

# African Drought Risk Pay-Out Benchmarking

COLM CONNAUGHTON<sup>1†</sup>, and EMILY KAWABATA<sup>2</sup>

<sup>1</sup> *University of Warwick, UK*

<sup>2</sup> *University Edinburgh*

*(Communicated to MIIR on 13 April 2021)*

**Study Group:** ESGI130, University of Warwick

**Communicated by:** Hilary Ockendon

**Industrial Partner:** Willis Towers Watson

**Presenter:** Jon Gascoigne

**Team Members:** Colm Connaughton, University of Warwick; John Herman, University of Warwick; Adam Johansen, University of Warwick; Emily Kawabata, University of Edinburgh; Robert Kerr, University of Warwick; Michael Pegg, University of East Anglia; Jeremy Reizenstein, University of Warwick and Alan Turing Institute; Piotr Sakrajda, Warsaw University of Technology; Nick Tawn, University of Warwick; Luke Whincop, University of Warwick.

**Industrial Sector:** Finance

**Tools:** Risk analysis, statistical inference, multivariate statistics, extreme value theory

**Key Words:** Insurance, rainfall, extreme events, El Niño

**MSC2020 Codes:** 60, 62, 91

## Summary

We analyse rainfall and Water Resource Sufficiency Index (WRSI) data provided by African Risk Capacity (ARC) to assess the predictability of African droughts. We compare historical WRSI benchmarks with the observed WRSI values and conclude that they broadly align. We then compare various linear time series models based on their ability to fit and forecast the WRSI data, concluding that simpler models are more appropriate. We also investigate the use of the SARIMA and TBATS models to forecast the patterns observed in rainfall data, concluding that either can generate structured forecasts that reflect seasonal variability. Attempts to measure the influence of the El Niño-Southern Oscillation on rainfall are inconclusive for the areas studied. Finally we apply univariate Extreme Value Theory to rainfall data: further investigation is necessary to understand how the Ethiopian famine in the early 1980's would be reflected in the data were a similar event to occur today.

† Corresponding Author: [C.P.Connaughton@warwick.ac.uk](mailto:C.P.Connaughton@warwick.ac.uk)

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Background and problem statement</b>	<b>2</b>
2.1	African climatology and drought risk . . . . .	2
2.2	Mutualising drought risk: Africa Risk Capacity (ARC) . . . . .	4
2.3	Quantifying drought: Water Resource Sufficiency Index (WRSI) . . . . .	5
2.4	Simplified description of ARC insurance contracts . . . . .	6
2.5	Overview of mathematical problems in contract benchmarking . . . . .	6
<b>3</b>	<b>Assessment of benchmark WRSI values</b>	<b>8</b>
3.1	A simple model contract . . . . .	8
3.2	Actuarially fair pricing . . . . .	9
3.3	Empirical estimates of benchmarks . . . . .	9
<b>4</b>	<b>WRSI forecasting</b>	<b>12</b>
4.1	Time series forecasting . . . . .	12
4.2	Fitting some common time series models to WRSI data: in-sample errors . . . . .	12
4.3	Forecasting WRSI data with some common time series models: out-of-sample errors . . . . .	13
<b>5</b>	<b>Analysis of ENSO and rainfall time series</b>	<b>14</b>
5.1	ARIMA model of the ENSO time series . . . . .	14
5.2	SARIMA and TBATS models of rainfall time series . . . . .	16
5.3	Correlation between rainfall and ENSO time series . . . . .	19
<b>6</b>	<b>Extreme value analysis of rainfall data</b>	<b>19</b>
6.1	Issues arising in applying extreme value statistics to rainfall data . . . . .	21
6.2	Summary of theoretical foundations of Extreme Value Theory . . . . .	21
6.3	A simple extreme value analysis on Ethiopian rainfall data . . . . .	23
<b>7</b>	<b>Conclusions and possibilities for future work</b>	<b>25</b>
	<b>References</b>	<b>26</b>

# 1 Introduction

Insurance works by mutualising the risk of financial or economic losses caused by unforeseen events. Insurance is usually sought against losses that are infrequent at the population level but catastrophic at the individual level. A group of individuals wishing to insure against a particular type of loss each pay an annual premium into a common fund called a pool. In the subsequent year, the pool pays full compensation to insured individuals who suffer a loss in that year and nothing to those who don't. The size of the pool grows in proportion to the population size since every participant pays an agreed premium every year. Losses however are random. If the probability of suffering a loss is independent for each member of the population, the Central Limit Theorem ensures that the variance of the total losses grows only as the square root of the population size. Therefore, when population size is large enough and the premium is chosen appropriately, it is possible to pay full compensation to all individuals suffering losses while making the probability of exhausting the pool arbitrarily small. The assumption of independent losses is crucial here and may not be strictly true in practice. It can be approximately true if the insured population exhibits diverse exposures to the risk factors underlying the insured risk. This makes it unlikely that a large fraction of a the population will suffer a loss at the same time.

Insurance companies provide the infrastructure to manage the pool and the associated processes. Profits are made by applying a surcharge to premiums in addition to what is required to maintain the pool. Insurance companies must identify markets that are both large enough and diverse enough in terms of risk exposure to allow the Central Limit Theorem to work. They must also set premiums at levels where market participants will buy insurance. If premiums are too high, insurance will be perceived as too expensive and individuals will not buy it. If premiums are too low, the risk of exhausting the pool becomes non-negligible, putting the insurance company at risk of bankruptcy. Most insurance companies themselves buy insurance against the risk of exhausting the pool. This is called reinsurance and is provided by a number of large global reinsurance companies like Munich Re, Swiss Re, Hannover Re, Berkshire Hathaway or Lloyds of London. Mathematics becomes operationally important in setting insurance and reinsurance premiums. Theoretically, premiums can be determined from the probability distribution of losses. In practice however this distribution is an unobservable quantity. It can only be estimated from data and models. Mathematics also plays a central role in quantifying and managing the diversification of risk within the market. This report considers some of the mathematical issues arising in the quantification and management of risk associated with insurance against catastrophic drought in Africa.

## 2 Background and problem statement

### 2.1 African climatology and drought risk

Agriculture is the primary economic activity for a large proportion of the African population. The economies of many African states, and the livelihoods of many of their people, are therefore exposed to risks associated with the impact of climatic variability on agricultural yields. Extremes of climatic conditions such as droughts and floods can be strongly detrimental to agricultural productivity. The subsequent economical and social consequences can be severe. Extreme drought in particular can affect large numbers of people over extended geographical areas. Significant loss of life due to drought remains a real threat for millions of people.

The African continent includes a number of very distinct climatic zones due to its large size, varied topography and adjacency to different oceans and atmospheric circulation zones [7]. This variability is reflected in rainfall patterns across the continent. Throughout this report, we use data from a small number of vulnerability areas distributed across the different African climatic zones:

- Bekily, Madagascar
- Benishangul, Ethiopia
- Dakar, Senegal
- Jarra West, Gambia
- Sila, Chad (separate data for Maize, Millet and Sorghum)
- Zaka, Zimbabwe

We hope that these are reasonably representative although we did not do an exhaustive analysis. Fig. 1 shows some representative time series of rainfall for these regions. These plots illustrate the extent of the variation in average annual rainfall, seasonal patterns and inter-annual variation between different regions. Local agricultural practices in different regions tend to be well adapted to the average rainfall levels and the seasonal patterns. It is the inter-annual variability that can cause drought. Lower than average rainfall in a particular year, failure of seasonal rains or disruption of the seasonal pattern of rainfall can all result in drought or crop failure. Over most of Africa, this inter-annual variability is unpredictable in character. Climate research has found that it correlates with indicators of global inter-annual climate variability such as the El Nino Southern Oscillation (ENSO) or the Indian Ocean Dipole (IOD). ENSO refers to a quasi-periodic difference in sea-surface temperature between the eastern and western Pacific Ocean. The IOD is a similar phenomenon in the Indian Ocean. The time-series of the ENSO phenomenon is shown in Fig. 2(A). The corresponding autocorrelation function shown

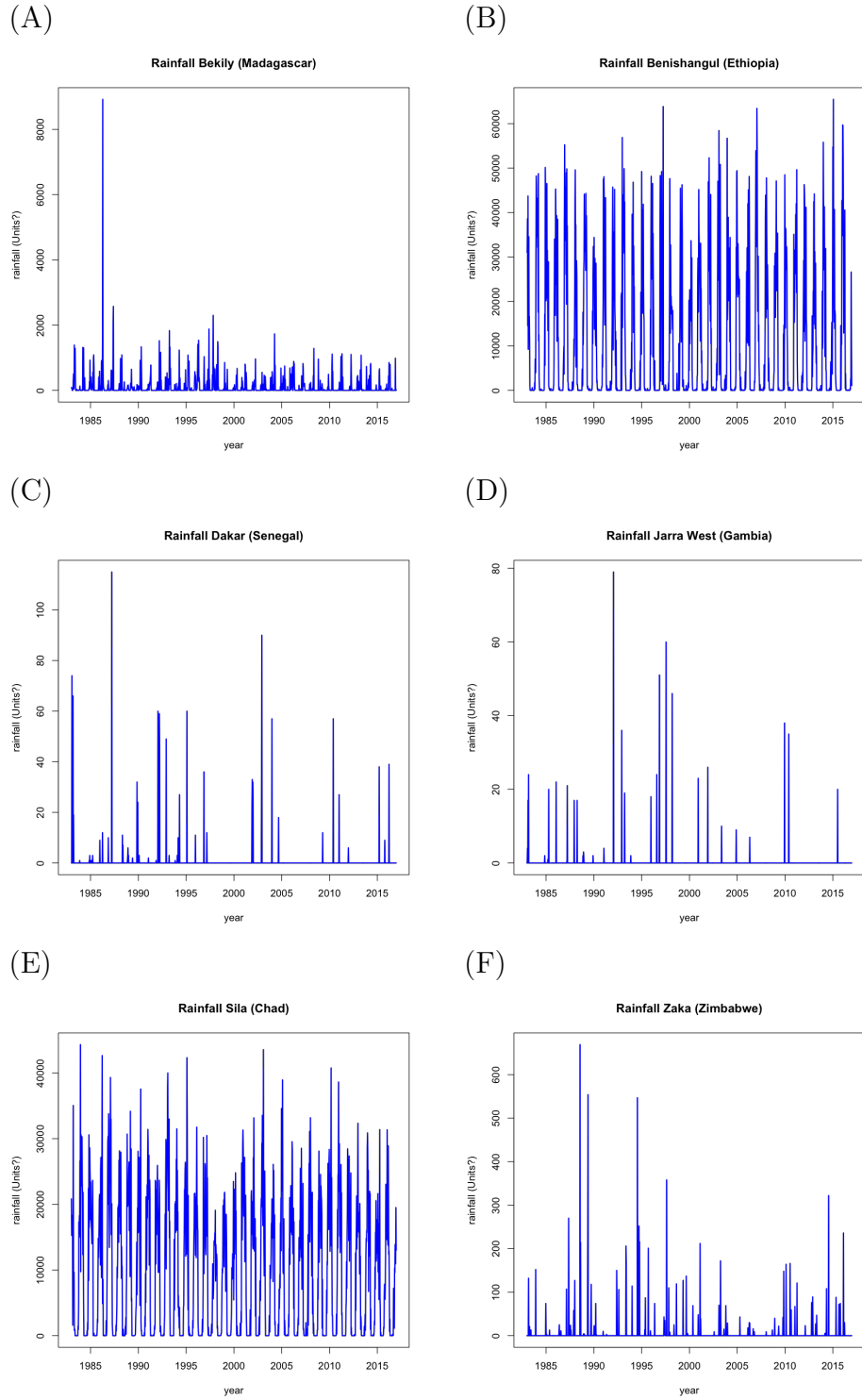


Figure 1: Time series of rainfall for a selection of administrative regions (the units got messed up during the process of data aggregation) : (A) Bekily (Madagascar), (B) Benishangul (Ethiopia), (C) Dakar (Senegal), (D) Jarra West (Gambia), (E) Sila (Chad), (F) Zaka (Zimbabwe).

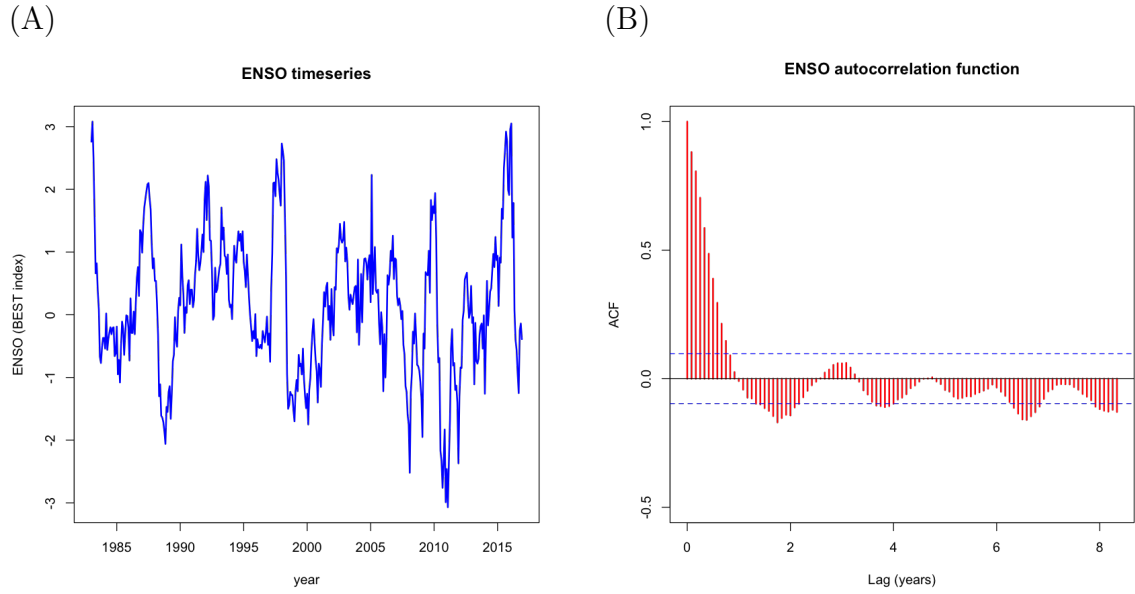


Figure 2: (A) Bivariate EnSo Time Series (BEST) signal. (B) Autocorrelation function of the BEST ENSO signal.

in Fig 2(B) illustrates the quasi-period nature of the phenomenon and indicates a characteristic timescale of a few years.

In contrast to other large scale natural catastrophes like earthquakes and tropical storms, regional droughts are slow-motion disasters. Their effects unfold on a timescale of months. An organised response by national governments can therefore be very effective at mitigating the most extreme consequences of droughts. Experience shows that the best outcomes are obtained by providing prompt financial assistance to affected populations to replace the income lost due to crop failure. Drought insurance is one mechanism for providing such prompt financial assistance.

## 2.2 Mutualising drought risk: Africa Risk Capacity (ARC)

Africa Risk Capacity (ARC) is an insurance pool providing insurance against drought to African farmers. The members of the pool are governments of nation states rather than individual farmers. This is because use of insurance to manage risk is not widespread in many African countries. This is particularly the case in rural areas where many farmers are too poor to prioritise payment of insurance premiums over more immediate requirements even when doing so would clearly be in their long term financial interest. The countries making up the membership of the pool are geographically diverse. This helps provide the required diversity of risk exposure: it is highly unlikely that all members of the pool will suffer a drought in the same year.

Willis Towers Watson (WTW) is acting as a reinsurance broker on behalf of ARC.

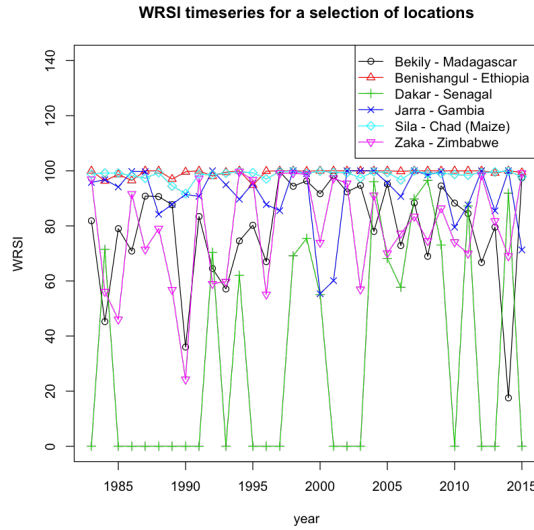


Figure 3: Time series of observed values of the WRSI for a selection of administrative regions: Bekily (Madagascar) [O], Benishangul (Ethiopia) [ $\Delta$ ], Dakar (Senegal) -[+], Jarra West (Gambia) [ $\times$ ], Sila (Chad) [ $\diamond$ ], Zaka (Zimbabwe) [ $\nabla$ ].

WTW is interested in using climate data (such as ENSO indices) and meteorological data (such as rainfall measurements) to improve estimates of drought risk. This is of interest both to WTW and to ARC in order to provide better understanding of the underlying probability distribution of losses for the purposes of setting appropriate insurance and reinsurance premiums. Insurance works on an annual cycle. The main interest is therefore in risk estimation with a 1-year forecasting horizon.

### 2.3 Quantifying drought: Water Resource Sufficiency Index (WRSI)

Drought insurance requires the ability to define drought objectively so all parties involved can reach contractual agreement. ARC have devised a methodology for doing this based on calculating a quantity called the Water Resource Sufficiency Index (WRSI) originally developed by the UN Food and Agriculture Organisation. The WRSI is calculated for each administrative region of each country participating in the insurance pool. It is a number between 0 and 100 that indicates the extent to which each region has received sufficient rainfall for its agricultural needs in a given year. A WRSI of 100 means that a region received enough rainfall to fully meet its agricultural requirements. A WRSI value less than 100 indicates a water deficit expected to lead to reduced yields. A value of zero indicates insufficient rainfall to plant any crops at all. The lower the value of the WRSI, the greater the likelihood of drought. The WRSI is calculated from a combination rainfall measurements with relevant location-specific data such as normal potential evapotranspiration, soil water holding capacity and information on crop types and cropping calendars.

The method of calculation is prescribed. Two parties starting from agreed input data should obtain the same value of WRSI for a given administrative region for a given year. Some sample time-series of the WRSI historical WRSI values obtained from the rainfall data presented in Fig. 1 are shown in Fig. 3. It is clear that the the dynamics of the WRSI varies a lot between different regions. Some regions like Benishangul (Ethiopia) and Sila (Chad) consistently attain values at or close to 100. Others like Bekily (Madagascar) and Zaka (Zimbabwe) show much more erratic fluctuations in WRSI. Dakar (Senegal) frequently experiences runs of years in which the WRSI is zero.

## 2.4 Simplified description of ARC insurance contracts

The methodology for calculating WRSI allows for some customisation to suit local conditions and allows a dynamic value to be calculated almost in real time as rainfall data is gathered throughout the year. For the purposes of this report, we make some simplifying assumptions. Firstly we assume that local parameters remain fixed for each region. Secondly we assume that only one single WRSI value is calculated at the end of each year once the rainfall data for that year has been observed. In these simplified terms, ARC insurance contracts work as follows. At the beginning of each year, members of the pool specify a WRSI benchmark in the drought insurance contract for each of its administrative regions. This benchmark is used by ARC to determine the premium that should be paid for the contract. At the end of the year, observed rainfall measurements are used to compute the WRSI values for the year by region. If the computed WRSI value for a region is below the benchmark set in the contract, the insurance contract provides a payment to the member intended to compensate farmers in that region for reduced yields caused by water deficit. For social and political reasons, ARC insurance contracts are expected to pay out about one in every four years. It is important both for ARC’s risk management and for the purposes of negotiating competitive reinsurance premiums that WRSI benchmarks are set appropriately by participating countries.

## 2.5 Overview of mathematical problems in contract benchmarking

As mentioned above, the central conceptual problem in determining insurance premiums is that the underlying probability distribution of losses is not directly observed and must be estimated from historical data. In the case at hand, losses are determined from WRSI values as described, in simplified terms, in Sec. 2.4 above. WRSI in turn is determined from a single risk factor, annual rainfall. The fundamental problem therefore comes down to estimating the probability distribution of rainfall across Africa. The standard statistical approach is to postulate a model of the data in the form of a stochastic process whose realisations should “look like” rainfall data. This model will necessarily contain parameters. These parameters

should be estimated by fitting to historical observations. The probability distribution of rainfall can be explored by generating many sample realisations from the fitted model. The practical difficulties arise from the fact that rainfall is both a non-stationary and a spatially extended process. In order to understand the issues, it is worth briefly considering why these properties introduce fundamental difficulties.

**Non-stationarity:** The statistical properties of rainfall vary with time, for example as a result of unpredictable drivers like ENSO. This is a problem for parameter estimation since the parameters must also vary in time. There is an unavoidable conflict between the desire for the fitted parameters to be representative of the true current parameter values and the desire to minimise the uncertainty in those values. The former is achieved by fitting parameters using only the most recent observed values that are presumably most representative of the current state. On the other hand, the latter is achieved by increasing the number of observed values used to fit parameters in the hope of reducing the uncertainty in those parameters. This necessarily means including older observations which are likely to be less representative of the current state.

**Spatial extent:** at each observation time, rainfall measurements are obtained at many different spatial locations. Values measured at different locations should not be modelled as independent: nearby locations, for example, are likely to be strongly correlated. Furthermore, even far away locations can have statistically significant correlations, albeit usually with a time lag. This phenomenon is referred to as a “teleconnection”. The correlation coefficients between observations at different locations should therefore be included as model parameters. However, if we observe data at  $n$  locations, there are in principle  $\frac{1}{2}n(n-1)$  correlation coefficients leading to an enormously under-determined problem.

Many questions can be asked about the probability distribution of rainfall. Different stakeholders are likely to emphasise different questions. We can organise these questions into the following hierarchy ordered in increasing level of difficulty:

- *Univariate statistics of typical events:* the simplest modelling task is to model the typical statistical properties of rainfall observations at a single location. This is likely to be of interest to individual pool members who are concerned with setting benchmarks appropriate to their own local needs. Since ARC insurance contracts are expected to pay out about once every four years, there would be limited incentive for pool members to characterise the tails of the local rainfall distribution from the perspective of setting benchmarks.
- *Multivariate statistics of typical events:* the next level of difficulty is to model correlations between rainfall time-series at different locations to understand what the typical fluctuations of the total losses drawn from the pool in a given year. This is of central importance to ARC’s risk managers.
- *Multivariate statistics of extreme events :* the hardest questions relate to characterising the tails of the distribution of spatially correlated fluctuations in rainfall and the corresponding likelihood of extreme losses big enough to ex-

haust the pool. This is likely to be of importance to WTW when brokering a reinsurance contract for ARC.

During the study group we did some exploratory data analysis to address some of these issues although within the time available, we barely scratched the surface. In particular, we did not address the spatially extended nature of the problem at all. This report summarises our findings and contains some recommendations for possible future directions for deeper work. The remainder of the report is laid out as follows. In Sec. 3 we assess whether the historical WRSI benchmarks set by the pool members are reasonable based on some simple assumptions on contract pricing and an empirical estimate of the underlying risk based on the historically observed WRSI values. In Sec. 4 we look at a selection of fairly standard linear time series models and assess the extent to which they can be used to do univariate forecasting of WRSI for individual regions one year ahead. We conclude from this that such forecasting is not very effective and propose that it could be more useful to forecast the underlying rainfall time series. Forecasting of rainfall time series is explored in Sec. 5. This section also looks at the extent to which the influence of ENSO can be detected in the rainfall data. We then demonstrate some simple univariate applications of Extreme Value Theory in Sec. 6 to estimate the return times of low rainfall events in Ethiopia. Surprisingly for the data we examined, the catastrophic famine of the early 1980's did not stand out as being particularly extreme in this analysis. Finally Sec. 7 concludes with a summary of findings and some suggestions for future work.

### 3 Assessment of benchmark WRSI values

The appropriateness of the WRSI benchmarks set by ARC members cannot be assessed independently of the premiums paid. High WRSI benchmarks should attract a higher premium than lower ones since the likelihood of a payout is higher if a higher benchmark is set. If premiums did not increase with the WRSI benchmark there would be no reason for pool members to set a WRSI benchmark lower than 100. We did not have access to premium data during the study group. However we do know that ARC contracts are expected to pay out about once in every 4 years. It is possible to work backwards from this by making some additional assumptions about fair pricing of contracts to estimate what would be a reasonable WRSI benchmark value. In this section, we carry out this analysis and conclude that the WRSI benchmarks set by ARC members are reasonable albeit slightly higher than they should be. That is to say, if the various assumptions that we make are true in practice, ARC is systematically generous to its members in setting its premiums.

#### 3.1 A simple model contract

Let us consider the following very simple model of an ARC insurance contract.

- Each country chooses a benchmark  $B$  for the following year and \$1 of assets are insured at a cost of  $r$ .
- The next year, the value of WRSI,  $w$ , is computed from measured rainfall.
- The insurance pays out \$1 if there is insufficient rain (when  $w < B$ ) and pays nothing otherwise (when  $w \geq B$ ).

To begin with, suppose we know  $\mathbb{P}(w)$  and the corresponding Cumulative Distribution Function,  $\mathbb{F}(w)$ . What value of  $B$  should be chosen?

### 3.2 Actuarially fair pricing

Let  $p(B) = \mathbb{P}(w < B)$  (probability of pay-out given  $B$ ). The expected value of the contract for ARC is:

$$V = p(B) \times (r - 1) + (1 - p(B)) \times r. \quad (1)$$

The notion of actuarially fair pricing for insurance contracts asserts that we should set  $V = 0$ . The thinking underlying this notion is similar to the no-arbitrage principle in the pricing of financial derivatives. If  $V < 0$  then ARC would make a loss on each contract and the pool would eventually be depleted. On the other hand, if  $V > 0$  then the market facilitate the emergence of a competitor offering the same insurance at a lower cost, undermining ARC's business. Hence the only sustainable possibility is  $V = 0$ . If we accept these arguments, we can solve Eq. (1) for  $r$  to give:

$$r = p(B) = \mathbb{F}(B). \quad (2)$$

This is the main conclusion of the theory of actuarially fair pricing: for binary “all or nothing” insurance contracts like the model contract described in Sec. 3.1 above, the premium should equal the pay-out probability. We know already that  $p(B) \approx 0.25$  since ARC contracts are expected to pay out once in every four years on average. Hence we should have  $r \approx 0.25$ . We have assumed that we know the CDF,  $\mathbb{F}(w)$ , so we can find the fair benchmark by solving Eq. (2) for  $B$  in terms of  $r$ :

$$B = \mathbb{F}^{-1}(r). \quad (3)$$

The problem is we *don't* know  $\mathbb{P}(w)$  or  $\mathbb{F}(w)$ . They must be estimated from historical data.

### 3.3 Empirical estimates of benchmarks

Let us assume that the WRSI,  $w$ , (after rescaling by 100 to put it in the interval  $[0, 1]$ ) is Beta distributed:

$$\mathbb{P}(w) \sim \text{Beta}(\alpha, \beta) = w^{\alpha-1} (1-w)^{\beta-1} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)},$$

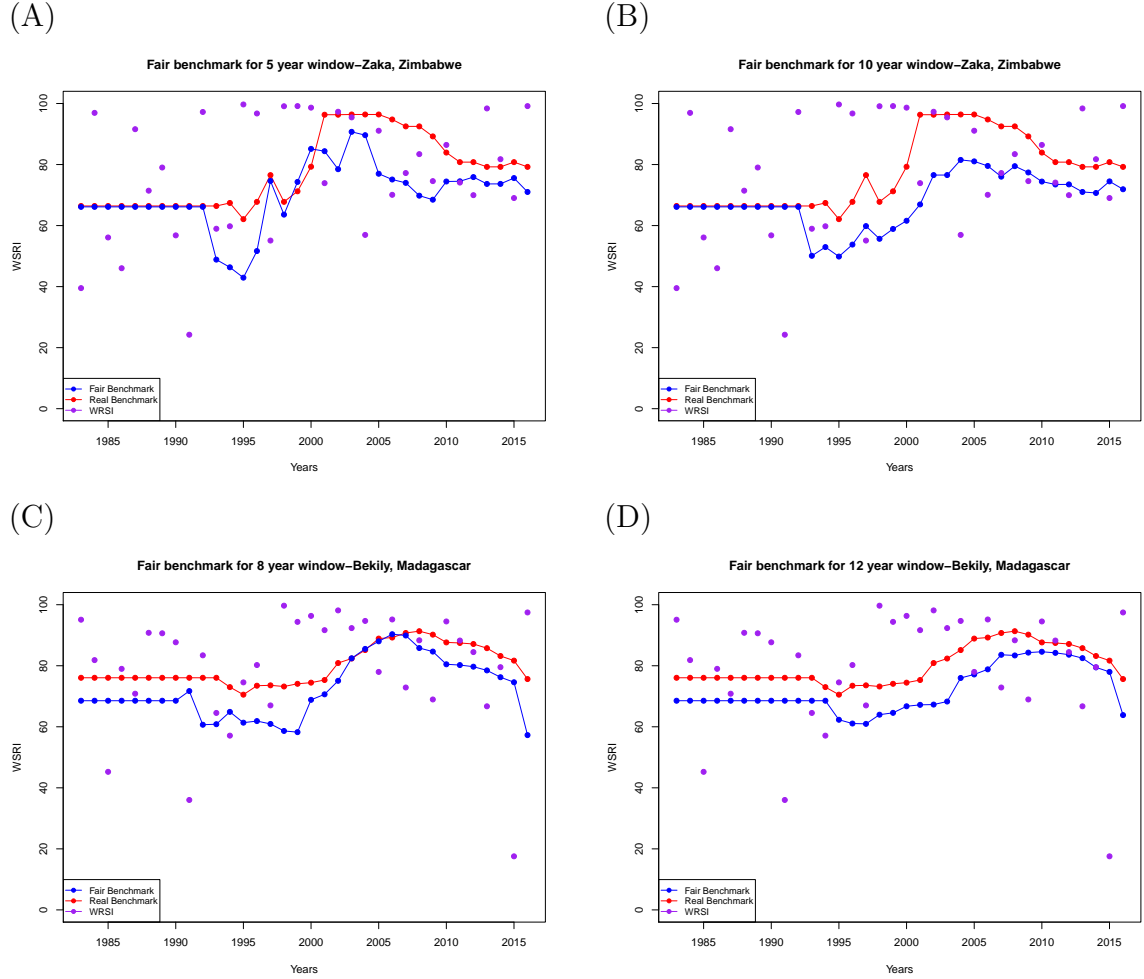


Figure 4: Comparisons between the fair WRSI benchmarks estimated using the procedure described in Sec. 3.3 (blue circles), the WRSI benchmarks set by ARC (red circles) and the actual observed WRSI values (magenta circles). (A) benchmark comparison for Zaka, Zimbabwe using 5 year estimation window, (B) benchmark comparison for Zaka, Zimbabwe using 10 year estimation window, (C) benchmark comparison for Bekily, Madagascar, using 5 year estimation window, (D) benchmark comparison for Bekily, Madagascar, using 10 year estimation window.

where  $\Gamma(x)$  is the Gamma function and  $\alpha$  and  $\beta$  are parameters. Given historical data on WRSI values, we can estimate the best values of  $\alpha$  and  $\beta$  using the method of maximum likelihood. The choice of a Beta distribution is solely because it provides a flexible distribution to describe a random variable taking values on the unit interval. It is a very poor model in some cases - for example for locations like Dakar where WRSI frequently takes a zero value. With more time, better models could certainly be found.

The Beta model for WRSI now allows us to empirically estimate the fair benchmark as a function of time from historical data. The procedure is as follows:

- for any given year, take the previous  $n$  observed values of WRSI and use them to estimate the values of the parameters  $\alpha$  and  $\beta$  for that year.
- Insert these values into Eq.(3) with  $r = 0.25$  to obtain the benchmark value. We need the CDF of the Beta distribution:

$$\mathbb{F}(w|\alpha, \beta) = \frac{\int_0^w t^{\alpha-1}(1-t)^{\beta-1} dt}{\int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt}.$$

This integral is a tabulated special function known as the regularised Beta function but it has no explicit inverse. Eq. (3) is therefore solved numerically using Mathematica.

As already discussed in Sec. 2.5 above, if the underlying distribution of WRSI is non-stationary, there is a trade off to be made in selecting the number,  $n$ , of past values of observed WRSI values. We would like to choose  $n$  to be large in order to reduce uncertainty in the estimated values of the parameters  $\alpha$  and  $\beta$  but if  $n$  is too large historical values of WRSI are included in the parameter estimation that are not representative of the current distribution of WRSI values. We don't know enough to say what should be the optimal value of  $n$  should be so we just checked some different values of  $n$  just to explore the parameter sensitivity. Fig. 4 shows the results of this analysis for two locations, Bekily, Madagascar (Fig. 4(A) and (B)) and Zaka, Zimbabwe (Fig. 4(C) and (D)). The blue circles show the fair benchmarks computed using the procedure described above. Results are shown for two different choices of the number of historical values to use in the parameter estimation:  $n = 8$  ((A) and (C)) and  $n = 12$  ((B) and ((D)). The results are qualitatively similar. The red circles show the actual WRSI benchmarks that were set by the ARC pool members for these two locations. Finally, the purple circles show the actual values of WRSI observed. We see that the benchmark values that were set by the ARC pool members are in reasonable agreement with the fair benchmarks calculated using the above procedure although they are systematically higher. This means either that ARC is systematically generous in the premiums it sets or that some of the assumptions underpinning our calculation of the fair benchmark do not hold in practice.

Location	SMA	Med	SES	HW-T	HW-T-S	TBATS
Bekily	19.64	17.79	17.86	13.94	18.58	17.33
Benishangul	1.85	1.61	1.65	1.66	2.30	1.62
Dakar	34.66	36.89	32.42	24.00	36.45	31.75
Jarra	6.03	5.66	5.30	3.85	7.90	5.15
Sila	2.33	1.83	2.11	1.28	3.09	2.04
Zaka	23.77	19.93	23.20	22.26	29.24	23.03

Table 1: In-sample RMS error for a variety of common time series models fitted to the historical WRSI time series in Fig. 3 for the period 1983-1999. The best fitting model for each location is coloured green. The worst fitting model for each location is coloured red.

## 4 WRSI forecasting

### 4.1 Time series forecasting

An alternative approach to setting benchmark WRSI values is to use time series forecasting techniques to try to model temporal correlations in the time series of observed WRSI values. After such models are fitted to historical time series data, they can be used to make forecasts of future values. Most time series models can also estimate the uncertainty in those forecasts. Thus in principle, a well-fitted time series model could provide a substitute for the assumed Beta distribution of WRSI values in the previous section. We did not get that far but restricted ourselves to investigating a suite of standard time series models against WRSI data for different regions to assess which models provide the best fit to the data and the lowest 1-year ahead forecasting error.

### 4.2 Fitting some common time series models to WRSI data: in-sample errors

We use time series data consisting of annual WRSI values from 1983 to 2015. We tested the following linear time series models:

- Simple moving average (SMA)
- Moving median (Med)
- Simple exponential smoothing (SES)
- Holt Winters model with linear trend (HW-T)
- Holt Winters model with linear trend and simple seasonality (HW-T-S)
- TBATS model with multiple seasonality (TBATS) [2]

Location	SMA	Med	SES	HW-T	HW-T-S	TBATS
Bekily	19.15	20.56	21.27	28.20	22.81	20.99
Benishangul	0.52	0.26	0.75	0.38	0.63	0.91
Dakar	46.69	60.38	47.14	61.08	52.47	49.70
Jarra	16.73	15.31	17.52	19.42	20.13	16.34
Sila	1.01	1.01	1.10	1.26	1.37	1.2215398
Zaka	13.62	16.19	13.04	18.42	18.93	14.33

Table 2: Out-of-sample RMS 1-year ahead forecasting error for a variety of common time series models fitted to the historical WRSI time series in Fig. 3 for the period 2000-2015. The best performing model for each location is coloured green. The worst performing model for each location is coloured red.

There are clearly some problems with proposing these models to represent WRSI. The standard versions of all of them assume normally distributed noise. Consequently they do not respect the constraint that WRSI takes values in the interval  $[0, 100]$ . They could even produce negative values. We proceed nevertheless.

For each time series model we use R to fit the WRSI values in the window 1983-1999 for each of the test locations and then calculate the root mean square difference between the data and the fitted model. This in-sample statistic provides a measure of how well the models fit the data. The results are shown in Table 1. It is difficult to compare results from different regions since the intrinsic variability of the WRSI varies a lot between regions as is clear from Fig. 3. The best fit model for each region is coded green in Table 1 and the worst fit is coded red. The winner comes out being HW-T although in many cases, we doubt that the differences between the best and worst fitting models are statistically significant given the short time series used in the fits. A specific disadvantage of the HW-T-S model is that the seasonal period must be specified in advance and it is not clear whether there really should be any inter-annual periodicity in the WRSI time series. In the results presented here, we set the period equal to 4 years based on the characteristic timescale seen in the ENSO ACF in Fig. 2(B). Based on the lack of a clear period, it is not surprising that HW-T-S fits badly.

### 4.3 Forecasting WRSI data with some common time series models: out-of-sample errors

Next we assessed the forecasting error of these models. To do this we considered the period 2000-2015. For each year,  $n$ , in this window, we refitted each model using the data from the years 1983 to  $n - 1$  (i.e. all the data available up until that year) for each of the regions. We then used the fitted model to forecast the WRSI value expected for year  $n$ . We then calculate the root mean square difference between the observed WRSI values and the forecast values. This is an out-of-sample statistic and should be more representative of how a forecasting model would be used in

practice. The results are shown in Table 2 with the best and worst performing models again coded in green and red respectively. We note that the models which were the best fits in the previous section mostly do not come out providing the best forecasts. In particular, the simplest forecasting models like SMA and Med seem to perform best. It could be interesting to investigate why this is so since it is not what one would normally expect. It is amusing that for Dakar, the forecast error for the HW-T is greater than 50, meaning that the model is even worse than random guessing.

## 5 Analysis of ENSO and rainfall time series

In this section we apply time series modelling to the underlying ENSO and rainfall time series rather than to the WRSI time series which is a derived quantity. There is more structure in the rainfall time series in particular and quite a lot more data since rainfall is reported every dekad (10 days) rather than once per year as with WRSI. We suggest that it may be more informative to convert rainfall forecasts into WRSI forecasts by passing them through the WRSI calculation in order to obtain a distribution of likely future WRSI values.

### 5.1 ARIMA model of the ENSO time series

We adopt a more systematic approach to time series modelling in this section. Rather than simply test a suite of candidate models as we did for the WRSI time series in the previous section, we propose models based on examining the autocorrelation function of the time series and accept a model when a Ljung-Box test of the residuals of the corresponding fit indicates a sufficient degree of independence.

We begin by constructing a model of the ENSO time series shown in Fig. 2 (A). The shape ACF of the ENSO time series shown in Fig. 2(B) suggests trying to fit an ARIMA model (AutoRegressive Integrated Moving Average) to the data. This is easily done in R using the `auto.arima` function from the `forecast` package. The best fitting model turns out to be an ARIMA(3,0,3):

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \theta_3 \epsilon_{t-3}, \quad (4)$$

where the  $\epsilon_t$  are iid Gaussian random variables with mean zero.

The R commands (where `ENSOts` is a time series object containing the ENSO data)

```
require(forecast)
fit.ENS0 = auto.arima(ENSOts)
summary(fit.ENS0)

give the output
```

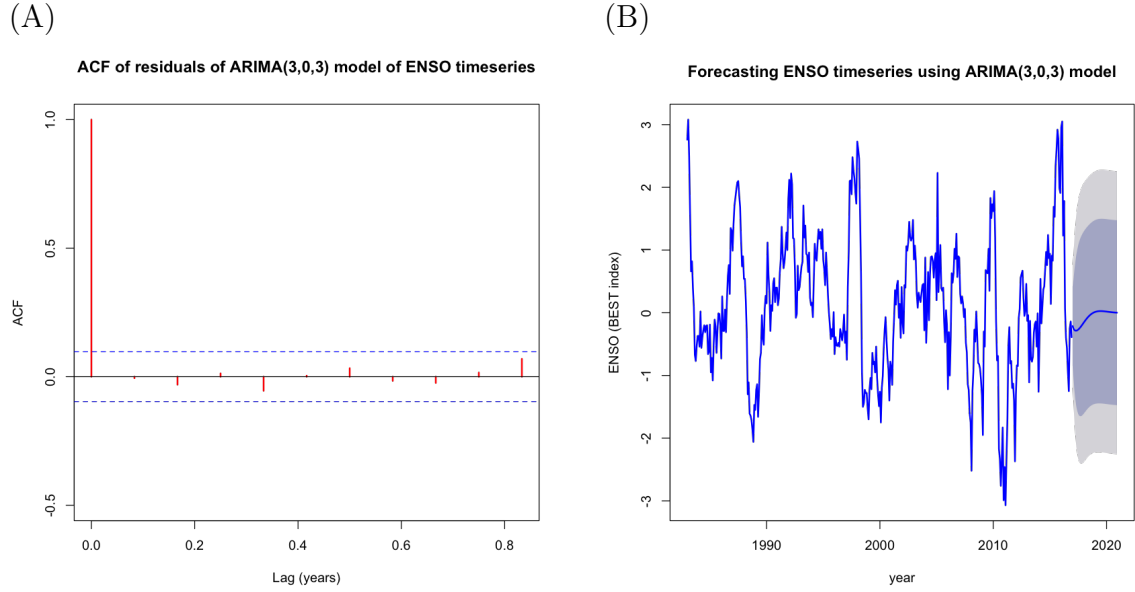


Figure 5: (A) Autocorrelation function of the residuals of the best fit ARIMA (3,0,3) model for the ENSO time series. (B) 2 year forecast of the ENSO time series using the fitted ARIMA(3,0,3) model. The dark shaded region corresponds to the 90% confidence interval and the lighter shaded region corresponds to the 95% confidence interval.

```
Loading required package: forecast
Series: ENSOts
ARIMA(3,0,3)          with zero mean
```

Coefficients:

	ar1	ar2	ar3	ma1	ma2	ma3
	2.1295	-1.4077	0.2641	-1.3925	0.7166	-0.2368
s.e.	0.3474	0.6021	0.2718	0.3417	0.3293	0.1056

sigma<sup>2</sup> estimated as 0.2426: log likelihood=-287.91

AIC=589.83 AICc=590.11 BIC=617.9

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	0.01948458	0.4888942	0.3752545	NaN	Inf	0.312239	-0.005795502

Referring to Eq. (4), the parameters ar1, ar2, ar3 and ma1, ma2, ma3 correspond to  $\phi_1$ ,  $\phi_2$ ,  $\phi_3$  and  $\theta_1$ ,  $\theta_2$ ,  $\theta_3$  respectively. The ACF of the residuals of the fitted ARIMA(3,0,3) model are shown in Fig. 5(A). We accept the model if these residuals are statistically indistinguishable from the ACF of white noise since that would indicate that there are no further temporal correlations present in the data beyond those that are described by the fitted ARIMA(3,0,3) model. Visually it seems plausible that this is the case. The Ljung-Box test is a formal statistical test to

assess the independence of residuals. The null hypothesis for the Ljung-Box test is that the data exhibit no auto-correlation up to the specified lag.

Before applying the test, let us briefly recall the notion of a p-value for the benefit of non-statisticians:

- A small p-value (typically  $\leq 0.05$ ) indicates strong evidence against the null hypothesis, so you reject the null hypothesis.
- A large p-value ( $> 0.05$ ) indicates weak evidence against the null hypothesis, so you fail to reject the null hypothesis.

Thus a small p-value is evidence against independent residuals. There is also the question of how many residuals,  $h$ , to choose. Hyndman suggests the following "rule of thumb" [3]:

- For non-seasonal time series, use

$$h = \min(10, T/5)$$

- and for seasonal time series, use

$$h = \min(2m, T/5),$$

where  $T$  is the length of the time series and  $m$  is the period (if seasonal). We adopt this rule of thumb here when applying the Ljung-Box test. The p-value comes out to be 0.9111 so we do not reject the hypothesis that the residuals are independent.

Fig. 5(B) shows forecasts made using the fitted ARIMA model out to a 4-year horizon. As expected for an ARIMA model, the forecast reverts to the mean and the uncertainty quickly grows to become comparable to the intrinsic variability of the data. Nevertheless, there is non-trivial information in the forecast, particularly in the first year or two.

## 5.2 SARIMA and TBATS models of rainfall time series

We now investigate models of the rainfall time series. We examine the rainfall from the Ethiopian region of Benishangul just to see whether we can handle the strong seasonality evident in both the data and the ACF shown in Fig. 6 (A) and (B). We try two models capable of capturing seasonality: the Seasonal ARIMA (SARIMA) model and the TBATS model. The SARIMA model has already been shown to be appropriate for modelling rainfall in Iran [6]. The best fitting SARIMA model comes out to be a SARIMA (3,0,1)(1,0,0)[12]. The fitted model and associated forecasts out to a 4-year horizon are shown in Fig. 7(B). The ACF of the residuals of the fitted model are shown in Fig. 7(A). There seems to be some significant correlations remaining in the residuals and the subsequent Ljung-Box test gives a p-value of  $5.395 \times 10^{-05}$  indicating strong evidence to reject the hypothesis that the residuals are independent.

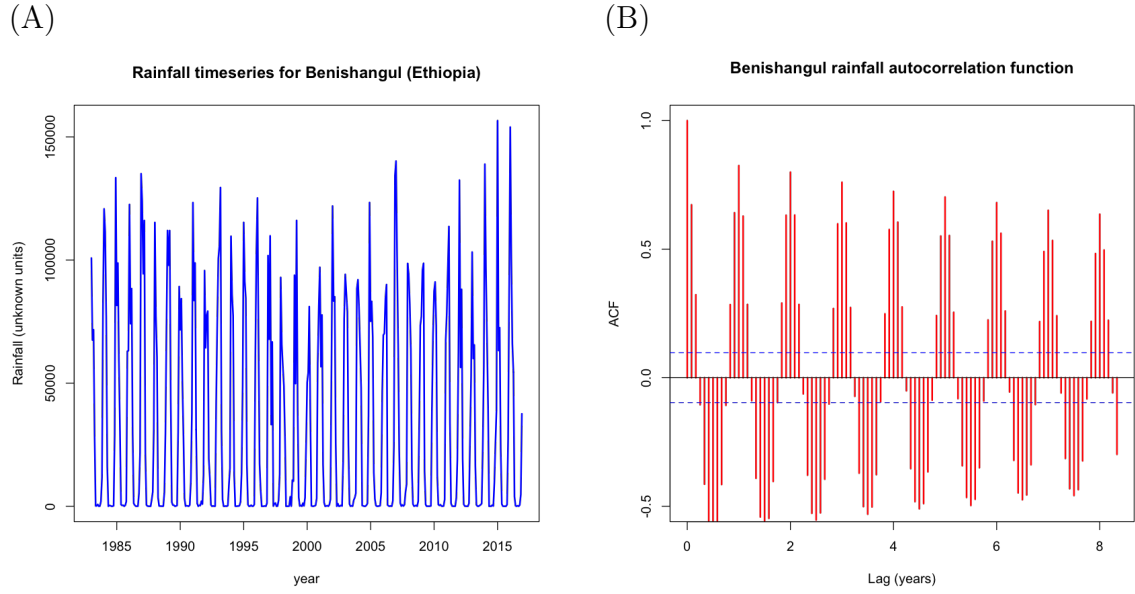


Figure 6: (A) Time series of the rainfall in the Benishangul region of Ethiopia. (B) Autocorrelation function of rainfall in the Benishangul region of Ethiopia.

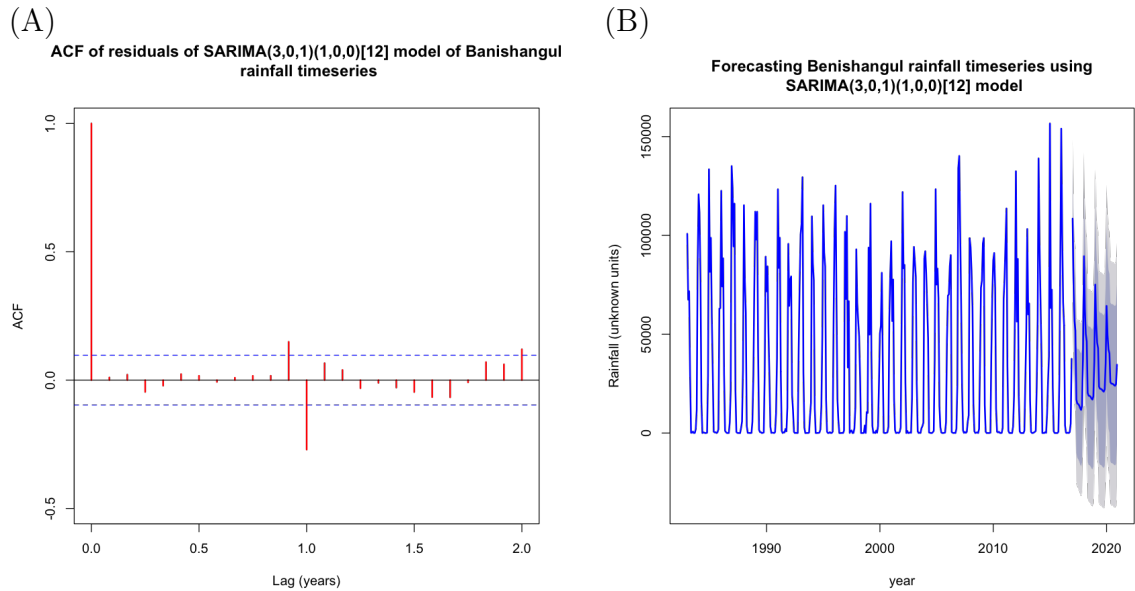


Figure 7: (A) Autocorrelation function of the residuals of the best fit SARIMA (3,0,1)(1,0,0)[12] model for the rainfall time series for the Benishangul region. (B) 2 year forecast of Benishangul rainfall using the fitted SARIMA model. The dark shaded region corresponds to the 90% confidence interval and the lighter shaded region corresponds to the 95% confidence interval.

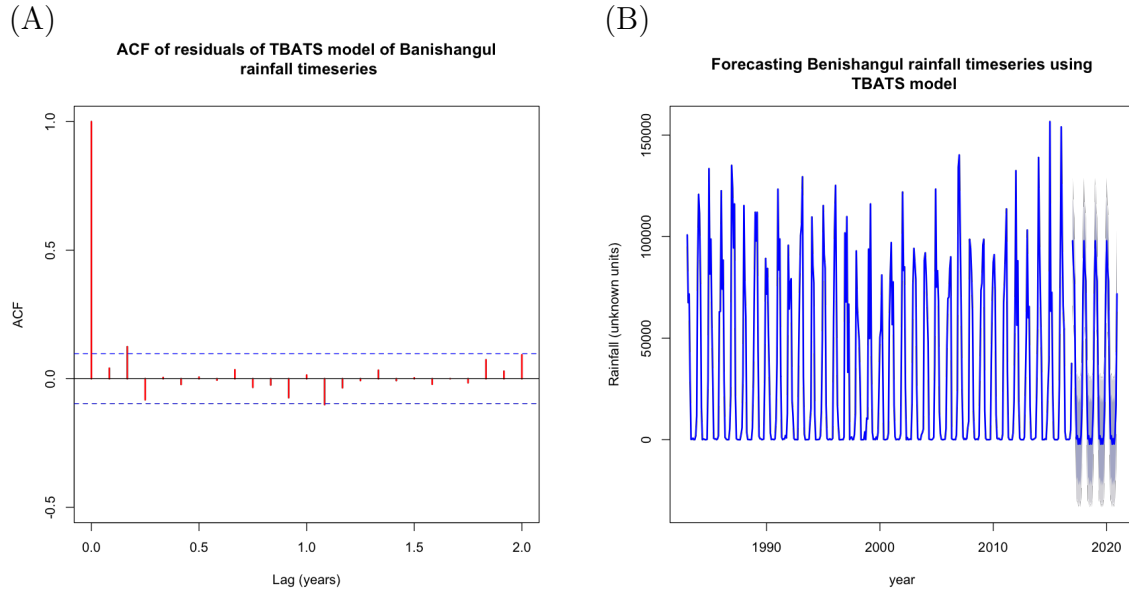


Figure 8: (A) Autocorrelation function of the residuals of the best fit TBATS model for the rainfall time series for the Benishangul region. (B) 2 year forecast of Benishangul rainfall using the fitted TBATS model. The dark shaded region corresponds to the 90% confidence interval and the lighter shaded region corresponds to the 95% confidence interval.

The best fitting TBATS model and the associated forecasts are shown in Fig. 8(B) and the ACF of the residuals in Fig. 8(B). The residuals look much better than for the SARIMA model and the subsequent Ljung-Box test gives a p-value of 0.359 indicating weak evidence for rejecting the hypothesis that the residuals are independent. Hence we take the TBATS model as our preferred model rather than the SARIMA model.

We note that both models significantly capture the seasonal structure in the time series and are capable of propagating this structure into the future when run in forecast mode. We note that although the forecasts are qualitatively similar for short horizons, their long term behaviours differ significantly. In particular, the SARIMA model (by construction) reverts to the mean level for long forecast horizons with the seasonal signal damping out whereas the TBATS model treats the seasonality as a deterministic component of the time series and propagates it unchanged into the future. The latter seems intuitively a more appropriate way of modelling the data. We note that both models suffer from the drawback that they do not understand the rainfall needs to be positive. This problem could probably be addressed by creating a multiplicative version of the model but we did not have time to investigate this.

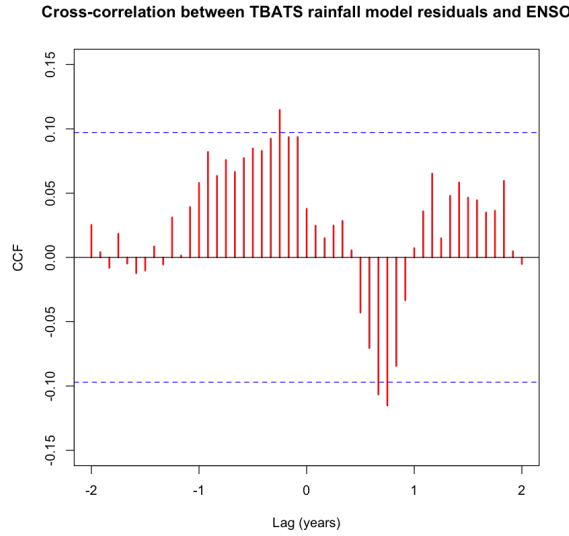


Figure 9: Cross correlation of the residuals of the best fit TBATS model for the rainfall time series for the Benishangul region with the ENSO time series.

### 5.3 Correlation between rainfall and ENSO time series

To finish this section on time series analysis we consider briefly how one might begin to quantify the effect of ENSO on rainfall patterns in particular regions. Attempts to find a direct link, for example by doing a regression of rainfall on ENSO values, exhibited too much scatter to be informative. Instead we suggest looking at the Cross Correlation Function (CCF) of the ENSO signal with the residuals of the TBATS model constructed in the previous section. If significant cross correlations are found, it means that the ENSO time series contains information that can explain the rainfall time series that is not already captured by the TBATS model. The results are shown in Fig. 9. While we see some statistically significant correlations and some regularity in the lag structure of these correlations, it is not immediately evident that these correlations are beyond what one would expect at random. A multi-variate version of the Ljung-Box test exists to help answer this question but we did not get to implement it. This inconclusive result is perhaps to be expected since the TBATS model did a pretty good job of fitting the data. It would be interesting to repeat this analysis for some other regions where the influence of ENSO is expected to be stronger.

## 6 Extreme value analysis of rainfall data

We are interested in estimating the likelihood of a particular location experiencing an extreme event, i.e., a drought, in a given year. The standard statistical methods (such as fitting a Gaussian distribution) are inappropriate for modelling extreme events due to underestimation. A well-developed theory called extreme value theory

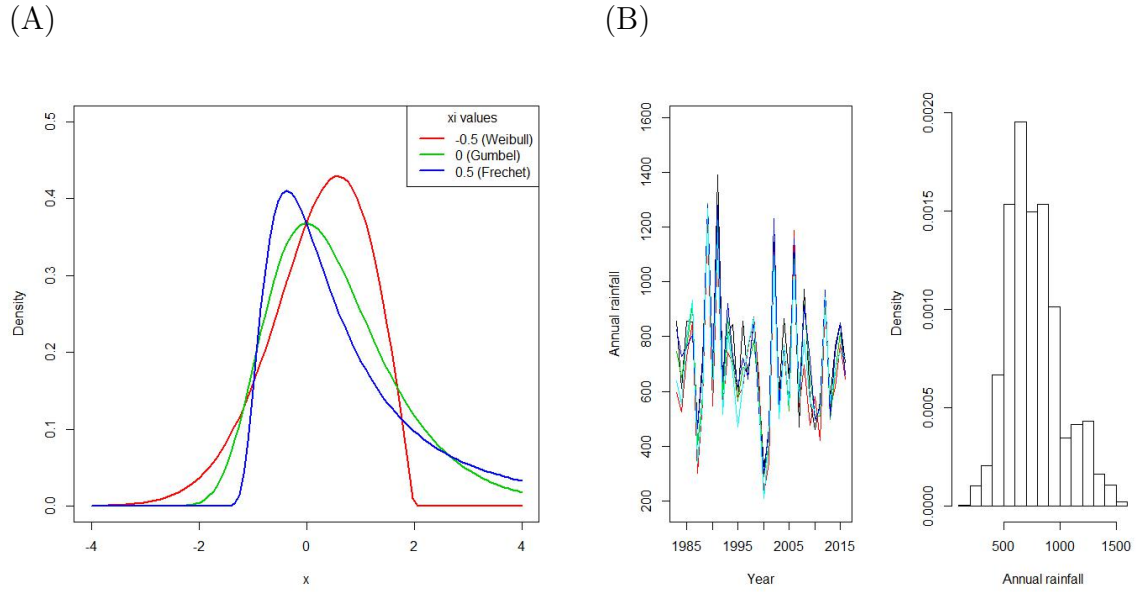


Figure 10: (A) Plots of the 3 types of extreme value distributions, Eq. (6), Eq. (7) and Eq. (8). (B) Histogram of rainfall time series for the Benishangul region of Ethiopia.

is appropriate here. Extreme value theory is based on the idea that the most extreme event has not yet been observed in a (finite) sample of observations. We use this theory to assess the probability of a more extreme event (than any previously observed) based on the observations. The distribution developed within EVT is the generalised extreme value distribution and this approach is extremely useful when we are interested in the tail of a distribution and not the median/mean. This topic has a large number of applications in many disciplines, such as engineering, finance and earth sciences.

Extreme value theory usually focuses on the upper tail of a distribution but the minimum of a sequence can be considered as the maximum of minus of the sequence so the same method can be applied whether we are interested in the maximum or minimum. The use of EVT for modelling the minimum value is illustrated in [1], which looks at the likelihood of a drought in Zimbabwe. Their motivation is much similar to ours in that since agriculture is a large part of their economy, it is important to study the probability of a drought in the country. They model the mean annual rainfall, which is the average of 62 weather stations in the country. This lets them avoid the question of spatial correlations between regions and independence between the annual measurements, hence they have a simple univariate problem.

A number of R packages are available to perform some part of an extreme value analysis:

- evd <https://cran.r-project.org/web/packages/evd/evd.pdf>
- extRemes <https://cran.r-project.org/web/packages/extRemes/extRemes>.

pdf

- fExtremes <https://cran.r-project.org/web/packages/fExtremes/fExtremes.pdf>
- gPdtest <https://cran.r-project.org/web/packages/gPdtest/gPdtest.pdf>
- isevd <https://cran.r-project.org/web/packages/ismev/ismev.pdf>
- ismev <https://cran.r-project.org/web/packages/ismev/ismev.pdf>

## 6.1 Issues arising in applying extreme value statistics to rainfall data

However, there are a few issues arising in applying EVT to rainfall data:

- Rainfall on the original scale may not be an ideal measure, especially if there are any observed measurements of 0.
- As we said in Sec. 2.1, rainfall alone does not capture the occurrence of a drought. It does not take into account other factors such as ENSO and IOD, which also affect the occurrence of a drought.
- The use of WRSI may be appropriate as it is an objective measure and captures all these aspects but it is measured between 0 and 100 so a transformation is required to apply EVT.

## 6.2 Summary of theoretical foundations of Extreme Value Theory

Let's assume that the annual rainfall measurements from a particular region are independent from one year to the next with no spatial correlations with neighbouring regions and consider the univariate case. Suppose  $X_1, X_2, \dots, X_n$  are independent and identically distributed random variables with a common cumulative distribution function  $F$ . Let  $M_n = \max\{X_1, X_2, \dots, X_n\}$  to represent the maximum of the  $n$  variables. Then

$$\begin{aligned}\Pr(M_n \leq x) &= \Pr(X_1 \leq x, X_2 \leq x, \dots, X_n \leq x) \\ &= F^n(x).\end{aligned}$$

If our interest was in the minimum, then we can consider it to be the maximum of minus of all  $X_i$  as in  $m_n = \max\{-X_1, -X_2, \dots, -X_n\}$  so the same method applies.

Unfortunately,  $F$ , which is the exact distribution of the maximum, is unknown but we can investigate how  $F^n$  behaves as  $n$  tends to  $\infty$ . The Fisher-Tippett-Gnedenko theorem applies if real values  $a_n > 0$  and  $b_n$  exist such that

$$\lim_{n \rightarrow \infty} \Pr\left(\frac{M_n - b_n}{a_n} \leq x\right) = F^n(a_n x + b_n) \rightarrow G(x), \quad (5)$$

where  $G$  is nondegenerate. If constants  $a > 0$  and  $b$  exist such that two extreme value cdfs  $G$  and  $G^*$  are  $G^*(x) = G(ax + b)$  for all  $x$ , then we say that  $G$  and  $G^*$  are of the same type. The Fisher-Tippett-Gnedenko theorem states that if the nondegenerate  $G$  exists, it belongs to either the Gumbel (Type I),

$$G(x) = \exp \left\{ - \exp \left[ - \left( \frac{x-b}{a} \right) \right] \right\} \text{ for } x \in \mathbb{R}, \quad (6)$$

the Frechet (Type II)

$$G(x) = \begin{cases} 0 & x \leq b \\ \exp \left[ - \left( \frac{x-b}{a} \right)^{-a} \right] & x > b \end{cases}, \quad (7)$$

or the Weibull (Type III)

$$G(x) = \begin{cases} \exp \left\{ - \left[ - \left( \frac{x-b}{a} \right) \right]^a \right\} & x < b \\ 1 & x \geq b \end{cases}. \quad (8)$$

These three types of distributions can be combined into a single distribution called the generalised extreme value distribution  $\text{GEV}(\mu, \sigma, \xi)$  with the CDF:

$$G(x) = \exp \left[ - \left( 1 + \xi \frac{x-\mu}{\sigma} \right)^{-1/\xi} \right], \quad (9)$$

defined when  $\mu, \xi \in \mathbb{R}$ ,  $\sigma > 0$  and  $1 + \xi(x-\mu)/\sigma > 0$ .  $\mu$  is the location parameter,  $\sigma$  is the scale parameter and  $\xi$  is the shape parameter that controls the tail behaviour of Eq. (9).

When  $\xi > 0$ , the above corresponds to the Frechet (Type II). When  $\xi < 0$ , it corresponds to the Weibull (Type III). As  $\xi \rightarrow 0$ , it reduces to the Gumbel (Type I). Fig. 10(A) plots  $\text{GEV}(\mu, \sigma, \xi)$  with three different  $\xi$  values, i.e.,  $\{-0.5, 0, 0.5\}$ , keeping  $\mu = 0$  and  $\sigma = 1$  fixed. The Gumbel (green) has an exponential tail. The Frechet (blue) has a heavy tail. The Weibull (red) has a short tail with an upper bound.

Alternatively, we can consider the likelihood of observing  $X$  that exceeds a particular threshold  $u$ . The distribution function of exceedances above threshold  $u$  is:

$$F(x) = \Pr(X - u \leq x | X > u)$$

for  $x \geq 0$ . The distribution of the exceedances  $\{X_1 - u, X_2 - u, \dots, X_{n_u} - u\}$  provided  $X_j \geq u$  for all  $j$  can be approximated by a generalised Pareto distribution  $\text{GPD}(\mu, \sigma, \xi)$  [5].

$$G(x) = \begin{cases} 1 - \left( 1 + \frac{\xi(x-\mu)}{\sigma} \right)^{-1/\xi} & \xi \neq 0 \\ 1 - \exp \left( - \frac{x-\mu}{\sigma} \right) & \xi = 0 \end{cases}$$

defined for  $x \geq u$  when  $\xi \geq 0$  or  $0 \leq y \leq -\sigma/\xi$  when  $\xi < 0$ .  $\mu, \xi \in \mathbb{R}$  and  $\sigma > 0$ .  $\mu$  is the location parameter,  $\sigma$  is the scale parameter and  $\xi$  is the shape parameter as before.

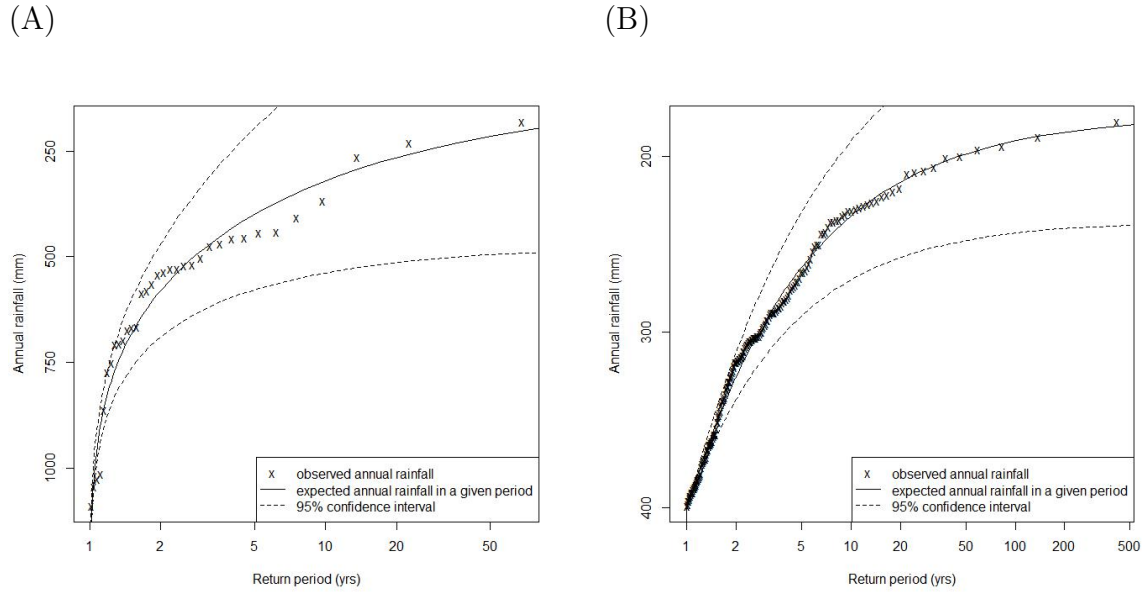


Figure 11: (A) Return periods for annual rainfall for East Harerghe computed using the block maxima method. (B) Return periods for annual rainfall for East Harerghe computed using peaks over threshold method. Note that the vertical axes are decreasing from bottom to top so the top right of each plot corresponds to extremely low levels of annual rainfall.

### 6.3 A simple extreme value analysis on Ethiopian rainfall data

We apply the above theory to 34 years of rainfall data from 196 locations in East Harerghe, Ethiopia, between 1983 and 2016. We chose this area because it was strongly affected by the catastrophic famine that struck Ethiopia in the early 1980's. We were interested to see the extent to which this event was reflected in the data. The observations were made three times a month, a total of 36 measurements each year. And we define the observed annual rainfall to be the sum of all 36 observations from that calendar year, for each location.

The left line plot in Fig. 10(B) shows the observed annual rainfall measurements from five randomly selected locations in East Harerghe. There seem to be some spatial and temporal correlations as expected. We aggregated annual measurements from all 196 locations into the histogram on the right of Fig. 10(B). It shows a right-skewed distribution but our interest is in the left tail.

#### The block maxima method:

The block maxima is a popular approach that breaks up the sequence of data into blocks of an equal length (e.g., one year) and uses only the maximum observed measurement from each block. So with our rainfall data, because we're interested in the minimum annual rainfall, for each year, we extract the minimum annual rainfall

measurement across the region. The table below shows the maximum likelihood estimates of the parameters of the fitted generalised extreme value distribution on the negative of the annual rainfall data and their corresponding standard errors. The shape parameter  $\xi$  is negative, indicating that the annual rainfall data belong to the Weibull family.

Parameter	Estimated value	StdError —
Location	-660	44
Scale	236	32
Shape	-0.4	0.1

A return level plot showing the expected return period for observing annual rainfall at different thresholds is shown in Fig. 11(A). The return period is defined as the inverse of the probability of observing rainfall at most the corresponding threshold. The expected return period for observing annual rainfall less than 320mm is 10 years, with a 95% confidence interval of 538mm and 51mm. Similarly, the expected return period for observing annual rainfall less than 215mm is 50 years. In other words, in a given year, the probability of observing annual rainfall of 320mm or less (worse) is 0.1 and the probability of observing annual rainfall of 215mm or less is 0.02. It was surprising to us that the rainfall level observed in 1983, at around 500, is not particularly extreme given the severity of the famine at that time. It would be interesting to investigate further to understand why since it suggests that WRSI is not necessarily a good indicator of the socio-economic impact of low rainfall.

#### **The peak over threshold method:**

The biggest disadvantage of the block maxima method is that it only uses the one (minimum) rainfall measurement from each block (year) and 'wastes' the rest of the measurements. An alternative method to avoid this is the peak over threshold method, which fits the generalised Pareto distribution. Here, we consider the likelihood of observing annual rainfall given that it does not exceed some preset drought threshold  $u$ . Here, we picked  $u = 400$ mm because this is considered to be a drought in Zimbabwe in [1].

The table below shows the maximum likelihood estimates of the parameters of the fitted generalised Pareto distribution, with a threshold  $u = 400$ , on the negative of the annual rainfall data and their corresponding standard errors.

Parameter	Estimated value	StdError —
Scale	132	10
Shape	-0.6	0.05

The corresponding return time plot is shown in Fig. 11(B) and is broadly consistent with the results obtained from the block maxima method, given the uncertainties.

## 7 Conclusions and possibilities for future work

We have done some exploratory data analysis on rainfall and WRSI data provided by ARC to assess the predictability of droughts at a number of locations in Africa. Based on some simple assumptions on the pricing of insurance contracts, we assessed the appropriateness of the historical WRSI benchmarks set by ARC members in the light of the historically observed values of WRSI for different regions. We conclude that the benchmarks are broadly sensible. Subject to the assumptions made in Sec. 3, ARC is probably slightly generous to its members in setting premiums. In Sec. 4 we compared a number of standard linear time series models for forecasting the WRSI values from the WRSI time series alone. The models tested were the Simple Moving Average, Moving Median, Simple Exponential Smoothing, Holt Winters with linear trend, Holt Winters with linear trend and simple seasonality and TBATS. This analysis produced mixed results and could be done more systematically. A fairly robust conclusion is that the simpler forecasting methods (Simple Moving Average and Moving Median) produced better forecasts than the more complex models. Although we did not have time to do a detailed statistical analysis, we suspect that the uncertainties in the WRSI forecasts are very large. As an alternative, in Sec. 5 we explored in a more systematic way the forecasting of the rainfall time series underlying the WRSI. Due to the strong seasonality of this data, we assessed the Seasonal AutoRegressive Integrated Moving Average (SARIMA) and TBATS models since both are capable of capturing complex seasonality in time series data. We found that both models capture the seasonal structure of the data and are able to make forecasts that encode this structure. The TBATS model was preferred (at least for the Ethiopian rainfall data we examined) since it was able to capture more of the temporal correlations in the data than the SARIMA model. We conclude that it would probably be informative to pass forecasts of rainfall data through the WRSI calculation in order to obtain predictions and confidence intervals for WRSI rather than trying to forecast the WRSI time series directly. Our attempts to detect the influence of ENSO on the rainfall levels were inconclusive, probably because we examined a region only weakly affected by ENSO. Finally in Sec. 6, we showed a simplified application of Extreme Value Theory to the rainfall data for Ethiopia to estimate return periods of extreme low rainfall values. Our main conclusion from this analysis was that we were surprised that the catastrophic famine of the early 1980's did not appear as a particularly extreme event in the rainfall data we examined. This also warrants further investigation.

In terms of future work, the most important direction to explore would be to properly account for the spatial extended nature of the rainfall data as was done for extremes of temperature in Australian droughts in [8]. It would be productive to adapt the methodology of [8] to describe extremes of low rainfall in Africa. Another direction for future work would be to improve the time series models for rainfall since the models proposed here do not respect some basic properties such as positivity of the data. It is also important to establish how to incorporate external climate signals like ENSO into these forecasts. We have proposed that using

such rainfall forecasts to produce WRSI forecasts could be informative. We suggest exploring how uncertainties (including uncertainties in static input parameters) propagate through the WRSI calculation. Finally we would like to understand why our extreme value analysis did not pick out the early 1980's drought as having been particularly extreme in the Ethiopian data. The catastrophe that unfolded in that period seems like a natural historical "stress test" scenario for the ARC system so it would be useful to understand whether data analytics would detect it if a comparable event were to reoccur today.

Finally, another idea which we discussed at the study group but did not have the opportunity to develop is to build a low dimensional dynamical system model to help reason about the coupling of the different large scale climate processes influencing the African climate. Such models are usually too simplified to provide robust forecasting but by combining such a model with a data assimilation algorithm, it could be possible to make informative inferences about the effect of multiple influences such as coupling the incommensurate forcing from the Indian and Atlantic Oceans. That is, on each coast: Wind direction and the temperature and humidity of the air. Each affecting the 4 major climate regions identified in [7], with some coupling between the regions. First a linear model with the different forcings to see how that could be parameterised. Then add some non-linearity. Parameters would be tuned based upon the statistical data.

## References

- [1] D. Chikobvu and R. Chifurira. Modelling of extreme minimum rainfall using generalised extreme value distribution for Zimbabwe. *South African Journal of Science*, 111(9-10):01–08, 2015.
- [2] A. M. De Livera, R. J. Hyndman, and R. D. Snyder. Forecasting time series with complex seasonal patterns using exponential smoothing. *J. Amer. Stat. Assoc.*, 106(496):1513–1527, 2011.
- [3] R. J. Hyndman. Thoughts on the Ljung-Box test. <https://robjhyndman.com/hyndsight/ljung-box-test/>.
- [4] Y. Min and A. Agresti. Modeling nonnegative data with clumping at zero: a survey. *Journal of the Iranian Statistical Society*, 1(1):7–33, 2002.
- [5] J. Pickands III. Statistical inference using extreme order statistics. *Ann. Statistics*, pages 119–131, 1975.
- [6] S. Soltani, R. Modarres, and S. S. Eslamian. The use of time series modeling for the determination of rainfall climates of Iran. *International Journal of Climatology*, 27(6):819–829, 2007.
- [7] M. Tadross and T. Hauser. Review of Africa-wide regional weather/climate co-variability and associated processes. Technical report - unpublished.

- [8] H. C. Winter, J. A. Tawn, and S. J. Brown. Modelling the effect of the El Niño-Southern Oscillation on extreme spatial temperature events over Australia. *Ann. Appl. Stat.*, 10(4):2075–2101, 2016.