Long-term variations in the net inflow record for Lake Malawi

K. Sene[1] & B. Piper[1] & D.Wykeham[1] & R.T.McSweeney[1]
[1]Atkins Limited, Water & Environment, United Kingdom

W.Tych[2] & K.Beven[2]
[2]Lancaster Environment Centre, Lancaster University, United Kingdom

ABSTRACT

Lake Malawi is the third largest lake in Africa and plays an important role in water supply, hydropower generation, agriculture and fisheries in the region. Lake level observations started in the 1890s and anecdotal evidence of variations dates back to the early 1800s. A chronology of lake level and outflow variations is presented together with updated estimates for the net inflow to the lake. The inflow series and selected rainfall records were also analysed using an Unobserved Component approach and, although there was little evidence of long-term trends, there was some indication of increasing inter-annual variability in recent decades. A weak quasi-periodic behaviour was also noted with a period of approximately 4-8 years. The results provide useful insights into the severity of drought and flood events in the region since the 1890s and the potential for seasonal forecasting of lake levels and outflows.

KEYWORDS

Malawi, Lake Malawi, southern Africa, climate, trend, variability, rainfall, lake levels

INTRODUCTION

Lake Malawi – with a mean surface area of approximately 28,760 km^2 - is the third largest lake in Africa and occupies approximately 20% of the land area of Malawi. The lake is used for water supply, fisheries and navigation and is a major tourist attraction. The river Shire, which is the sole outflow from the lake, supports extensive areas of irrigation in the Lower Shire valley together with the water supply to Malawi's second largest city, Blantyre, as well as being a major tributary of the Zambezi. The hydropower schemes in the middle reaches of the Shire supply more than 90% of the national electricity output.

There have been several studies of the water balance of the lake and much of the early work was linked to investigations of the potential for water supply, irrigation and hydropower. Drayton (1984) provides a useful summary of historical developments up to the 1980s which included landmark studies by Cochrane (1956), Pike (1964) and WMO (1976). The latter study was subsequently updated and extended in the early 1980s (WMO 1983) and subsequent research and operational studies include those by Neuland (1984), Calder *et al.* (1994), Spigel and Coulter (1996), Shela (2000), MoIWD (2001), Jury and Gwazantini (2002), Kumambala and Ervine (2010) and Lyons *et al.* (2011).

The methods used have varied widely in terms of record lengths, simulation intervals (daily, monthly, annual), and approaches to infilling and extending rainfall and tributary flow records and estimating lake evaporation. Typically, the tributary inflow terms are estimated from records for key flow gauging stations and where necessary scaled up to the full catchment area using regression relationships. Similarly, lake rainfall estimates are normally derived from an area average of raingauge values from around the lakeshore and the lake evaporation by averaging Penman estimates of open water evaporation from lakeshore meteorological stations. Some studies have included additional terms in the water balance such as a groundwater component (e.g. WMO, 1983, Lyons *et al*., 2011) or investigated individual components such as the lake rainfall and evaporation in more detail (e.g. Nicholson and Yin 2002). Several have also used rainfall-runoff models to explore the sensitivity of lake levels and outflows to factors such as land-use changes (Calder *et al*.,1994) and climate variations (e.g. WMO,1983, Kumambala, 2009).

An important finding throughout has been the extreme sensitivity of lake levels and outflows to changes in the net inflow or net basin supply to the lake, which is often called the ‘free-water’ in studies of Lake Malawi. On account of its length, this record also provides useful insights into climate variability in the region, particularly in the early 1900s before raingauge networks were first established (e.g. WMO 1983, MIWD 2001). Here, we use a stochastic signal extraction technique (Young *et al.* 1999) to explore the trends and interannual variations in this record in more detail. For comparison, the same technique is applied to indicative updates to previous estimates for the lake rainfall (WMO 1983). The findings are also compared with the results from several other studies regarding the variability in lake levels and rainfall in Malawi and other parts of southern and eastern Africa.

THE STUDY AREA

The lake catchment (Fig. 1) has a land-surface area of nearly 100,000 km^2 of which approximately 67% is in Malawi, 27% in Tanzania and 6% in Mozambique. The main inflows arise from the rivers Bua, South Rukuru, Dwangwa and Linthipe in Malawi, the Ruhuhu and Kiwira in Tanzania, and the Songwe, which forms the border between Malawi and Tanzania. These mainly originate in the highland areas surrounding the lake which reach elevations of 2500-3000m before dropping down the rift valley escarpment to the lakeshore, which is typically at an elevation of about 500m. In the Malawi section of the catchment there are also extensive areas of plateau above the escarpment, which are typically at an elevation of 1000-1500m.

In Malawi the predominant climate-type is temperate (dry winters, hot summers), with regions of arid savannah and arid steppe in the south (Peel *et al*. 2007). Variations in both rainfall and river flows are linked to the passage of the Inter-Tropical Convergence Zone (ITCZ) and intrusions of Atlantic air via the Congo basin. Additional influences sometimes include the remnants from tropical cyclones in the Indian Ocean and local convectively driven rainfall associated with the annual arrival of the ITCZ and the onset of the southeasterly trade winds as it departs towards the north. These various influences result in a main rainfall season from November to April or May over much of the lake catchment. Approximately 95% of the annual rainfall typically falls in that period although there is some evidence of a transition to a mid-latitude rainfall regime during February at many locations, resulting in a temporary reduction in rainfall intensity during that month (Nicholson *et al*. 2014). The lake is large enough to have a local influence on the diurnal atmospheric circulation and differential heating of the water and land surfaces often results in onshore winds in the afternoon and offshore winds in the morning, with lake rainfall tending to occur in the late night and/or early morning (e.g. WMO 1983, Nicholson and Yin 2002).

Due to topographic influences, the annual rainfall is often greater at the lakeshore and escarpment (1500-2000mm typically) than over the higher elevation plateau areas (700-1000mm typically) and can exceed 3000mm near the northwestern part of the lake due to wind funnelling effects in lakeside valleys (UNDP 1986). Further north and to the east, in the Tanzanian and Mozambique portions of the catchment, the annual rainfall is typically in the range 1000-2500mm. The lakeshore and/or escarpment, plateau areas, highlands and lower Shire rift valley are therefore often considered as distinct climatic zones in Malawi, although the boundaries and names used differ between individual studies (e.g. Fry *et al.* 2004, Nicholson *et al.* 2014).

The tributary inflows to Lake Malawi follow a similar seasonal pattern to the rainfall, typically reaching a peak in February or March and then reducing to low or zero values by the end of the dry season. Several studies have shown that due to these distinct wet and dry seasons – with little over-year storage - there is a strong correlation between rainfall and runoff on an annual basis (e.g. WMO, 1983, Drayton, 1984); also that, due both to the higher rainfall and topographic influences, the contribution to inflows from the smaller Tanzanian portion of the catchment exceeds that from areas in Malawi and is typically slightly more than half of the total tributary inflow to the lake (WMO, 1983, UNDP, 1986).

Lake levels have been recorded since the 1890s and some notable events since then (Table 2) include the near cessation of outflows for more than twenty years up to 1935, unusually high levels and outflows in the late 1970s and in 1980 which caused flooding of lakeshore communities and areas immediately downstream, and unusually low levels and outflows associated with a widespread regional drought in the early 1990s. Since 1965, the lake outflows have been controlled at a barrage – the Kamuzu Barrage – which is situated near Liwonde about 83km downstream from the lake outlet. Some estimates suggest that by the 1990s the cumulative influence of the temporary bunds built during construction of the barrage and then during subsequent operations led to lake levels being up to 0.4-0.8m higher than they would have been otherwise (Drayton, 1984, Shela, 2000).

METHODOLOGY

Lake water balance

Fig. 1 shows the main catchment area for Lake Malawi. For a given time interval the water balance for the lake can be expressed as:

$$\Delta S = P - E + Q_{in} + Q_{GW} - Q_{out} \qquad (1)$$

where ΔS is the change in storage, P is the lake rainfall, E the lake evaporation, Q_{in} and Q_{GW} are the catchment and groundwater inflows, and Q_{out} is the lake outflow to the Shire. Here, all flow terms are expressed in terms of a depth per unit lake area and a constant area is assumed. This assumption is an approximation but a reasonable one since based on level-area estimates presented in Lyons *et al*. (2011) at current levels the area varies by less than 1% per metre rise or fall.

Eq. (1) can be rewritten in the form:

$$N = \Delta S + Q_{out} = P - E + Q_{in} + Q_{GW} \qquad (2)$$

where N is the net inflow, net basin supply or free-water. This expresses the balance between two terms of which the first is based on levels and outflows whilst the second is based on quantities which are more difficult to estimate or observe. For example some observational challenges include the small number of long-term meteorological stations around the lake, a lack of groundwater observations, the large number of lake tributaries – some of which are ungauged - and the large spatial variations in rainfall around the lake catchment.

Table 1 illustrates the range of mean values suggested for the terms in the right hand side of Eq. (1). As might be expected, these types of study usually also show that estimates for the lake evaporation vary least both seasonally and from year to year. For example, based on the values presented in WMO (1983), the annual lake evaporation typically varies over a range within about 4-5% of the mean value, but the corresponding value for lake rainfall is about 24-25%; likewise the coefficients of variation for annual values are about 0.02 and 0.14 respectively. However, as can be seen from the table, the mean values across these studies typically span a wide range, although this in part reflects the different averaging periods and datasets used.

Derivation of the net inflows

Given these difficulties, for this study the net inflow was estimated from the lake level and outflow terms in the water balance. These calculations were performed using published data up to the 1980s (WMO 1983, UNDP 1986) and more recent records provided by the Ministry of Irrigation and Water Development (MIWD) in Malawi. Until 1915, levels were only documented twice per year and for a single gauge, with monthly values obtained by interpolation; however since then measurements have been made daily at three gauges (Chilumba, Monkey Bay, Nkhata Bay) and an average value computed, representing the mean lake level (Shela, 2000).

Outflows from the lake are usually recorded at the Liwonde gauge which is situated close to the Kamuzu Barrage. This river gauge – established in 1948 - was the first in Malawi and the observations are important both for operation of the barrage and for management of the hydropower and irrigation schemes further downstream. The flow record is generally considered to be of good quality and the few periods of missing data were infilled by linear interpolation in the present study. The gauge record is also a good surrogate for the outflow from the lake since this reach of the Shire is very flat, only dropping by 1-2m between the lake outlet and the barrage and with only a few minor tributary inflows, although possibly with some losses due to seepage and evaporation in Lake Malombe which lies between the lake outlet and Liwonde. An investigation of these influences (WMO 1983) suggested that on an annual basis they tend to cancel out and that even the largest seasonal differences have a negligible influence on flows at Liwonde.

For the period before the Liwonde gauge was established, an alternative approach needs to be used to estimate the lake outflows. Regarding the cessation of flows, some studies (e.g. WMO 1983) have suggested that outflows first stopped in 1915 but – in perhaps the most detailed review to date of historical accounts - MIWD (2001) suggest that this began in 1908. The blockage was possibly caused by sediment washed in during floods from tributaries downstream from the lake outlet and theories vary regarding its nature; for example ranging from a distinct sandbar formed at the lake outlet to more extensive sediment deposition in the river channel further downstream.

Lake levels then rose by 3-4m over the following 2-3 decades until outflows resumed in 1935, with the blockage cleared by 1938.

This time sequence of events has also been adopted here, re-computing the outflows in the periods 1899-1908 and 1935-1948 using a weir formula based on the channel bed (or sill) levels assumed by MIWD (2001). The outflows were assumed to be zero in the intervening years and the observed values were used from 1948 until 2010, which was the latest year for which records were available in this study. Comparisons suggested that, on an annual basis, the results were similar to those reported previously for 1899/00 to 1989/99 in MIWD (2001) and from 1954/55 to 1979/80 in WMO (1983). Here the notation 1979/80 etc. refers to the Malawi hydrological year which extends from November to October.

To help to assess the sensitivity of the results to these assumptions, for some of the analyses a second version of the record was used which omitted the period up to 1915 - when only two lake level readings were made per year - and from 1935 to 1947 when outflows were estimated from the weir formula. This record is called the partial net inflow series in the following text. It is also worth noting that, during the time that the blockage was present, there may have been some flood flows due to overtopping of the sandbar and/or outflows due to seepage through or beneath it; however these effects could not be quantified and are therefore an additional source of uncertainty in the analyses.

Lake rainfall estimates

Whilst the focus in this paper is on the long-term net inflow record, it was also considered useful to make some comparisons with previous estimates for the lake rainfall. However, as noted earlier there are many challenges in deriving these values; in particular due to the sparse raingauge coverage in early years and the influence of the lake on local rainfall.

Perhaps some of the most detailed studies to date are those reported by WMO (1983), which was one of the final outputs from more than a decade of hydrometeorological studies in Malawi. In that study, the following two long-term rainfall records were derived:

- Lake rainfall - monthly values for the period November 1954 to October 1980 derived on the basis of a weighted average of 17 raingauge records from around the lakeshore, including 4 stations in Tanzania and one on an island in the lake
- Climate index series - annual values for the years 1920/21 to 1979/90 based on a weighted average of 10 long-term raingauge records which was derived to provide an indication of the long-term variability in catchment and lake rainfall

Due to limitations on the raingauge data available before the 1950s, the index series was based only on records from Malawi and, of necessity, made use of records for several more distant gauges which were not used in the lake rainfall estimation procedure. Regarding the lake rainfall series, some limitations that were noted included the sparse nature of the raingauge network in the middle section of the lake due to lakeshore access difficulties, and the logistical challenges in obtaining rainfall data from islands in the lake.

As part of the present study, the feasibility of extending these records using the same methodology was investigated based on raingauge records obtained from more recent studies (e.g. IFAD, 2001) and from the Department of Climate Change and

Meteorological Services in Malawi. However, this proved not to be possible; for example for the climate index series only five of the ten gauges used in the original study appeared to have more recent data and of those records were only available for two gauges before the 1950s: Nkhota Kota and Mzimba (Table 3).

Instead alternative estimates were derived based on this smaller number of raingauges and the net inflow record itself. Table 4 summarises the approaches that were used which were as follows:

- WMO 1983 climate index (present study) – monthly rainfall values estimated from the WMO (1983) annual series using a typical seasonal profile
- Raingauge regression model – a multiple regression relationship between the WMO (1983) monthly lake rainfall and the records for the Nkhota Kota and Mzimba gauges
- Net inflow regression model – a linear regression relationship between the net inflows and the WMO (1983) monthly lake rainfall record

As part of this work double mass and time series comparisons were also made of the two raingauge records versus that for the only other gauge in the lake catchment with records dating from the 1920s, at Kasungu, and these checks showed no obvious major discrepancies.

For the regression analyses a Dynamic Linear Regression technique was used (Young et al. 1999) which is closely linked to the stochastic techniques described in the following section. For the purpose of estimating annual rainfall values some minor infilling of monthly values was also performed based on records for nearby gauges, where available, or long-term mean values.

Based on these analyses the mean values for the individual series ranged from about 1414 to 1573mm for the period in common (1954/55 to 1979/80). When compared to the monthly lake rainfall series, the Nash-Sutcliffe efficiencies were about 0.90 and 0.86 respectively for the raingauge and net inflow regression models and 0.86 for the WMO 1983 climate index series.

Investigations of trend and variability

There are many approaches to estimating the temporal characteristics of hydrological records and some commonly used techniques include linear regression (with time), tests based on sign (e.g. the Mann-Kendall test), subtracting an assumed cyclical component, and comparisons of mean values for different averaging periods. Some typical challenges include the limitations of short record lengths, dealing with missing data values, and the identification of statistically significant behaviour.

An approach which avoids many of these problems is to adopt methods based on the unobserved component signal extraction techniques developed for the analysis of non-stationary observations. For the analyses of the net inflow and lake rainfall records derived in this study, the Dynamic Harmonic Regression technique (UC-DHR) of Young *et al.* (1999) was used and can be considered as an extension of the classical Fourier series approach which in addition allows for time-varying parameters. This provides a powerful and computationally efficient technique for data exploration with few prior assumptions required about the nature or magnitude of any trends or quasi-periodic behaviour. The method has been used for trend identification, interpolation of missing data and forecasting for a wide range of environmental and economic applications, including investigations of the impacts of land use change on runoff in the UK and Malaysia (Chappell *et al.* 2012).

The methodology is described in detail in the papers cited so only key details are provided. In essence though the approach used is to assume a functional form for the time varying nature of a series involving estimating changing coefficients of a harmonic regressive model by optimal filtering/smoothing operations using a combination of Kalman Filter and a Fixed Interval Smoother (KF/FIS). A recursive formulation - essentially time stepping through the data in both directions – provides both a mathematically elegant and computationally efficient approach accommodating any missing data and outliers within the methodological framework. Measures of uncertainty of the estimation results are an inherent part of the stochastic nature of this model.

In addition to correlation coefficients, additional more complicated performance measures known as information criteria are used to help identify optimum model metrics. Other important elements of the method include the assumed variance parameters of the stochastic model (Noise Variance Ratios in the KF/FIS formulation); these parameters define the time scale of the parametric variation.

Regarding the model formulation, various forms are available and the version used for this study had the following form:

$$y_t = T_t + S_t + e_t \qquad (3)$$

where y_t is the observed time series, T_t is a stochastic trend or low frequency component, S_t is a seasonal component, and e_t is an ‘irregular’ component, arising from factors such as the observation error. This approach is sometimes referred to as spectral decomposition, as the signal is split into the following three components:

- A very slow, low frequency trend component T_t
- The specific periodicity or periodicities (seasonal, diurnal, cyclic – as required in the model, and their harmonics in S_t)
- An unmodelled component e_t covering the rest of the spectrum, interpreted as the model residual

The seasonal component is represented by a combination of sine and cosine functions:

$$S_t = \sum_{i=1}^{N} \{a_{i,t}\, \mathrm{cost}\, (\omega_i t) + b_{i,t} \sin(\omega_i t)\} \qquad (4)$$

where $a_{i,t}$ and $b_{i,t}$ are stochastic time-varying amplitude parameters and ω_i, i=1,2,..N are the fundamental and harmonic frequencies associated with the periodicity in the series, in this case on an annual or sub-annual basis. Other possibilities – not required here – include the options to specify a longer-term quasi-cyclical (extra-annual) component and/or a vector of external input (i.e. exogenous) variables.

The extent, if any, to which each term in Eq. (3) is statistically significant is then assessed using confidence intervals computed as an inherent part of the estimation procedure, thus providing the vital model/data uncertainty information and allowing for assessment of the significance of any or all of the components of the model. The input data can include missing values if required and can be analysed for any desired time interval, including daily, monthly or annual values.

RESULTS

Annual variations in net inflows

Fig. 2 shows the estimated values for the annual net inflows using the full record from 1899 to 2010, with the years with the greatest uncertainties in lake levels and outflows highlighted. Values are expressed as an equivalent depth over the lake surface assuming a mean surface area of 28760km^2; as noted previously the change in area per metre rise or fall is thought to be small (less than 1%). Over this period, the estimated mean annual inflow value was approximately 0.3m but in some years dropped below zero, most probably when the losses exceeded the combined rainfall and runoff into the lake. Also, it can be seen that two of the key events in recent times – the low flows of the early 1990s and the 1979/80 floods – were some of the most extreme in this record, with comparable dry periods only occurring in 1900/01 and 1948/49, and the high flow period unmatched.

Table 3 summarises these events and a number of others in the history of Lake Malawi, based on the observational record and earlier traveller's reports of variations in lake levels during the 19th century (UNDP, 1986, Nicholson and Yin, 2001). Interestingly, there is evidence that lake levels were also exceptionally low in the early part of the 19th century and Nicholson (1998) notes that a drought – defined as unusually low rainfall – prevailed for most of the period from the start of the century to the 1860s and was particularly intense in the 1820s and 1830s, affecting major lakes throughout Africa.

More generally, the drought of the early 1990s was widespread in southern Africa and has been linked to El Niño Southern Oscillation (ENSO) events in the period 1991-

1995 (e.g. Jury and Mwafulirwa, 2002). By contrast, the increase in levels in 1997/98 is thought to have been due to increased rainfall in the eastern catchments of Lake Malawi and in eastern Africa, which caused increases in lake levels as far north as Ethiopia and Sudan (Birkett *et al*., 1999). Again there may have been an El Niño influence since this tends to cause above normal rainfall in East Africa but droughts in southern Africa (Nicholson and Selato 2000). Indeed, studies based on reanalyses from atmospheric models have shown that this event was linked to both ENSO and the Indian Ocean dipole (e.g. Reason and Jagadheesha 2005).

The years 1961 and 1962 also saw exceptional rainfall in East Africa with significant rises in the levels of lakes such as Lake Victoria; however, although there was also an increase in the net inflow series for Lake Malawi, this was significantly less than for the 1979/80 event. During the years 2002 and 2005 there were also major droughts in Malawi (World Bank, 2009) but in terms of the net inflow do not appear particularly abnormal on an annual basis, although this may mask seasonal variations. It is also worth noting that, during the 2001/02 growing season, crop damage from short-lived heavy rainfall events may also have been a factor in the food shortages which occurred.

Trends and variability in net inflows

The long-term variations in flows are also of interest and a first step in applying Eq. (4) was to select the fundamental frequency and harmonic periods to use. Following inspection of the autoregressive spectrum, intervals of 1 year and 6, 4, 3 and 2.4 months were identified. The Nash-Sutcliffe efficiency of the resulting model was about 0.87 for the full series and 0.89 for the partial series and the corresponding values for the coefficient of determination were about 0.89 and 0.90.

From the annual time series of net inflows (Fig. 2) there is the visual impression of an increasing trend, although perhaps with a return towards average values since the unusually dry period in the 1990s. However, when using monthly values, for the full series the model (Eq. 3) suggested a sustained positive trend until the 1930s and then another increase in the period leading up to the unusually wet years of the late-1970s. This was then followed by a precipitous fall to the 1990s and then a subsequent increase in the following years. The partial record showed similar variations. However in neither case were the changes significant when compared with 95% confidence intervals. The estimates for the trend slope, shown for the full series in Fig. 3, illustrate an additional point, which is that the rate-of-change in the trend is rarely stable and sustained changes can occur over periods of years or even decades, reflecting the long periods of drought and above average rainfall which occur in this region. Again the partial series had a similar response.

The model also provides estimates for the seasonal components in net inflows and Fig. 4 shows the estimated amplitudes for the three largest terms (Annual and 6 and 4 months) based on the full net inflow series. As might be expected, given the distinct wet and dry seasons around the lake, the response is dominated by the annual component. For the parts of the record in which there is most confidence (i.e. based on the partial record) the model suggests that the calendar years with the largest annual amplitudes were 1950, 1963, 1979, 1989 and 2001 whilst the lowest values were in 1953, 1966, 1967 and 1991.

Although there is always a danger of seeing periodic behaviour when there is none, the annual amplitudes do sometimes seem to alternate between high and low periods, with increasing variability since the 1940s. For example, considering the main turning points in the record, the highest ‘peaks’ and ‘dips’ seem to be clustered around intervals of around 4-8 years, as shown later. In comparison, in a study of storage variations alone for Lake Malawi, Jury and Gwazantini (2002) found a biennial oscillation of 2-2.6 years and a weaker oscillation of around 5.6 years. These periods are typical of those often reported for the El Niño Southern Oscillation although northern Malawi is thought to lie near a transition zone between the separate regions of influence in southern and eastern Africa mentioned earlier (e.g. Jury and Mwafulirwa 2002). There are also indications that cold (La Niña) events affect rainfall in southern Africa (e.g. Nicholson and Selato 2000) together with influences from the Indian Ocean (e.g. Saji *et al.* 1999, Nicholson 2007, Manatsa *et al.* 2011, Jury 2013) although the interactions between these various mechanisms remain an active area for research.

Trends and variability in lake rainfall

Similar techniques were used to analyse the long-term lake rainfall records. Again monthly values were considered and for convenience a logarithmic transformation was used in the analyses.

Since combining the series might mask underlying signals, the records derived in the present study were initially analysed separately, with similar results for all three series. For the amplitudes, the annual component was again by far the largest and again there seemed to be little evidence of an increasing or decreasing trend in the periods of record either from the trend slope results or the trend values. As for the net inflows, the late 1970s again appear as a high rainfall period and the early 1990s as a low rainfall period.

The lack of any definite trend has also been found in other studies of rainfall in Malawi and surrounding regions using different datasets and techniques. For example, for the period 1960-2001, Ngongondo *et al.* 2011 found a roughly equal split between an increasing or decreasing trend in annual rainfall for the 42 raingauge records considered in Malawi, although this was only statistically significant for three of those stations. Similarly, based on an analysis of records for 71 raingauges in Malawi, including locations outside the lake catchment, Nicholson *et al.* (2014) found no long-term trends in the period 1900-2010, although noted that rainfall in the northern lakeshore and plateau areas was generally below normal in the 1990s and 2000s. Some differences were also noted in both the interannual variability and spatial coherence in records between the early and later parts of the rainfall season, which were attributed to long-term changes in atmospheric circulation.

In contrast, for the southern highlands of Tanzania, including parts of the Lake Malawi catchment, in an analysis for 16 raingauge records from 1970-2010 Mbululo and Nyihirani (2012) found that the wettest years were 1977/78, 1978/79, 1984/85, 1988/89 and 1997/98 whilst the driest years were 1976/77, 1987/88, 1990/00, 2002/03 and 2005/06. It therefore appears that there are some differences in high and low rainfall years when compared to those for Malawi, perhaps indicating a different rainfall response in this part of the lake catchment; however there were insufficient long-term records to investigate this aspect further.

As for the net inflow analyses, the annual amplitude values also provided some useful insights into quasi-cyclical behaviour, and a similar pattern was exhibited in all three series; in particular there appeared to be unusually low amplitudes (‘dips’) in hydrological years 1968, 1983, 1991 and 1990 in all three series and high values (‘peaks’) in 1956 and 1978.

This effect was less apparent in the individual rainfall records, although there were some periods with high or low values at two or more raingauges; for example lows were experienced in 1967 and 1968 and highs in 1979 and lows in 1999 for two of the three gauges. The irregular components of the rainfall series – as defined by Eq. (3) – also suggested a change in pattern towards more extreme values in more recent years for the Nkhota Khota and Kasungu gauges but the results were more mixed for the Mzimba gauge. So, although there might be some signs of increasing variability in recent decades, this did not appear to be a general result, based on this small sample of gauge records.

To provide a more quantitative estimate for this cyclical behaviour, typical turning points were identified manually and the time intervals between them estimated. A similar exercise was also performed for the net inflow amplitude series (in Fig. 4) and Fig. 5 shows the results of these analyses, which cover about 100 turning points in total. The distributions for the net inflows and lake rainfall were generally similar and the ranges spanned were 2-6 and 2-7 years for the lake rainfall 'peaks' and 'dips' respectively and 2-10 and 3-8 years for the corresponding values for the net inflows.

Although subjective, this again illustrates a possible linkage to phenomena occurring on timescales of a few years, such as the El Niño Southern Oscillation or Indian Ocean Dipole. Here, before performing this analysis, the individual lake rainfall series were combined into a single annual record which, although not a statistically homogenous series, still provides some information on the relative magnitudes of rainfall in different periods, and whether dry or wet years tend to occur in succession. This series was constructed as follows, again using the terminology defined in Table 4 (and shown here as *period – series used*):

- 1899/00-1919/20 – Net inflow regression model (present study)
- 1920/21-1953/54 - WMO 1983 annual index (present study)
- 1954/55-1979/80 - WMO 1983 lake rainfall
- 1980/81-2008/09 – Rainfall regression model (present study)

For exploring long-term variations it is also convenient to plot the annual values for this series (Fig 6). Here Figure 6(a) shows a comparison of this combined record with the net inflows, standardised in terms of the mean values and standard deviations, and Figure 6(b) shows the rainfall series itself, in terms of the percentage departures from the mean.

In general terms Figure 6(a) shows a close correspondence between the standardised rainfall and net inflow series, although with some notable exceptions, such as in the late 1920s and in 1983/84 and 1992/93. This helps to confirm the value of the net inflow as an indicator of regional rainfall and at a more basic level, adds confidence in the underlying records used to calculate these values. The differences that are observed could be a real-effect and/or related to errors in lake levels, outflows and/or individual raingauge records; for example, the net inflow also responds to variations in evaporation and catchment runoff which may vary in different ways to the lake rainfall in some years. In these comparisons values for the period 1899/00-1919/20 should of course be ignored since the rainfall estimates are based on the net inflows in those years (and were also ignored when considering the turning points summarised in Figs. 5(a) and 5(b)).

From Figure 6(b), it is also interesting that some of the most notable events in the observational records for levels and outflows appear to have been caused by rainfall shortfalls or excesses that were not extreme in terms of magnitude, but did occur over a period of years. From the records available it therefore appears that major changes in levels and outflows tend to occur from prolonged periods of above or below average rainfall, rather than single unusually dry or wet years. However there is always the potential for an extreme rainfall event in an individual year to lead to a rapid rise or fall in levels.

DISCUSSION AND CONCLUSIONS

These results illustrate a number of interesting features regarding the long-term variations in the net inflows to Lake Malawi and the rainfall in its catchment area. In particular, in the 20th century, the most extreme periods in the observational record to 2010 appear to have been the dry years of the early 1990s and the high inflows during the 1979/80 floods. Some other notably low inflow years were 1900/01 and 1948/49 although it is of interest that the blockage at the lake outlet in the early 1900s seems to have resulted from a sustained period of low rainfall and inflows rather than from any one particular event.

Based on the model outputs, overall there seems to have been a slight but statistically insignificant increasing trend in the net inflows since the start of observations. However, this has been swamped by periods of low and high inflows, which can last for a decade or more in some cases. Other complicating factors may also have played a role such as changes in land use and water abstractions on the tributaries flowing into the lake. These are difficult to quantify although it is worth noting the lake catchment area remains largely rural with few major irrigation or dam schemes to date, although with widespread clearance of natural vegetation for agricultural and other purposes (e.g. Chavula *et al.* 2011). There was also little discernible trend in the rainfall records although with some evidence of increasing variability in recent decades.

Regarding the seasonal component of net inflows, as expected the model suggested that this was dominated by the annual contribution. There was also some evidence that the highest amplitudes seem to recur at intervals of about 4-8 years. As noted earlier these are typical of the timescales which are often cited for the El Niño Southern Oscillation and other quasi-periodic variations in the Pacific and Indian Oceans. This raises the interesting possibility of improving seasonal forecasts for the net inflows and hence lake levels and outflows based on ocean and atmospheric conditions or indices linked to these phenomena, such as the Southern Oscillation Index (e.g. Ropelewski and Jones, 1987) and the Dipole Mode Index (e.g. Saji *et al.* 1999). For example, for the lake storage alone, Jury and Gwazantini (2002) found that a regression approach based on sea surface temperatures and pressures and upper zonal winds could provide potentially useful results, and Jury (2014) – in investigations of a naturalized outflow record - found evidence that it should be possible to anticipate lake level changes by about two months for some choices of global climate variables.

Although this would be the most direct approach, another possibility would be to forecast net inflows from estimates for the individual terms in the water balance. This would entail using downscaled medium- to long-range meteorological forecasts for the region to estimate the lake rainfall combined with rainfall-runoff models for the tributary inflows and possibly an energy budget model for the lake evaporation. However some potential challenges in model calibration include major gaps in the flow observations for some sub-catchments and the large spatial variations in rainfall and runoff around the catchment. Previous studies have also suggested some enhancement of lake rainfall due to local variations in atmospheric circulation resulting from the temperature differences between the lake surface and the surrounding land, as has been observed on some other large lakes, such as Lake Victoria in East Africa.

In contrast, due to the large storage capacity of the lake, the net inflow represents an accumulation of these factors, helping to integrate or smooth out these effects. The results presented here also suggest that it varies in a similar way to the lake rainfall, providing another option for estimating that parameter in the first half of the 20th century, when few raingauge records were available. This then allows insights into the nature of variations in regional rainfall during the period in which lake outflows ceased, and for the previous decade.

Regarding forecasting techniques, both statistical and dynamical seasonal forecasting approaches have been used operationally in southern Africa since the 1990s, particularly for commercial agriculture operations (e.g. Jury 2013). For Lake Malawi, given the many uncertainties in observations and models, a probabilistic approach would be desirable and it could also be useful to update the net inflow estimates using data assimilation techniques based on near real-time observations of lake levels and outflows. For shorter-range forecasts, there might also be advantages in using daily or 10-day (decadal) values rather than monthly values, although the flow routing effects of the lake storage would become more apparent at these timescales. The application of this approach could then provide a more risk-based basis to decision-making for a number of applications, including water supply, hydropower, and irrigation operations.

ACKNOWLEDGEMENTS

Several of the hydrological datasets used in this paper were reviewed as part of the World Bank funded Malawi Water Resources Investment Strategy for the Ministry of Irrigation and Water Development in Malawi (2010-2011). We are grateful to the

Ministry both for providing data and for expert advice on the quality and history of gauging station records. The Department of Climate Change and Meteorological Services in Malawi also provided updates to some of the raingauge records used in the analyses. The WMO (1983) study was performed by Dr Chris Kidd and provides a valued contribution to understanding of the hydrology of Lake Malawi.

REFERENCES

Birkett, C., Murtugudde, R. & Allan, T. 1999 Indian Ocean climate event brings floods to East Africa's lakes and the Sudd Marsh. *Geophys. Res. Letters* **26(8)**, 1031-1034.

Calder, I.R., Hall, R.L., Bastable, H.G., Gunston, H.M., Shela, O., Chirwa, A., Kafundu, R. 1994 The impacts of land use change on water resources in sub-Saharan Africa: a modelling study of Lake Malawi. *Journal of Hydrology* **170**, 123-135.

Chappell, N.A., Tych, W. 2012 Identifying step changes in single streamflow and evaporation records due to forest cover change. *Hydrol. Process* **26**, 100–116.

Chavula, G., Brezonik, P., Bauer, M. 2011 Land Use and Land Cover Change (LULC) in the Lake Malawi Drainage Basin, 1982-2005. *International Journal of Geosciences* **2**, 172-178.

Cochrane, N.J. 1957 Lake Nyasa and the Shire River. *Proc. Instn Civ. Engrs* **8**, 363-382.

Drayton, R.S. 1984 Variations in the level of Lake Malawi. *Hydrological Sciences Journal* **29**, 1-3.

Fry, M.J., Houghton-Carr, H.A., Folwell S.S., Butao Uka, Z. 2004 Development and implementation of GIS water resources software In Southern Africa FRIEND Phase II 2000-2003 report (Eds: Meigh J, Fry M), UNESCO, Paris.

IFAD 2001 Smallholder Flood Plains Development Programme (SFPDP). Water Resources Assessment Study (WRAS). International Fund for Agricultural Development / Ministry of Agriculture and Irrigation.

Jury, M.R., Gwazantini, M.E. 2002 Climate variability in Malawi, part 2: sensitivity and prediction of lake levels. *International Journal of Climatology*, **22(11)**, 1303-1312.

Jury, M.R., Mwafulirwa, N.D. 2002 Climate variability in Malawi, part 1: dry summers, statistical associations and predictability. *International Journal of Climatology,* **22(11),** 1289–1302.

Jury, M.R. 2013 Climate prediction experiences in southern Africa 1990–2005 and key outcomes. *Natural Hazards,* **65**, 1883-1894.

Jury, M.R. 2014 Malawi's Shire River Fluctuations and Climate. Journal of Hydrometeorology, *October 2014*, 2039-2049.

Kumambala, P. 2009 Sustainability of Water Resources Development for Malawi with Particular Emphasis on North and Central Malawi. PhD thesis, University of Glasgow, Glasgow.

Kumambala, P.N., Ervine, A. 2010 Water Balance Model of Lake Malawi and its Sensitivity to Climate Change. *The Open Hydrology Journal* **4**, 152-162.

Lyons, R.P., Kroll, N., Scholz, C.A. 2011 An energy balance hydrologic model for the Lake Malawi Rift Basin, East Africa. *Global and Planetary Change* **75**, 83-97.

Manatsa, D., Matarira, C.H., Mukwada, G. 2011 Relative impacts of ENSO and Indian Ocean dipole/zonal mode on east SADC rainfall. *International Journal of Climatology* **31(4)**, 558–577.

Mbululo, Y., Nyihirani, F. 2012 Climate Characteristics over Southern Highlands Tanzania. *Atmospheric and Climate Sciences*, **2012(2)**, 454-463.

MIWD 2001 Integrated Water Resources Management System for the Lake Malawi and Shire River: Lake Malawi Level Control. Pre-feasibility study, Norconsult.

Neuland, H. 1984 Abnormal High Water Levels of Lake Malawi? – An attempt to assess the future behaviour of the lake water levels. *Geojournal* **9.4**, 323-334.

Ngongondo, C., Xu, C-Y., Gottschalk, L. 2011 Evaluation of spatial and temporal characteristics of rainfall in Malawi: a case of data scarce region. *Theor Appl Climatol.* **106**, 79–93.

Nicholson, S.E. 1998 Fluctuations of Rift Valley Lakes Malawi and Chilwa during historical times: a synthesis of geological, archaeological and historical information. In: Environmental Change and Response in East African Lakes (J.T. Lehman, ed.), Kluwer Academic Publishers, Dordrecht, pp. 207–231.

Nicholson, S.E. 2007. The nature of rainfall variability in Africa South of the Equator. *International Journal of Climatology*, **6(5)**, 515–530.

Nicholson, S.E., Selato, J.C. 2000 The influence of La Niña on African rainfall. *Int. J. Climatol*. **20**, 1761-1776.

Nicholson, S.E., Yin, X. 2001 Rainfall Conditions In Equatorial East Africa During The Nineteenth Century As Inferred From The Record Of Lake Victoria. *Climatic Change* **48**, 387–398.

Nicholson, S.E., Yin, X. 2002 Mesoscale patterns of rainfall, cloudiness and evaporation over the Great Lake of East Africa. In ‘The East African Great Lakes: Limnology, Paleoclimatolgy and Biodiversity’ (Eds E.O.Odada and D.O.Olago), Kluwer Academic Publishers, The Netherlands, pp. 93-119

Nicholson, S.E., Klotter, D., Chavula, G. 2014 A detailed rainfall climatology for Malawi, Southern Africa. *International Journal of Climatology* **34(2)**, 315-325.

Peel, M.C., Finlayson, B.L., McMahon, T.A. 2007 Updated world map of the Kὂppen-Geiger climate classification. *Hydrol. Earth Syst. Sci*. **11**, 1633–1644.

Pike, J.G. 1964 The Hydrology of Lake Nyasa. *Journal of the Institution of Water Engineers and Scientists* **18**, 542-564.

Reason, C.J.C., Jagadheesha, D. 2005 A model investigation of recent ENSO impacts over southern Africa. *Meteorol Atmos Phys* **89**, 181-205

Ropelewski, C.F. Jones, P.D. 1987 An extension of the Tahiti-Darwin Southern Oscillation Index. *Monthly Weather Review* **115**, 2161-2165.

Saji, N., Goswami, B., Vinayachandran, P., Yamagata, T. 1999 A dipole mode in the tropical Indian Ocean. *Nature* **401**, 360-363.

Shela, O. 2000 Naturalisation of Lake Malawi levels and Shire river flows: Challenges of Water Resources Research and Sustainable Utilisation of the Lake Malawi-Shire River System. Proc. 1st WARFSA/WaterNet Symposium: Sustainable Use of Water Resources, Maputo, 1-2 November 2000.

Spigel RH, Coulter GW (1996) Comparison of Hydrology and Physical Limnology of the East African Great Lakes: Tanganyika, Malawi, Victoria, Kivu and Turkana (with reference to some North American Great Lakes). In 'The Limnology, Climatology and Paleoclimatology of the East African Lakes' (Eds.TC Johnson. EO Odada), Gordon and Breach Publishers, pp. 103-140

UNDP 1986 Annex 2B, National Water Resources Master Plan. United Nations Development Programme Projects MLW-79-015/MLW-84-003.

WMO 1976 A Water Resources Assessment of Lake Malawi. World Meteorological Organisation, Geneva.

WMO 1983. A Water Resources Evaluation of Lake Malawi and the Shire River. World Meteorological Organization, Report No. MLW/77/012, Geneva.

World Bank 2009 Malawi: Economic Vulnerability and Disaster Risk Assessment. Drought and Flood Risk Atlas.

Young, P.C., Pedregal, D.J., Tych, W. (1999) Dynamic harmonic regression. Journal of Forecasting **18**: 369–394.

Table 1 – Examples of estimates for the annual water balance of Lake Malawi

Reference	Period	P (mm)	E (mm)	Q_{in} (mm)	Q_{GW} (mm)
WMO (1976) in Drayton (1984)	1953-74	1350	1610	653	-
WMO (1983)	1954-79	1414	2264	1000	380
Neuland (1984)	1954-79	1374	1605	693	-
Spigel and Coulter (1996)	Not stated	1350	1610	650	0
Nicholson and Yin (2002)	1956-80	1350	~1700-1900	-	-
Kumambala (2009)	1975-90	1272	1695	400	-
Lyons *et al*. (2011)	1992-07	955	1665	-	Negligible

Table 2 – Some key events which have influenced the levels and outflows for Lake Malawi (WMO, 1983, Drayton, 1984, UNDP, 1986, Shela, 2000, MoIWD, 2001 and other sources)

Period	**Description**	**Period**	**Description**
1800-1809	Levels were "…so low that local inhabitants traversed dry land where a deep lake now resides" and the Ruhuhu tributary "…was completely desiccated at some time early in the century". Levels may have been about 465m at the start of the century (Nicholson and Yin 2001)	1900-1909	Lake levels dropped with the outflow stopped by a sandbar in 1908 (MoIWD 2001)
1810-1819		1910-1919	No outflow. Minimum level reached in 1915 after which values rose by nearly 1m in the remainder of the decade
1820-1829		1920-1929	No outflow. Levels rose by nearly 2m over the decade
1830-1839		1930-1939	Levels rose by about 2.5m from 1930 to a peak in 1937. Outflows resumed from 1935
1840-1849	By mid-century "Lake Malawi had risen about 6m and maintained this level throughout the next few decades" (Nicholson and Yin 2001)	1940-1949	Country-wide drought in 1948/49. The lake level was about 1.5m below the 1937 peak
1850-1859		1950-1959	Temporary bund in place at the outlet from the lake from October 1956 to July 1957
1860-1869		1960-1969	Temporary bund placed across the Shire at Liwonde in 1965 during construction of the Kamuzu Barrage, which was also commissioned in 1965. Outflows regulated from that time
1870-1879	Lake level high in 1873 (~475m; Pike, in WMO 1983), but falling in the remainder of the decade	1970-1979	Peak annual levels of about 477m reached in the years 1978, 1979 and 1980 with inundation of lakeshore areas and high flows in the Shire
1880-1889	Lake level high in 1882 (~474m; Pike, in WMO 1983) but falling in the remainder of decade	1980-1989	
1890-1899	Lake level about 470m in 1890 but rising to the mid-1890s then falling again (Pike, in WMO 1983)	1990-1999	Levels declined by about 2m from 1989 to 1997 affecting flows in the Shire and hydropower generation, in part through temporary changes to the barrage operating rules
		2000-09	Unusual rainfall patterns in the 2001/02 crop season caused both drought and flooding. There was also a country-wide drought following rainfall deficits in the 2004/05 wet season. However, lake levels varied within a range of about 1m in this decade

Table 3 – Summary of raingauge records used in the analyses

Name	Climate zone	Approximate elevation (m)	Period selected	Approximate mean annual rainfall (mm)	Description
Kasungu Boma	Plateau	1036	1925-2009	800	Moved to Kasungu airport in 1983
Mzimba	Plateau	1350	1933-2009	870	Long established gauge in a plateau region to the west of Lake Malawi
Nkhota Kota	Lakeshore	500	1922-2009	1500	Long established gauge near the lakeshore in the northwest part of the lake

Table 4 – Summary of lake rainfall and rainfall index series discussed in the text

Series	Hydrological year (Nov-Oct)	Type	Basis of approach
WMO 1983 lake rainfall	1954/55-1979/80	Monthly lake rainfall estimates	Weighted average of 17 raingauge records of which 12 were around the lakeshore in Malawi and 4 along the Tanzanian lakeshore, with the remaining gauge on Likoma Island in the Malawi part of the lake. In the weighting scheme used, the gauge records from Malawi accounted for about 80% of the total
WMO 1983 climate index	1920/21-1979/80	Annual index series	Weighted average of 10 raingauge records, all from Malawi, of which 4 were used in the above estimation procedure and the remainder were of necessity from locations more distant from the lake, but within or near the lake catchment. Approximately two-thirds of the contribution to total values was from the following 4 gauges: Nkhota Kota, Livingstonia, Karonga and Chinteche
WMO 1983 climate index (present study)	1920/21-1979/80	Monthly index series	The annual WMO (1983) values disaggregated to monthly values using a seasonal profile. The profile for the Nkhota Kota gauge was used since a comparison with the WMO 1983 lake rainfall series showed this to be the most representative record, when compared with those for the Mzimba and Kasungu gauges. To help with infilling missing periods in the lake rainfall, the profile for the period to 1953/54 was used
Raingauge regression model (present study)	1933/34-2008/09	Monthly index series	A fixed parameter multiple regression relationship developed between the scaled logarithms of the Nkhota Kota and Mzimba records and the WMO 1983 lake rainfall record
Net inflow regression model (present study)	1899/00-2008/09	Monthly index series	A fixed parameter linear regression relationship between the net inflow record and the logarithm of the WMO 1983 lake rainfall record, with any negative estimated rainfall values set to zero for the purpose of this approximate analysis; the net effect of this assumption was to change the mean lake rainfall estimate by about 2-3-%

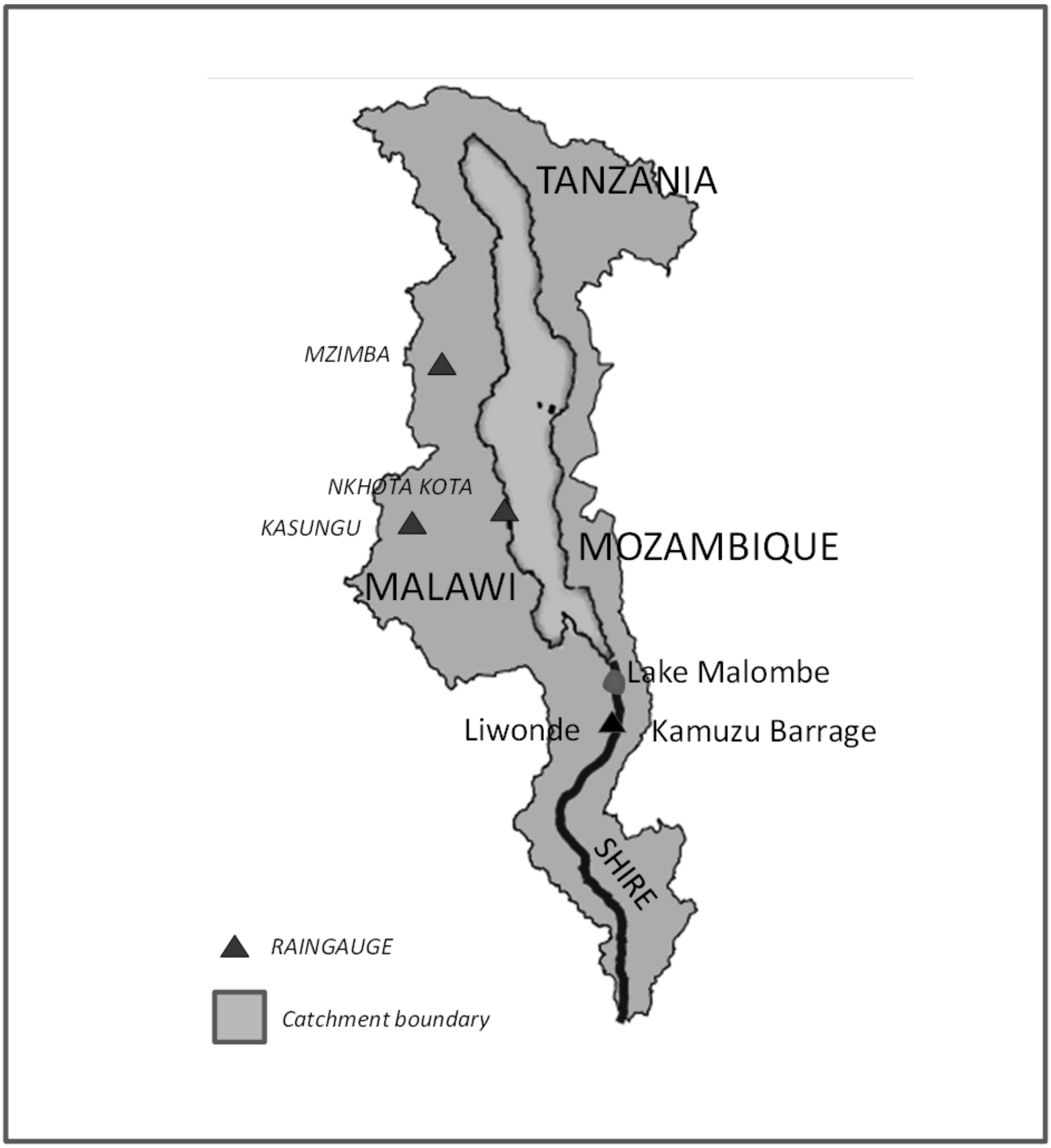


Figure 1 – Location map

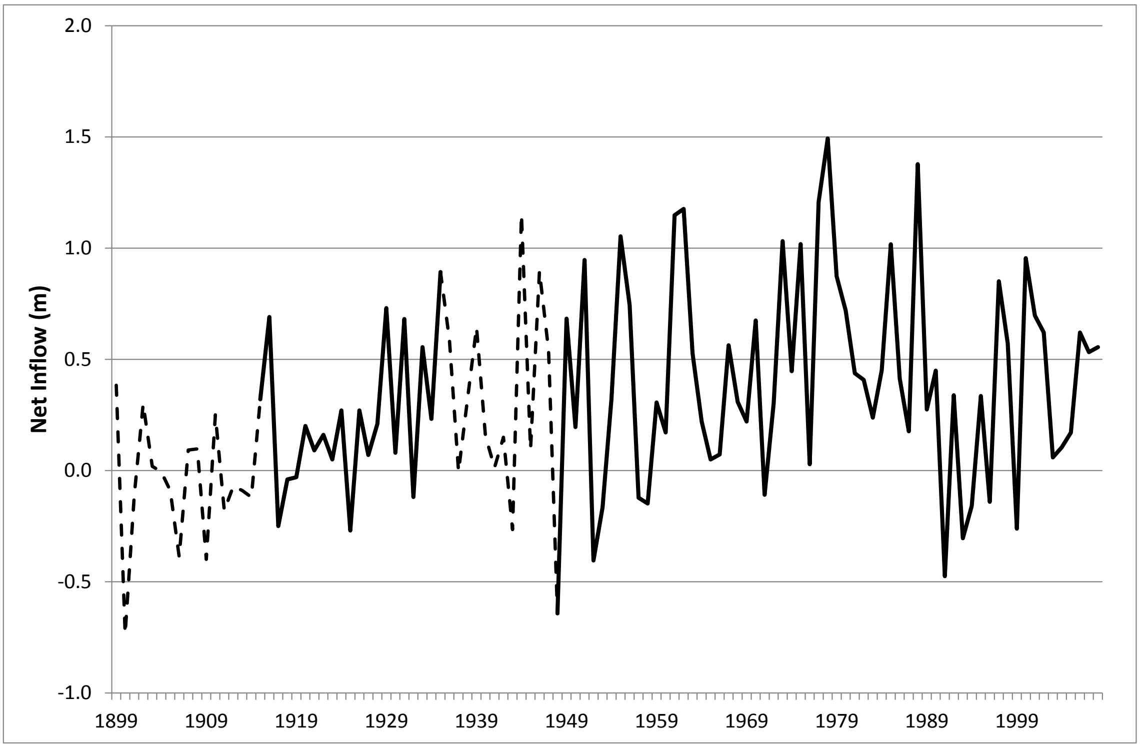


Figure 2 – Estimated annual net inflows for 1900-2008; the periods in which the lake outflows were estimated and/or levels only recorded twice per year are shown as a dotted line

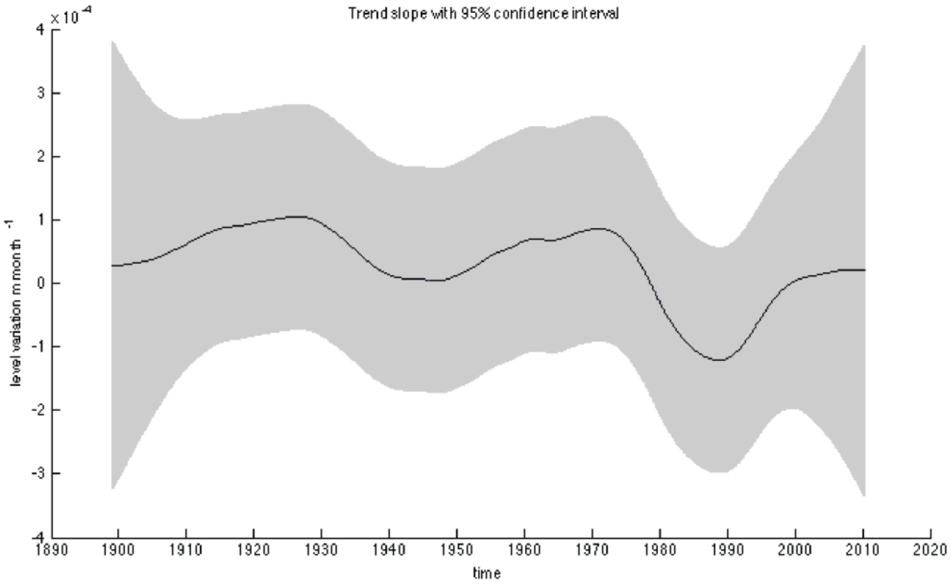


Figure 3 – Estimated trend slopes and confidence intervals for the full monthly net inflow estimates from 1899-2009

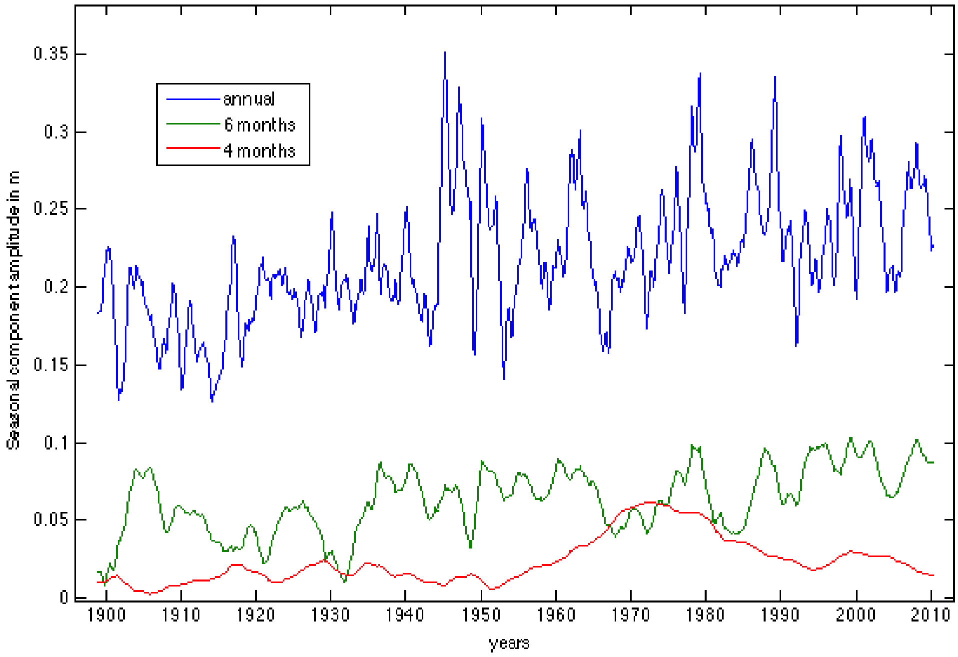
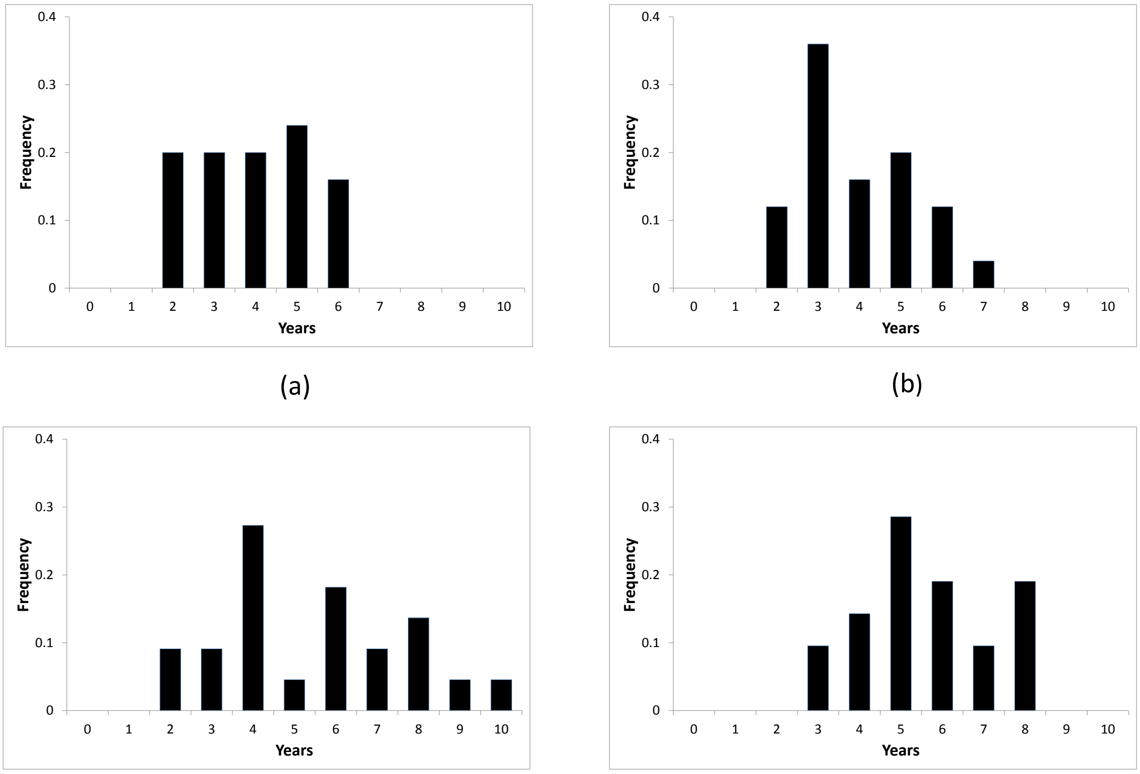


Figure 4 – Estimated amplitudes of the annual, 6-monthly and 4-monthly components for the full monthly net inflows from 1899-2009



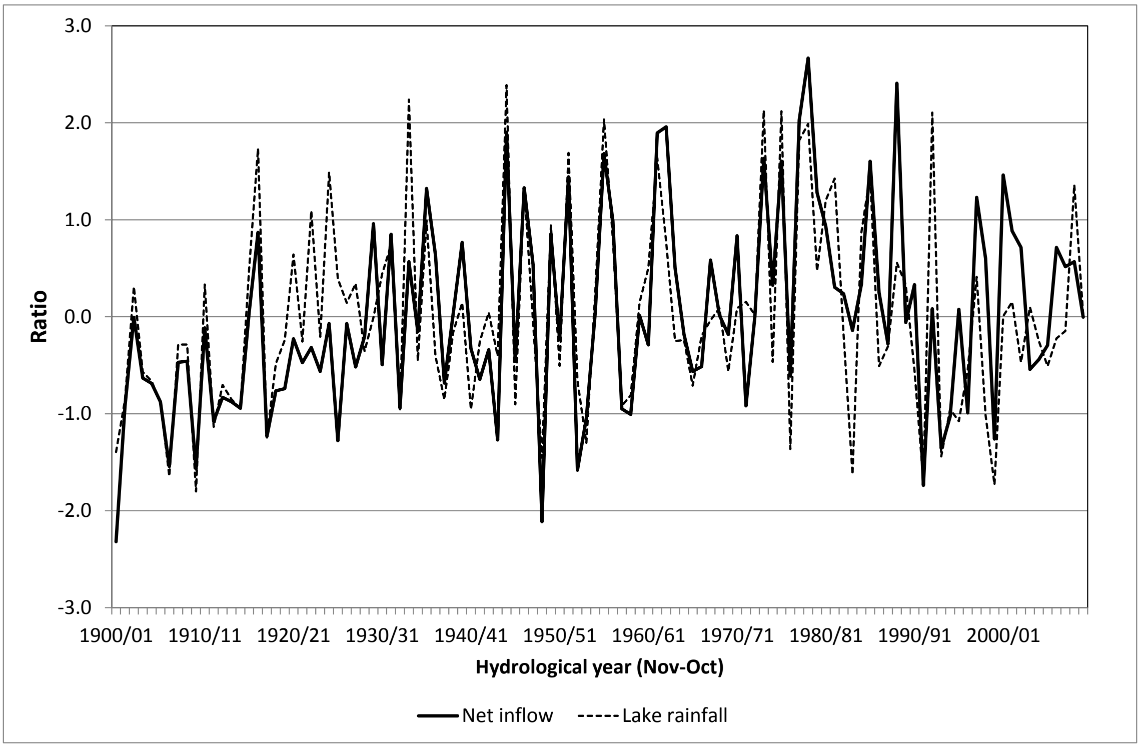
(a)

(b)

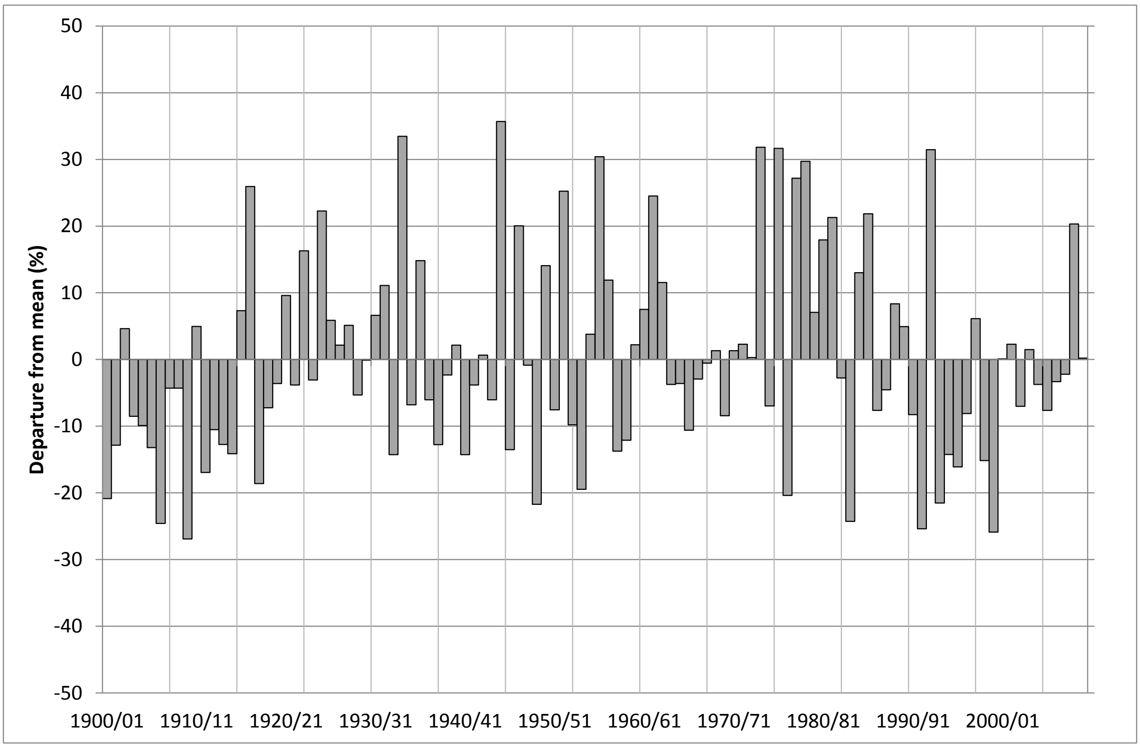
(c)

(d)

Figure 5 – Comparison of the frequencies of occurrence of peaks and dips in the annual amplitude series for the lake rainfall and net inflow series (a) lake rainfall (peaks) (b) lake rainfall (dips) (c) net inflow (peaks) (d) net inflow (dips)



(a)



(b)

Figure 6 (a) comparison of the standardised net inflow and combined lake rainfall series (b) annual percentage departures from the mean for the combined lake rainfall series