Hazards: Drought. In K. Smith & D. N. Petley (Eds.), *Environmental Hazards: Assessing risk and reducing disaster*: Routledge.

SouthTravles. (2001). Weather and Climate - Kenya. Retrieved October, 2010, from http://www.southtravels.com/africa/kenya/weather.html

Speranza, C. I., Kiteme, B., & Wiesmann, U. (2008). Droughts and famines: The underlying factors and the causal links among agro-pastoral households in semi-arid Makueni district, Kenya. [doi: DOI: 10.1016/j.gloenvcha.2007.05.001]. *Global Environmental Change, 18*(1), 220-233.

STAR. (2010). Global Vegetation Health Products. Retrieved Nov. 30, 2010, from http://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/vh_browse.php

Tadesse, T., Brown, J. F., & Hayes, M. J. (2005). A new approach for predicting drought-related vegetation stress: Integrating satellite, climate, and biophysical data over the U.S. central plains. *ISPRS Journal of Photogrammetry and Remote Sensing, 59*(4), 244-253.

Thenkabail, P. S., Gamage, M. S. D. N., & Smakhtin, V. U. (2004). *Use of remote sensing data for drought assessment and monitoring in Southwest Asia*. Colombo, Sri Lanka.

Umamaheshwaran, R., Bijker, W., & Stein, A. (2007). Image Mining for Modeling of Forest Fires From Meteosat Images. *Geoscience and Remote Sensing, IEEE Transactions on, 45*(1), 246-253.

Vicente-Serrano, S. M. (2007). Evaluating the Impact of Drought Using Remote Sensing in a Mediterranean, Semi-arid Region. *Natural Hazards, 40*(1), 173-208.

Wilhite, D. (2005). Drought. In J. E. Oliver (Ed.), *Encyclopedia of World Climatology* (pp. 338-341): Springer Netherlands.

Wilhite, D. A. (1992). Preparing for Drought: A Guidebook for Developing Countries. Nairobi, Kenya: United Nations Environment Programme.

Wilhite, D. A., & Svoboda, M. D. (2000). *Drought Early Warning Systems in the Context of Drought Preparedness and Mitigation*. Paper presented at the Early Warning Systems for Drought Preparedness and Drought Management, Lisbon, Portugal.

WMO. (2006). *Drought monitoring and early warning: concepts, progress and future challenges*: World Meteorological Organization.

Zadeh, L. A. (1968). Probability measures of Fuzzy events. *Journal of Mathematical Analysis and Applications, 23*(2), 421-427.