Projected Rise in Extreme Rainfall Intensity in Niamey, Niger: Implications for Future Climate.

Luke Dunkley

Module: MATH3001 Project in Mathematics

Student ID: 201429441

Supervisor: Douglas J. Parker

Date: 28/03/2024

UNIVERSITY OF LEEDS

You must sign this (digitally signing with your name is acceptable) and include it with each piece of work you submit.

I am aware that the University defines plagiarism as presenting someone else's work, in whole or in part, as your own. Work means any intellectual output, and typically includes text, data, images, sound or performance.

I promise that in the attached submission I have not presented anyone else's work, in whole or in part, as my own and I have not colluded with others in the preparation of this work. Where I have taken advantage of the work of others, I have given full acknowledgement. I have not resubmitted my own work or part thereof without specific written permission to do so from the University staff concerned when any of this work has been or is being submitted for marks or credits even if in a different module or for a different qualification or completed prior to entry to the University. I have read and understood the University's published rules on plagiarism and also any more detailed rules specified at School or module level. I know that if I commit plagiarism I can be expelled from the University and that it is my responsibility to be aware of the University's regulations on plagiarism and their importance.

I re-confirm my consent to the University copying and distributing any or all of my work in any form and using third parties (who may be based outside the UK) to monitor breaches of regulations, to verify whether my work contains plagiarised material, and for quality assurance purposes.

I confirm that I have declared all mitigating circumstances that may be relevant to the assessment of this piece of work and that I wish to have taken into account. I am aware of the University's policy on mitigation and the School's procedures for the submission of statements and evidence of mitigation. I am aware of the penalties imposed for the late submission of coursework.

Student Signature		Date	27/03/2024
Student Name	Luke Dunkley	Student Number	201429441

Contents

Abstract 1

1 Introduction 1

2 Methods 2

2.1 Data 2

2.2 Exploratory analysis 3

2.3 Discrete time Markov chain 4

2.4 Fitting the distribution of useful rainfall using the method of maximum likelihood 8

2.5 Fitting the generalised extreme value distribution to maximal rainfall using the method of maximum likelihood 10

3 Results 12

3.1 Exploratory analysis 12

3.2 Sahelian storm dynamics: 16

3.3 Intensity-duration analysis of wet spells 19

3.4 Extreme value analysis 22

4 Discussion 23

5 Conclusion 25

Ethical Considerations 26

Bibliography 26

Abstract

This paper analyses the effectiveness of convection permitting and parameterisation models in replicating the dynamics of Sahelian storms in Niamey. It then uses future runs of those models to try to predict how the climate will change in the future. This is done by comparing the intensity, duration, and meteorological dynamics of rainfall events in model runs with observational datasets. The focus of this paper is on extreme precipitation events and their frequency and likelihood. Rainfall data for 100-year periods is simulated using discrete time Markov chains to allow more accurate analysis of the return periods of extreme rainfall. It is hypothesized that an increase in global sea temperature, an expected result of global warming, will increase the likelihood and intensity of extreme rainfall events.

1 Introduction

Many countries in sub-Saharan Africa fail to include climate impacts in their short and long term plans due to a lack of knowledge and data, among other reasons (Jones et al., 2015). This paper aims to address this with the potential impacts of precipitation changes.

Niamey is the capital city of Niger, located south of the Sahara it is part of semi-arid region stretching across West Africa called the Sahel. The city of Niamey had a population of 397,437 in 1988 which is set to increase to 2,063,894 in 2030 (Manu et al., 2014). Around 99% of convection occurs in the Sahel's rainy season (D'Amato and Lebel, 1998) which is roughly between May and August. Niamey is situated on the banks of the Niger river, the third largest river in Africa (Mahamane, Oumarou and Piñeira Mantiñán, 2023) and a river that has experienced multiple flooding events since the end of the 20th century. Niamey district experienced 6 floods in the 18 years between 1998 and 2015 (Fiorillo et al., 2018). These floods resulted in the destruction of 20,000 houses and affected 160,000 people in the Niamey district alone (Fiorillo et al., 2018). As Niamey has high population growth rates, the need for people to live on areas with a high flood risk is increasing making the population more vulnerable (Fiorillo et al., 2018).

The cause of precipitation in Niamey in the wet season is the West African monsoon. Convection occurs in the Sahel as a series of convective storms (D'Amato and Lebel, 1998) referred to in this paper as Sahelian storms. Mesoscale convective systems (MCSs) that affect Niamey have often been generated over the Jos Plateau/AÏr mountain orography and the Damergou gap (Vizy and Cook, 2017). These MCSs are associated with extreme rainfall and strong winds (Schumacher and Rasmussen, 2020). The West African Sahel experiences interdecadal cycles in rainfall, since 1950 there have been 3 cycles, a wet epoch, a dry epoch and a moderate recovery (Bigi, Pezzoli and Rosso, 2018).

This analysis uses two climate models, one is convection parameterising, and the other is convection permitting. Stratton et al. (2018) suggests that convection permitting models have an increased dry bias over the Western Sahel, where Niamey is situated, and that the convection permitting models have brighter deep convective cores. Furthermore with a

particular focus on flooding, Berthou et al. (2019) suggests that CP4 is better at predicting rainfall intensity. In Niamey, both models underestimate rainfall intensity (Berthou et al., 2019). To analyse how effective the models are at predicting precipitation characteristics, they are compared to observational datasets from the region.

For each model there is a current and future simulation. Fitzpatrick et al. (2020b) shows that the two models being used in this paper predict that in August, the peak monsoon month for Niamey, temperatures will increase by between 3-5℃ in the future model runs. The Clausius Clapeyron relationship between temperature and rainfall suggests that the amount of moisture the air can hold will increase by around 6-7% per ℃ (Chan et al., 2016). Analysing whether this will increase precipitation, particularly the intensity of extreme precipitation, will be the focus of this investigation.

Mouhamed et al. (2013) and Panthou, Vischel and Lebel (2014) suggest that extreme rainfall events have become more common in the Sahel in recent times, therefore it is investigated to see if this pattern is set to continue in a future climate. Extreme rainfall is a key ingredient in the increase of flooding events (Kendon et al., 2014), therefore increases in the intensity of extreme rainfall would worsen the recent trend of more frequent and intense flooding in the region (Fiorillo et al., 2018). Highlighting the flood risks is vital as only 61 out of 297 flood damage models in the region were statistically significant (Fiorillo et al., 2018). Therefore this paper explores how changes in precipitation trends could affect the flood risk in Niamey.

2 Methods

2.1 Data

2.1.1 Model data

This analysis uses two climate models: The Met Office Unified Model Convection Permitting for Africa (CP4-A) and the R25-Africa Regional Model. The major difference between the two models is the resolution. The CP4 model as it is referred to in this paper has a resolution of 4.5km whereas the R25 model has a 25km resolution. The resolution determines how convection is modelled. The smaller resolution allows the CP4 model to be convection permitting, whereas the R25 model requires convection parametrisation. Data has been extracted from each model for two ten-year simulation runs, a current climate run and a future climate run. The current climate run for both models is programmed with the sea surface temperatures from Reynolds observational dataset (Reynolds et al., 2007), this run is supposed to simulate the climate from 1996 to 2006 . The future model simulation is based off an assumption made with high confidence by the Inter-governmental Panel for Climate Change (IPCC) that if we do not make active changes, the global temperature will warm by 3-6°C (Fitzpatrick et al., 2020b). Therefore, the future models have a blanket increase in sea surface temperature of 4°C to those in Reynolds et al. (2007). For this analysis, the model

data is only retrieved for Niamey, Niger. To note when looking at results, the model treats all months as 30 days and consequently all years as 360 days.

2.1.2 Observational data

Figure 1: Observation locations, map of Niamey, Niger (13.5116° N, 2.1254° E) and the surrounding area (Google Maps, 2024).

The observational data used in the analysis is from 3 stations around Niamey: Niamey Airport, Niamey ORSTOM and Banizoumbou. All three observation datasets are from the AMMA-CATCH observatory (AMMA-CATCH, 1990). Each observational data set contains some missing values, for means these are just ignored in the analysis and for useful rainfall analysis they are treated as 0's, this is because the majority of missing values are in the dry season anyway for precipitation. The data included in the analysis is analogous to the years of the current model runs, to allow for direct comparisons. Except for the composite analysis, where the data was unavailable for that time period so the ten-year period of 2005-2014 was used.

2.2 Exploratory analysis

2.2.1 Summary statistics

Basic summaries of the data are included in the results. Any mean described is the arithmetic mean and any error bars included on mean plots are the standard error from the mean. Standard error is defined as the standard deviation of the sample divided by the square root of the sample size.

Probability density functions are calculated using histograms where each bin has width 1mm, the density of each bin is multiplied by precipitation and then plotted against precipitation (daily precipitation totals) to see what proportion of the total rainfall each rainfall bin provides. The datasets are fit to common statistical distributions using methods described later, and the results for these distributions are superimposed onto the plots.

2.2.2 Composite analysis

Composite analysis of precipitation is a method that models the meteorological dynamics around a rainfall event, in this context a Sahelian storm. The analysis requires the identification of the start time of significant precipitation, this is defined as the hour when there is useful rainfall given that the hour before did not have useful rainfall. Useful rainfall is defined as an hourly precipitation value greater than 1mm, for this analysis any value of precipitation less than 1mm is set to 0mm, this threshold was also used in Klutse et al. (2018) and D'Amato and Lebel (1998). For each variable of interest, precipitation, temperature, and wind modulus, the means of each hour around the rainfall event are computed to produce a "mean" precipitation event for each model/observation dataset.

2.2.3 Intensity-duration analysis

This analysis requires the data to be split into a series of wet spells. These wet spells are defined as one or more days with daily rainfall greater than 1mm. The characteristics of these wet spells are examined in this section, particularly duration and intensity. Duration is the length in days of the wet spell and intensity is chosen firstly to be the maximum daily rainfall in the model, as was chosen in Kendon et al. (2014). The plots are 2 dimensional density plots with a colour scale to show the density. The number of bins was chosen to separate each duration into its own block. As there is 1 extra bin it causes a slightly misleading gap in some of the graphs that makes it look like a certain duration has no density, when it might not be the case. However the gain in being able to examine each duration separately outweighs this negative.

Table 3 and 4 analyse the mean cumulated rainfall of each spell, cumulated rainfall is defined as the total rainfall over the spell. To test the difference between different means in table 4, a Welch two sample t-test is used as the data is not paired. It is a two sided test and has the following null and alternative hypotheses

Equation 1

$$H_0 = The\ two\ data\ sets\ have\ the\ same\ mean$$

$$H_1 = The\ two\ data\ sets\ do\ not\ have\ the\ same\ mean$$

therefore, a rejection indicates that the means are different at the level of significance indicated.

2.3 Discrete time Markov chain

Creating a Markov chain of rainfall is a way to synthesize many years of precipitation data at much less computational expense than running the model for that amount of years. The type of Markov chain that is considered here is the discrete time two state Markov chain. This is a stochastic model that differentiates between the two states of "useful rainfall" defined in the model by state 1 and "no rainfall" defined in the model as state 0. For this analysis, useful rainfall μ is defined as rainfall greater than 1mm per time period (day or

hour) to be consistent with the other analysis in the paper. The Markov property that differentiates Markov chains from other stochastic models is described by

Equation 2

$$\mathbb{P}\{X_t|X_{t-1}\} = \mathbb{P}\{X_t|X_{t-1}, X_{t-2}, X_{t-3}, \dots\}.$$

This "memoryless property" states that the only determining factor in which state the chain goes to is the current state that it is in. As this is a two state Markov chain it requires 4 transition probabilities often described using a transition matrix A in Equation 3. A transition probability, say p_{01} , describes the probability that you move to state 1, given that you are in state 0. The subscript describes the state that chain has moved from and to.

Equation 3

$$A = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}.$$

To calculate these parameters the number of transitions between each state is computed. Here *n* is the total number of daily rainfall variables that the dataset has, and t denotes each day, with μ defined above. Now, n_{01} is the number of times that that there is a day with useful rainfall (state 1) directly following a day with no rainfall (state 0).

Equation 4 (a,b,c,d)

$$n_{00} = \sum_{t=1}^{n} (\text{rainfall}_t < \mu \mid \text{rainfall}_{t-1} < \mu)$$

$$n_{01} = \sum_{t=1}^{n} (\text{rainfall}_t > \mu \mid \text{rainfall}_{t-1} < \mu)$$

$$n_{10} = \sum_{t=1}^{n} (\text{rainfall}_t < \mu \mid \text{rainfall}_{t-1} < \mu)$$

$$n_{11} = \sum_{t=1}^{n} (\text{rainfall}_t > \mu \mid \text{rainfall}_{t-1} > \mu).$$

The transition probability estimators are then the calculated using the maximum likelihood estimation (MLE) formula for two state Markov parameters (Wilks, 2010). This is the probability that the chain moves from a given state to the state described.

Equation 5

$$\hat{p}_{ij} = \frac{n_{ij}}{n_{i\bullet}}$$

where $n_{i\bullet} = n_{i0} + n_{i1}$ and $n_{\bullet j} = n_{0j} + n_{1j}$ and i and j can take the value 0 or 1 for this two-state Markov chain. To synthesize a set of binary data for rainfall/non rainfall events the

transition matrix determines which state you will be in at time t+1 given the state that you are in at time t. This is done by generating a random variate (0 or 1) based on the transition probabilities in the row of the transition matrix that the chain is in (the first row is state 0 and the second row is state 1).

The statistical significance of the Markov chain is computed using a goodness of fit test. This test takes the null hypothesis to be that the data is serially independent (a set of Bernoulli random variables) with the converse true for the test hypothesis (Wilks, 2010).

Equation 6

$$H_0 = The\ data\ series\ is\ serially\ independent$$

$$H_1 = The\ data\ series\ is\ generated\ by\ a\ Markov\ chain$$

This test compares the expected value of each parameter if it was a Bernoulli random variable (H_0), with the observed count. The expected value is calculated using the following formula (Wilks, 2010)

Equation 7

$$e_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}.$$

The test statistic χ^2 for the two state Markov chain is then calculated by squaring the difference between observed counts and expected counts (if the data was serially independent), before dividing by the expected amount, then summing this value for all observed combinations (Wilks, 2010).

Equation 8

$$\chi^2 = \sum_{i=0}^{1} \sum_{j=0}^{1} \frac{(n_{ij} - e_{ij})^2}{e_{ij}}.$$

2.3.1 Daily Markov chain

The first application of a Markov chain is to synthesize daily rainfall. This method was done extensively and successfully in Wilks (1999). The two-state first order Markov chain proved the most effective stochastic model in replicating rainfall in North America when comparing BIC statistics (Wilks, 1999). The binary Markov chain was computed as described above, however the data was split into monthly sections because of the strong seasonal rainfall cycle described in Figure 4. Table 1 shows that many monthly Markov chains reject the null hypothesis. However, not rejecting the null hypothesis simply means that the data does not strongly exhibit the Markov property described in equation 2. This does not mean that it won't simulate the data effectively as qualitative checks of monthly means showed. However other methods could be explored such as an autoregressive model as was used in Rebora et al. (2006).

Table 1: χ^2 significance results (at a 95% confidence level) for daily Markov chain parameters. Separate tests done for each model/observation data set and for each month.

Month	CP4_C	CP4_F	R25_C	R25_F	Niamey Airport	Banizoumbou	Niamey ORSTOM
January	NA	NA	NA	Accept H0	NA	NA	NA
February	NA	NA	Accept H0	NA	NA	NA	NA
March	NA	Accept H0	Accept H0	Reject H0	Accept H0	Accept H0	Accept H0
April	Accept H0	Accept H0	Reject H0	Reject H0	Reject H0	Accept H0	Accept H0
May	Accept H0	Accept H0	Reject H0	Reject H0	Accept H0	Accept H0	Accept H0
June	Reject H0	Reject H0	Accept H0	Reject H0	Accept H0	Accept H0	Accept H0
July	Accept H0	Accept H0	Reject H0	Reject H0	Accept H0	Accept H0	Accept H0
August	Accept H0	Accept H0	Reject H0	Reject H0	Accept H0	Accept H0	Reject H0
September	Accept H0	Accept H0	Reject H0	Reject H0	Accept H0	Accept H0	Accept H0
October	Reject H0	Accept H0	Reject H0	Reject H0	Reject H0	Accept H0	Reject H0
November	Accept H0	Reject H0	NA	Reject H0	NA	NA	NA
December	NA	NA	NA	NA	NA	NA	NA

Table 1 shows that the two R25 model runs display the Markov property most strongly. However, the observations and CP4 model runs only have a few months that display the strong Markov property, indicating that the dynamics of the CP4 model are more similar to the observations than R25. Not displaying the Markov property strongly especially in the wet season indicates that rain events in the Sahel region occur on a time span of less than a day. Which is confirmed by the composite analysis in Figure 9 and 10 as after 24 hours the precipitation is back to its mean value. Therefore, rainfall on a daily scale is just as accurately modelled by a Bernoulli distribution.

For the extreme value analysis, 100 years of precipitation data was synthesized (36000 days) from this Markov chain.

2.3.2 Hourly Markov chain

The significance findings from the daily Markov chains lead the creation of a hourly Markov chain instead. This Markov chain was created from the data from the rainy season only defined here as June to September. As table 2 shows, the hypothesis that the Markov property applies to useful hourly rainfall data is true for all observational data and model iterations. Therefore, for the extreme value analysis, a Markov chain consisting of 100 rainy seasons was synthesized. Here the Markov chain was not fit monthly, and neither was the

gamma distribution for useful hourly rainfall. This is because all months have similar rainfall trends as they are all in the rainy season and the value of extra data to get a better distribution fit is preferable to modelling the small amount of monthly variability in this dataset.

Table 2: χ^2 significance results (at a 95% confidence level) for the wet season (JJAS) hourly Markov chain parameters, separate tests done for each model/observation data set.

CP4_C	CP4_F	R25_C	R25_F	Niamey Airport	Banizoumbou	Niamey ORSTOM
Reject H0	Reject H0	Reject H0	Reject H0	Reject H0	Reject H0	Reject H0

2.4 Fitting the distribution of useful rainfall using the method of maximum likelihood

Once the binary Markov chain is created, the second stage of creating a synthetic data set is to fit a distribution to the precipitation amounts. The first step in this process is to analyse the distributions of each dataset using a skewness-kurtosis graph (Cullen and Frey, 1999). As we only want to fit it for when rainfall is above 1mm, we fit the distribution only to useful

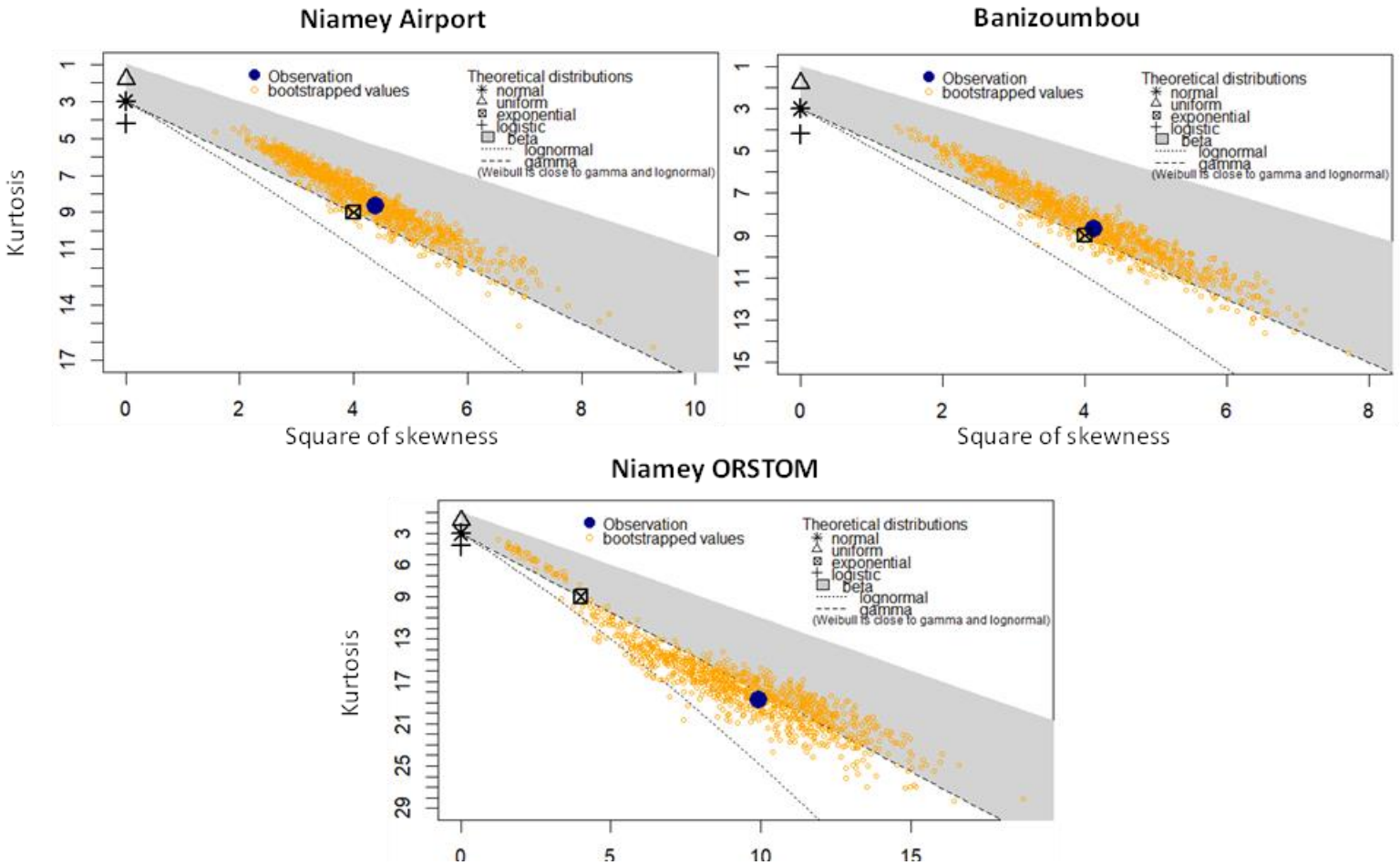

Figure 2: Cullen and Frey graph of kurtosis vs skewness along with 1000 bootstrapped samples for each useful rainfall dataset (precipitation above 1mm) for the four models: CP4 (right), R25 (left), current climate (top) and future climate (bottom).

rainfall values. The *"fitdistrplus"* package in r (Delignette-Muller, 2023) allows you to plot the position of each empirical distribution on this graph in relationship to where common distributions lie. To account for the uncertainty in the data the function *"descdist"* can bootstrap the data set to predict where simulated data with similar properties to that of the observations would lie as done in Pouillot and Delignette-Muller (2010). As Figure 2 and 3 and show, the Cullen and Frey graphs indicate that the distribution of useful rainfall can be predicted using a gamma distribution, particularly for all 3 observational data sets, and all model iterations except the CP4 Future run which seems to be better modelled by a beta distribution. This is interesting as the current CP4 run replicates the distribution of the observational data, and CP4 is agreed to model convection more accurately, therefore this could indicate that future precipitation is likely to follow a different distribution.

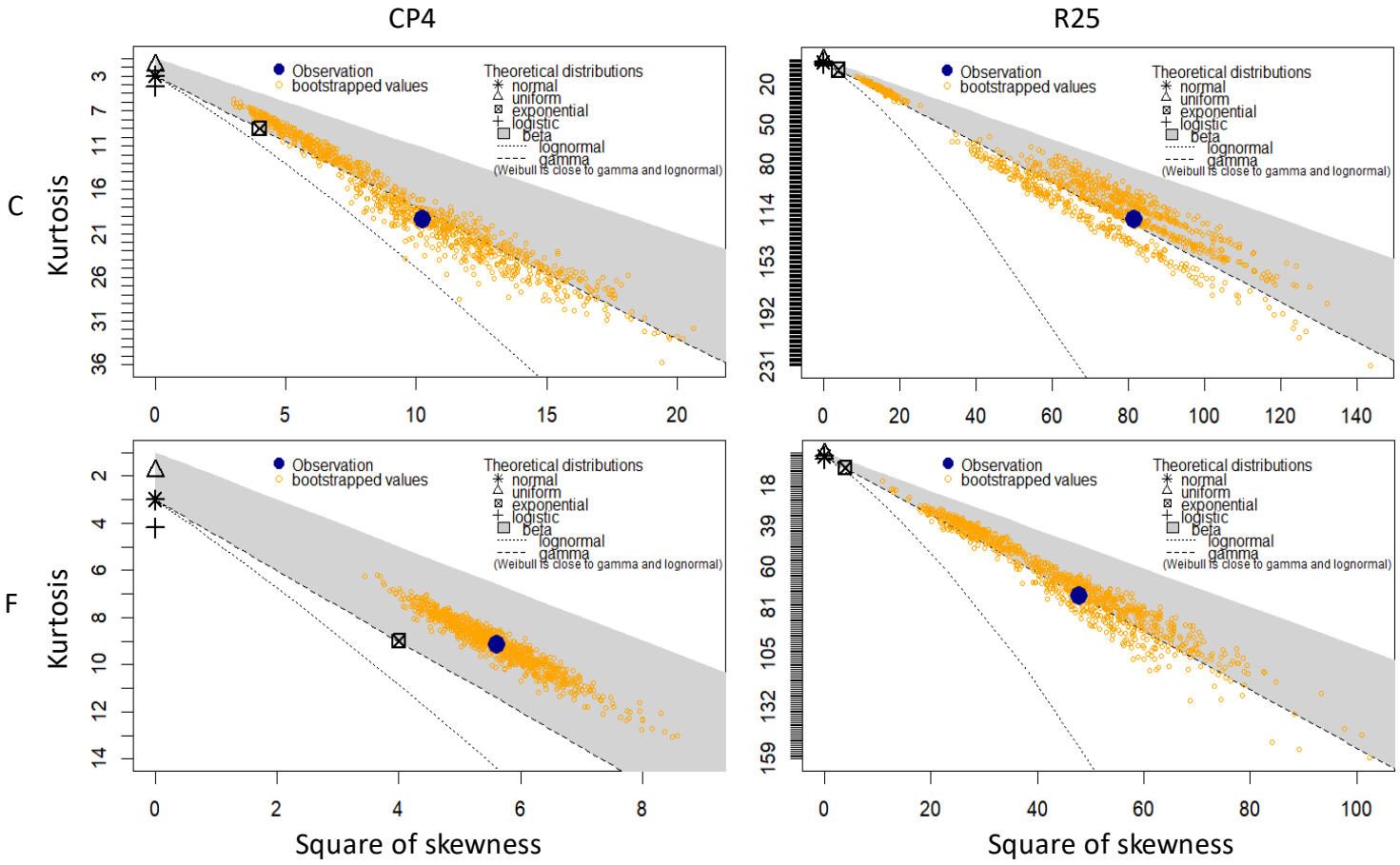

Figure 3: Cullen and Frey graph of kurtosis vs skewness along with 1000 bootstrapped samples for the useful rainfall datasets (precipitation above 1mm) for each of the three observation locations.

Gamma distributions are commonly used to model precipitation (Wilks, 1999) however in it has been shown that another distribution fit well to precipitation data is the mixed exponential distribution (Wilks, 1999) and this is not shown on the Cullen and Frey graphs. Fitting a mixed exponential distribution requires numerical methods which were too computationally advanced for this analysis, and the gamma distribution is qualitatively accurate.

The gamma distribution has the probability density function

Equation 9

$$f(x) = \frac{(x/\beta)^{\alpha-1}\exp(-x/\beta)}{\beta\Gamma(\alpha)} \quad \alpha, \beta, x > 0$$

(Wilks, 2010) where $\Gamma(\alpha)$ Is the gamma function

Equation 10

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt.$$

The gamma distribution was again fit using the method of maximum likelihood estimators. This method aims to calculate the values of the shape parameter β and scale parameter α that maximise the likelihood function. The likelihood function for a given distribution is the product of the likelihood functions for each individual observation. For a single observation x the likelihood function is identical to the pdf for that distribution (Wilks, 2010) with the difference being that in the pdf, x is variable whereas in the likelihood function the parameters β and α are variable and x is fixed. Therefore, the overall likelihood function for the set of precipitation values $(x_i)_{i=1,\dots,n}$, where n is the total number of precipitation values, is

Equation 11

$$\Lambda(\beta, \alpha) = \prod_{i=1}^{n} \frac{(x_i/\beta)^{\alpha-1}\exp(-x_i/\beta)}{\beta\Gamma(\alpha)}.$$

Calculating the exact values of β and α can only be done iteratively. This is possible using the function "*egamma*" from the package "*envstats*" (Millard and Kowarik, 2020).

To check the goodness of fit of daily useful rainfall to the gamma distribution, the Anderson-Darling test was performed. This test was chosen as it is a viable test of the goodness of fit of continuous distributions where the parameters are estimated from the data unlike other tests such as the Kolmogorov-Smirnov test (Özge Karadağ and Serpil Aktaş, 2016). This treats the null to be that the dataset does come from the gamma distribution and the alternative to be that it doesn't. This test was done using the command *"ad.test"*, from the "goftest" package (Faraway et al., 2021). The p-values are included on Figure 5 and 6 for each daily useful rainfall dataset, low p-values indicate worse fits to the data.

2.5 Fitting the generalised extreme value distribution to maximal rainfall using the method of maximum likelihood

The generalised extreme value (GEV) distribution can be used to predict how often an extreme daily precipitation event is to occur. The data that is fit to the distribution is a set of

maximal daily precipitation values, these can either be the maximum value of each period, in this case a year, or the m maximal values of the entire dataset. A similar analysis was done in Elagib et al. (2021) however they used the Weibull and Gumbel extreme value distributions instead of the GEV. For this analysis the singular maximum yearly rainfall value in a single day is chosen as the set of maximal values as was done in Chapman et al. (2022). The GEV has probability density function

Equation 12

$$f(x) = \frac{1}{\beta}\left[1 + \frac{\kappa(x-\zeta)}{\beta}\right]^{1-1/\kappa} \exp\left\{-\left[1 + \frac{\kappa(x-\zeta)}{\beta}\right]^{-1/\kappa}\right\}, \qquad 1 + \frac{\kappa(x-\zeta)}{\beta} > 0$$

(Wilks, 2010) and cumulative distribution function

Equation 13

$$F(x) = \exp\left\{-\left[1 + \frac{\kappa(x-\zeta)}{\beta}\right]^{-1/\kappa}\right\}$$

(Wilks, 2010) where ζ is the location parameter, β is the scale parameter and κ is the shape parameter.

The data is fit using the method of maximum likelihood, this process is explained in the previous section and the likelihood function is calculated in the same way. This was chosen over the method of L moments like in Chapman et al. (2022). The maximum likelihood estimators are again estimated iteratively, this was possible using the function *"fgev"* from the package *"evd"* (Stephenson and Ferro, 2018). This function also calculates the standard error of the estimators, which is used to plot error regions in Figure 12. The extreme value analysis is used to quantify how unlikely certain rain events are, or, given a probability of it happening, how much rainfall there will be. The second of those statements is calculated using the inverse of the CDF called the quantile function. This is of the form

Equation 14

$$F^{-1}(p) = \zeta + \frac{\beta}{\kappa}\{[-\ln(p)]^{-\kappa} - 1\}$$

(Wilks, 2010). Here the input is the probability. This formula was used in conjunction with the average return period formula

Equation 15

$$R(x) = \frac{1}{\omega[1 - F(x)]}$$

(Wilks, 2010) to create the rainfall return period plots. Here ω is the time scale of the maximal events that were chosen, in this case $\omega = 1\, year^{-1}$ and $F(x)$ is the cumulative

distribution function. This formula transforms the probabilities into return periods. For example, an extreme event that has a 50% cumulative probability would have a return period of 2 years whereas an extreme event with 95% cumulative probability would have a return period of 20 years. For this analysis the return period formula was inverted to calculate the cumulative probabilities of each yearly return period for year 2-100. This includes the right tailed, lower probability, high cumulative probability events, which are the extreme events focused on in this investigation. These probabilities are inputted into the quantile function to create the plots.

This analysis was performed with the ten year long data sets to begin with but the lack of sample size made the results have qualitatively unlikely features and they all had extremely large error ranges. Therefore 100 year simulations of the Markov chains, as described in methods 3.3, were created in order to decrease the uncertainty by increasing the sample size (from 10 to 100). However, this means that the GEV distribution becomes very dependent on the gamma distribution that is fit to each Markov chain, so any inaccuracies in the model will affect the final results, but comparisons can still be made even if the values are not exact. The only observational simulated dataset compared to is the Niamey Airport Markov chain, this is for ease of viewing and all observational models had similar results.

As will be discussed, extreme rainfall values can be the result of different hydrological processes, each of which in theory may come from a different distribution (Wilks, 2010). Therefore, this could invalidate the assumption underpinning extreme value theory that all events come from the same distribution (Wilks, 2010). The goodness of fit of the data to the GEV distribution is checked using the Anderson-Darling as was done in Chapman et al. (2022). This test treats the null hypothesis to be that the dataset does come from the GEV distribution. This test was done using the same commands as in methods 2.4, again setting in the function that the parameters were estimated from the distribution themselves. The yearly maximum rainfall data from all four models, along with the three observational versions, did not reject the null hypothesis. This suggests that they were suitable candidates for the GEV distribution. Among the simulated datasets, the only hourly or daily Markov chain whose maximal annual rainfall dataset did not fit well with the GEV distribution, rejecting the null hypothesis at a 5% confidence level, was the hourly R25 current simulated dataset. It even rejected the null hypothesis at a 1% confidence level, the reason for this remains unclear.

3 Results

3.1 Exploratory analysis

3.1.1 Monthly mean rainfall

Despite Stratton et al. (2018) saying that JJA mean rainfall over West Africa is greater in the CP4 model, this result was not observed in Niamey as Figure 4 shows. Mean monthly

precipitation for both current models is similar, with the R25 model having a lesser maximum value and a more rounded distribution.

The observational data still clearly shows the same overall pattern of rainfall with a singular peak in August. All three observation locations show similar results. Figure 4 (b) shows that there is no significant indication of one model having a lower error than the other. Model CP4 exhibits a higher error for the month of July in all three observation locations than R25, where both models underestimate the mean monthly rainfall. In the case of CP4 this is by around 40mm. Interestingly, for all 3 observation locations and in the case of both models, the peak rainfall month, August, was predicted more accurately than the two months beside it, especially for the Banizoumbou and Niamey Airport locations. The R25 model had its biggest difference from observation to model in the month of September, the beginning of the cessation of the monsoon. This was an underestimation, indicating that the actual annual rainfall distribution has a broader bell curve than both models suggest. As Berthou et al. (2019) suggests, the CP4 model displays more accurate means at the mature stage of the monsoon than in the initiation and cessation months. We can see from the large error bars on the absolute difference charts that we do expect these differences to be subject to quite a large amount of variability. R25 exhibited a larger error than CP4 for the month of May, this was an overestimation, against all three observation locations.

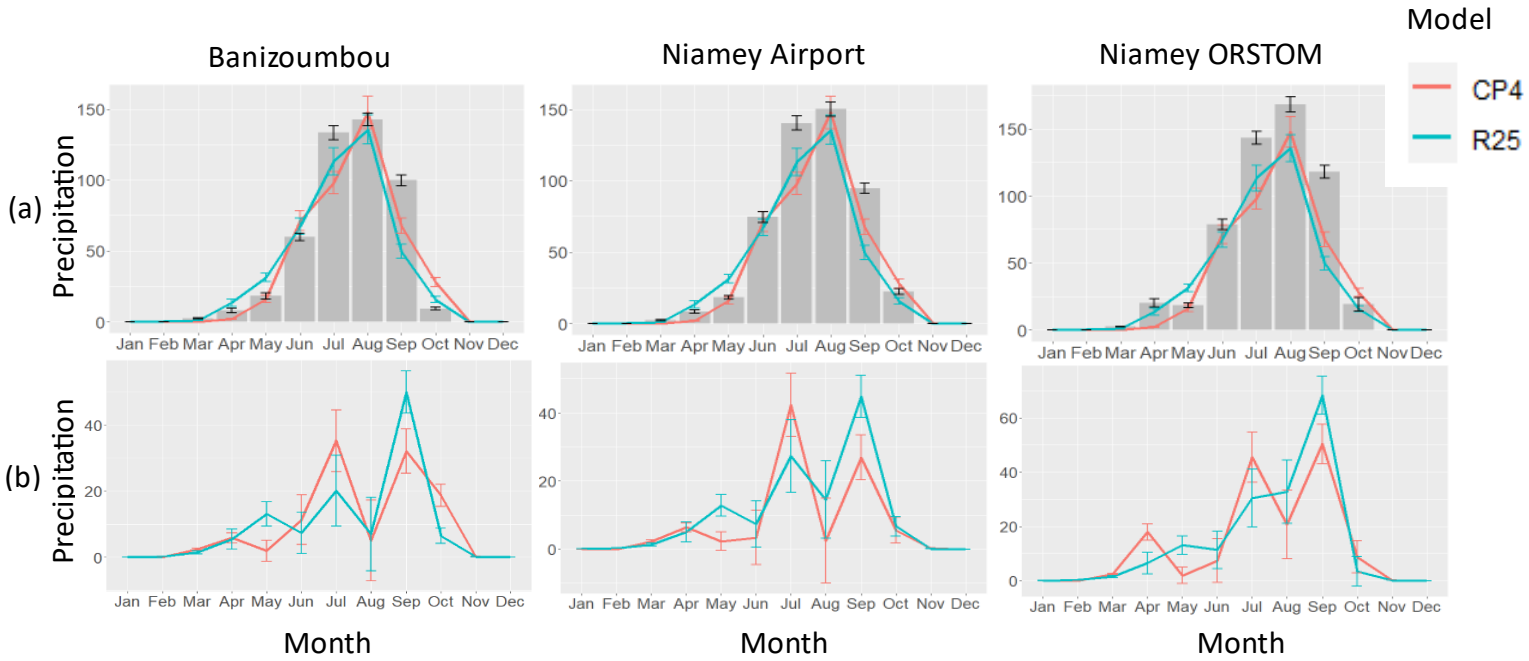

Figure 4: (a) Monthly mean precipitation (mm) for both current model runs over a 10-year period (1997-2006) plotted as lines against observational data from three sites: Banizoumbou (left), Niamey Airport (middle) and Niamey ORSTOM (right) over the same 10-year period (1997-2006) plotted as bars. The monthly mean for observational data is the hourly mean for each month multiplied by 24 to get the daily mean multiplied by 30 to get the monthly mean, this is to match the model runs where there are 30 days in each month and 360 days in a year, missing observation values are ignored in the mean. Error bars here again are standard error. (b) The error between the observational data and both current model runs for mean monthly precipitation (mm) for each observation location. The error bars are the combined standard error from the difference of the distributions calculated by: the square root of the sum of the two standard errors squared.

Figure 4 indicates that there is only 1 wet season in Niamey, Niger as is described in Bigi, Pezzoli and Rosso (2018). This wet season is known as the West African Summer Monsoon. This is when the Intertropical Convergence Zone (ITCZ) is over Niamey. The ITCZ is the ascending branch of the Hadley circulation. The Hadley circulation is an equatorward movement of the trade winds between tropical highs and the equator, caused by the thermally induced, meridional circulation of air rising at the equator and descending at tropical highs (Bjerknes, 1966), because of the change in surface temperature at these locations. The ITCZ is the central point of the cell where the south-east and north-east trade winds meet (Philander, 1983) and ascension occurs. Atmospheric ascension is associated with convection. Monsoon periods are characterised by the prevailing winds changing direction by over 120° and the temperature gradient is reversed. The ITCZ fluctuates in position because of these changing surface temperature gradients.

3.3.2 Proportion of total rainfall from extreme rainfall

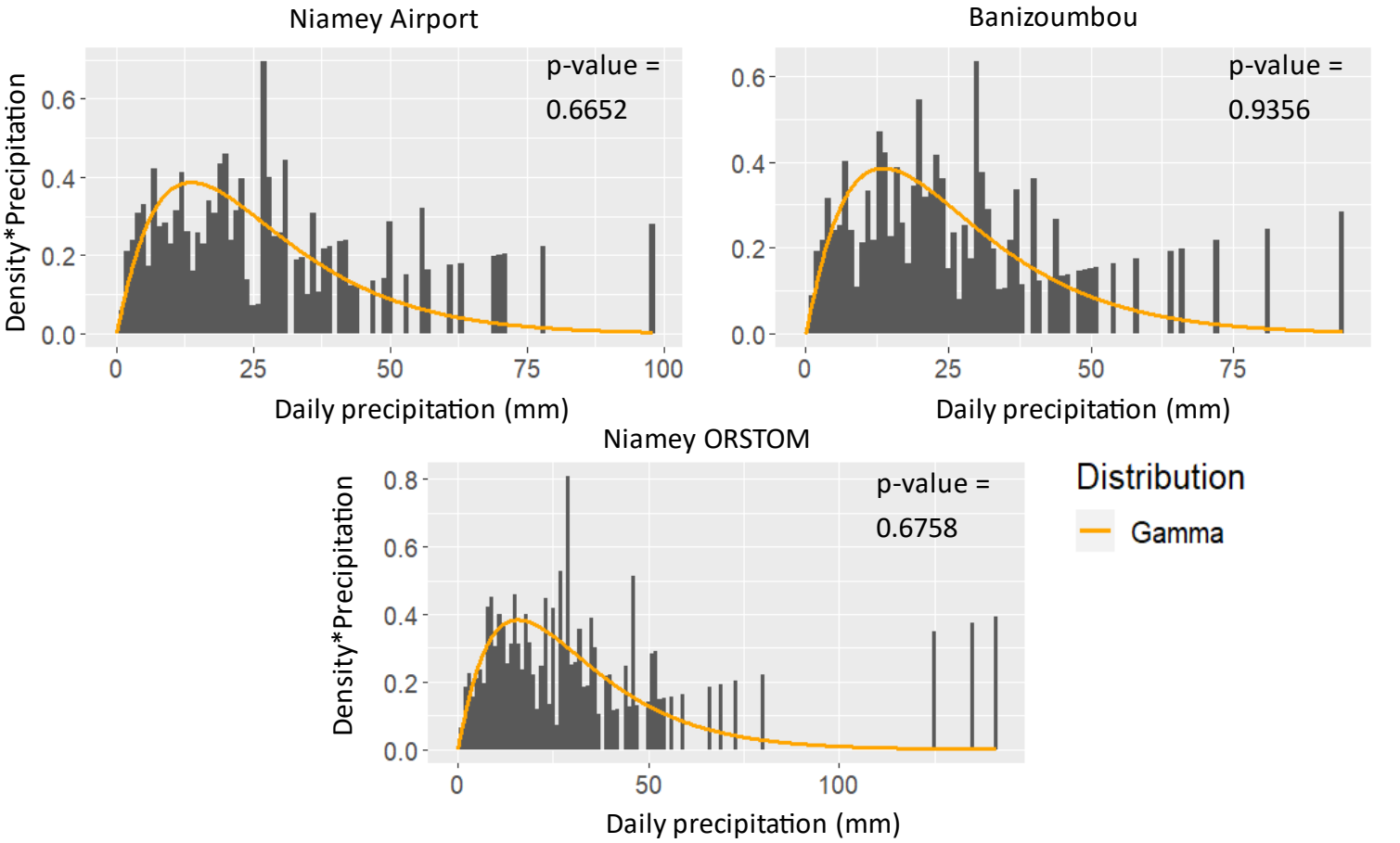

Figure 5: Probability density functions of useful daily precipitation (precipitation values greater than or equal to 1mm) multiplied by the median precipitation value for each bin plotted against precipitation for each of the three observation locations. Plotted as a histogram with bin widths of 1. Superimposed is the same calculation for the gamma distribution fit to the data using the method of maximum likelihood estimators. The p-value of the goodness of fit test of the gamma distribution has been included on each graph. The lower the p-value, the worse fitted to the distribution the data is, this is significant if $p < 0.05$.

The gamma distribution qualitatively matches the shape of the 4 model histograms as well as the three observational ones. From figure 2 it was hypothesized that CP4 future does not suit the gamma distribution very well as was shown, this is confirmed by having a p-value below 0.05, indicating that the gamma distribution fit to the data is not an accurate one. Both future models have a goodness of fit statistic that significantly suggests that the gamma distribution does not fit as well. Therefore exploring other distributions for the future runs would be a beneficial next step as this suggests that useful rainfall amounts are set to change distribution characteristics in a future climate.

The precipitation scales (x axis) show that both current models don't experience any daily rainfall greater than 100mm however the future models have a daily precipitation maximal value of near 150mm. This indicates that extreme rainfall events are becoming more extreme. Comparing the current model scales in Figure 6 with the observations in Figure 5, Niamey Airport and Banizoumbou have similarly maximal values to that of the model but Niamey ORSTOM has a larger maximal value (above 100mm).

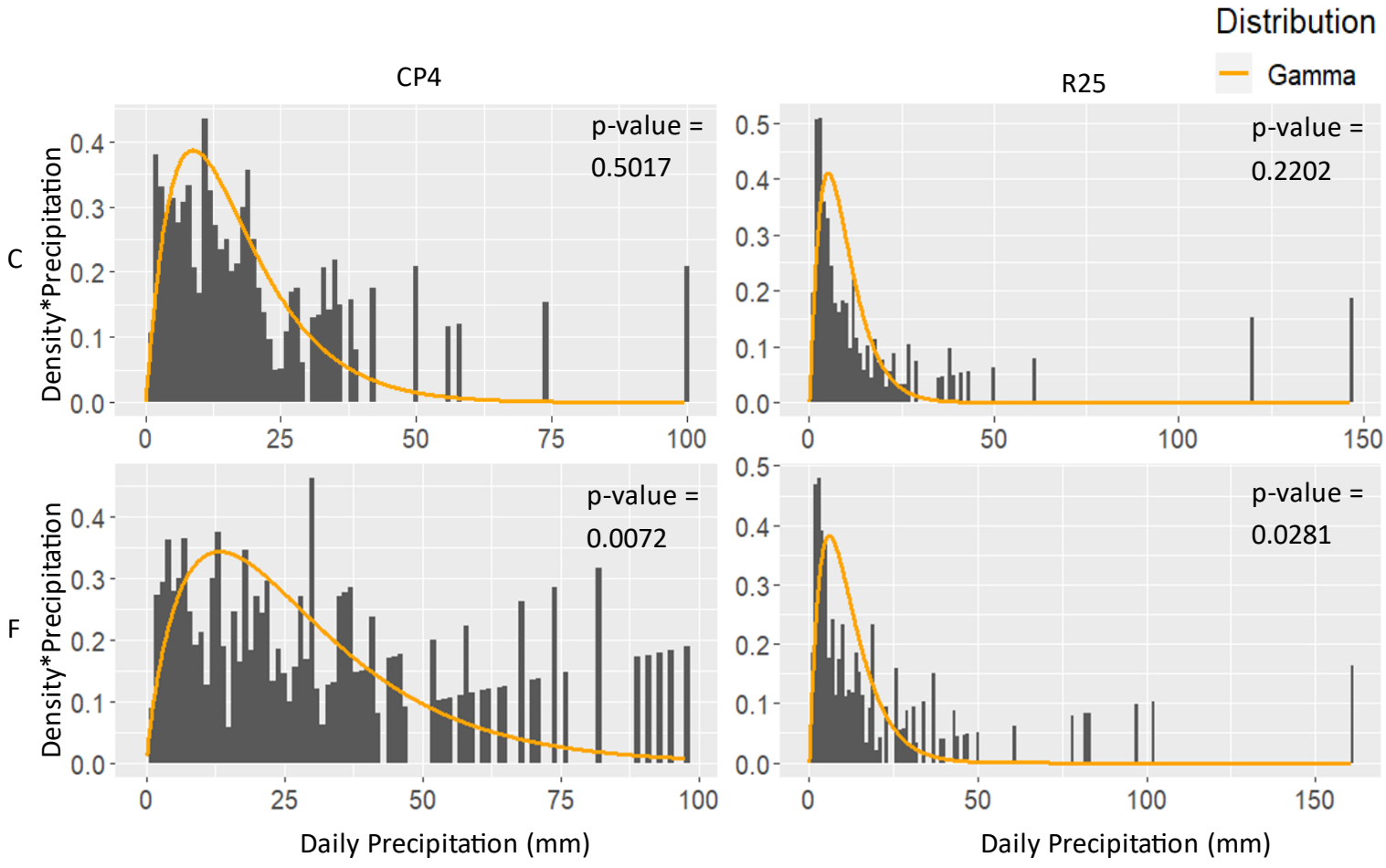

Figure 6: Probability density functions of useful daily precipitation (precipitation values greater than or equal to 1mm) multiplied by the median precipitation value for each bin plotted against precipitation for each of the four models: CP4 (right), R25 (left), current climate (top) and future climate (bottom). Plotted as a histogram with bin widths of 1. Superimposed is the same calculation for the gamma distribution fit to the data using the method of maximum likelihood estimators.. The p-value of the goodness of fit test of the gamma distribution has been included on each graph. The lower the p-value, the worse fitted to the distribution the data is, this is significant if $p < 0.05$.

The area of the histogram / under the distribution indicates the proportion of total rainfall from events with that amount of rainfall. Looking at the difference between the distribution of current models and observations it is clear that a larger proportion of rainfall in the models comes from low daily precipitation, whereas the models seem to get a greater proportion of rainfall from the higher (extreme) precipitation events. This difference indicates that extreme events are more common in the observations. Between the two current models, CP4 demonstrates this property qualitatively better than R25 however the outliers in the extremes in R25 could be misleading for the scale.

The future CP4 model exhibits a greater proportion of rainfall from extreme events than the current model. Indicating that extreme events are more common in a future climate. R25 model qualitatively agrees with this statement.

3.2 Sahelian storm dynamics:

The Sahel region of Africa experiences some of the most intense thunderstorms on earth (Zipser et al., 2006). To check how well our models simulate these mesoscale convective storms, the mean precipitation around rainfall events have been composited with temperature and windspeed to observe the meteorological dynamics of rainfall events in the models and observations.

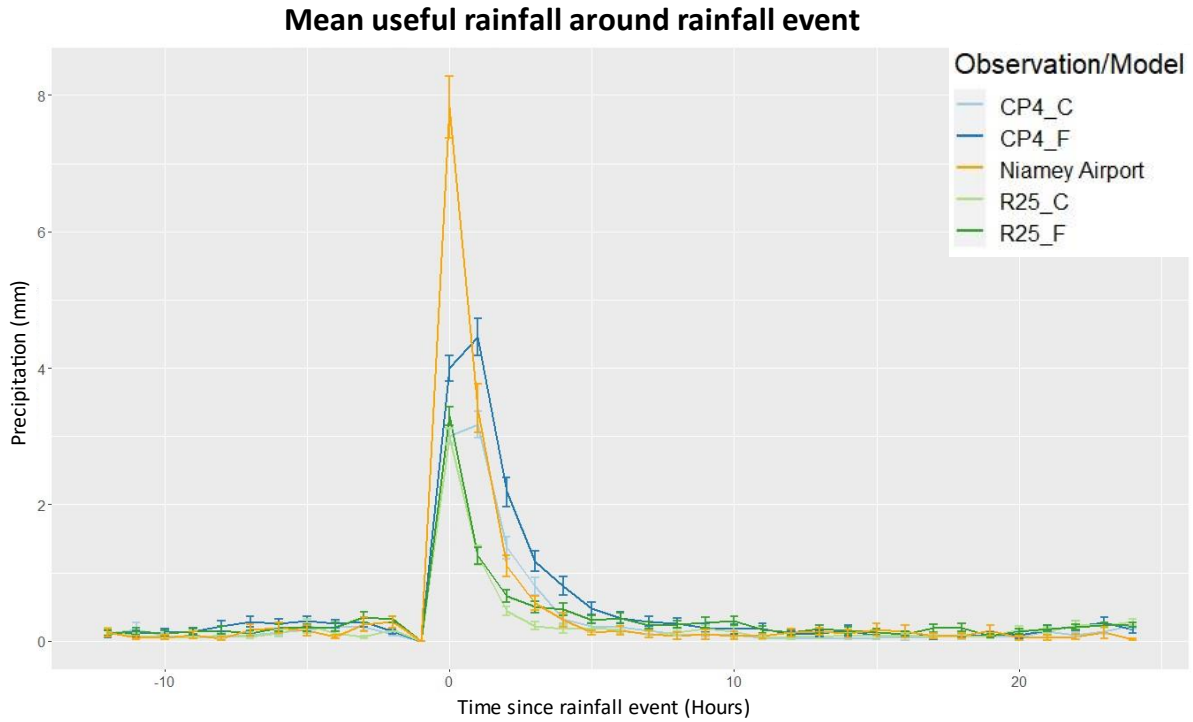

Figure 5: Mean useful rainfall (defined as rainfall above 1mm) around rainfall event (Methods 3.2) for the four model runs and the Niamey Airport observational dataset. Error bars are standard error.

Examining Figure 7, by isolating just the precipitation means of rainfall events, we observe that the observation data shows a very large spike of precipitation at the first hour of a

rainfall event. Both models underestimate that spike however CP4 does have a higher value at time 1, which may be due to 4.5-km simulations having brighter deep convective cores (Stratton et al., 2018). Model CP4 for both runs sees maximal rainfall the hour after the rainfall event begins. This lag time could fail to forecast the initial intensity of rainfall, the intensity of extreme rainfall is very important to the occurrence of flash floods in the Sahel (Elagib et al., 2021).

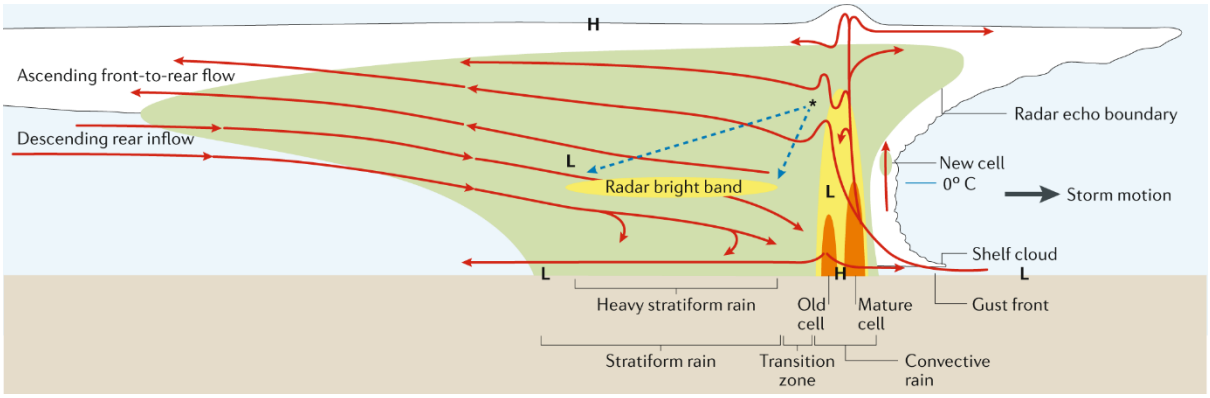

Figure 8: (Schumacher and Rasmussen, 2020) an adaptation of (Houze et al., 1989), showing the organisation of a mesoscale convective storm squall line. The image is a vertical cross section oriented perpendicular to the direction of convection. Pressure minima and maxima are represented by H and L..

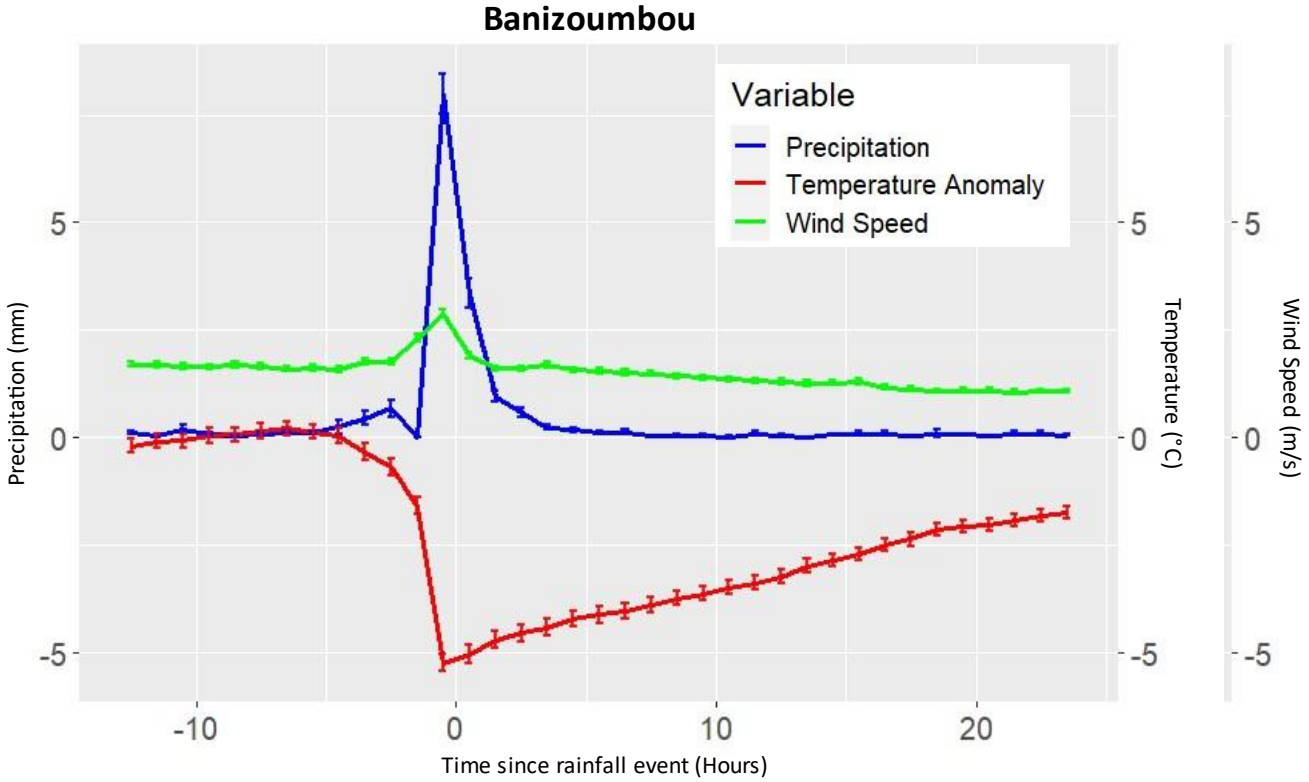

Figure 9: Mean precipitation (blue,) temperature anomaly (red) and wind speed (green) around a rainfall event (rainfall greater than 1mm given that the hour before had no rainfall) for Banizoumbou 2005-2014. All error bars are standard error.

The darker blue and green lines in Figure 7 show that future projections have an increase in the mean rainfall for a fewer hours after the rainfall events compared the current model runs, this difference is more noticeable for the CP4 model and reinforces the claim in Taylor et al. (2017) that higher temperatures intensify storm dynamics. Precipitation returns to the hourly mean level faster in the observational dataset, but all 5 datasets return to this level by around 10 hours after the precipitation event. Suggesting that it is intensity where the models differ from the observations most.

Figure 9 and Figure 10 show rainfall events composited with temperature and wind speed. Figure 9 shows the observational data, the "true" dynamics of a Sahelian storm. We observe a negative temperature anomaly as soon as the rainfall starts, this is as a result of the cold pool at the heart of the MCS that drives the convective processes such as the lifting (Schumacher and Rasmussen, 2020). A small spike in wind speed is shown in Figure 9 as the gust front, as the air moves from high to low pressure. Wind shear is a key ingredient in the formation of a mesoscale convective storm (Schumacher and Rasmussen, 2020). The lag time before precipitation returns to a low level is as a result of the stratiform rain in figure 8 lingering behind the storm front known as "precipitation drag" (Houze et al., 1989).

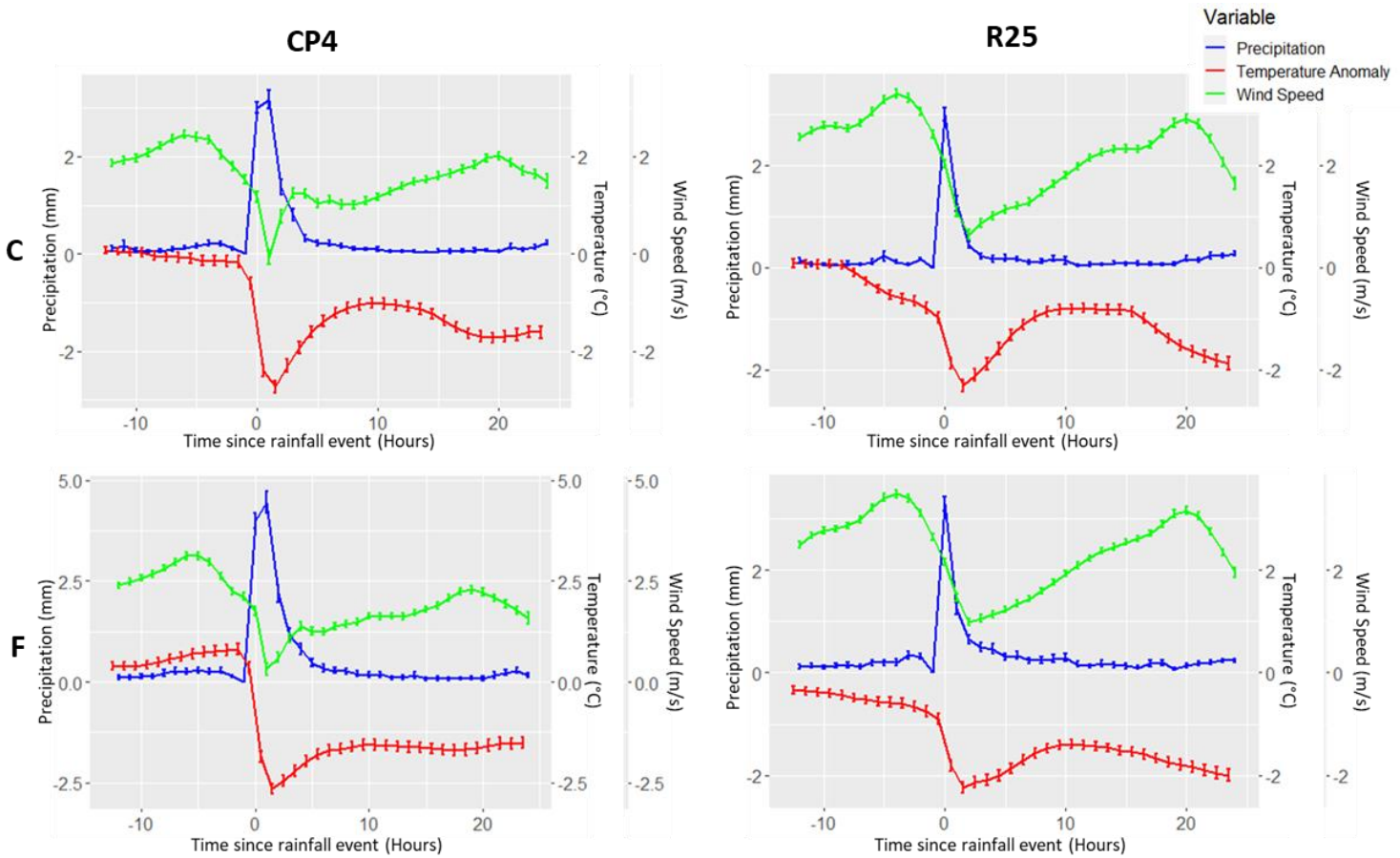

Figure 6: Mean precipitation (blue,) temperature anomaly (red) and wind speed (green) around a rainfall event (rainfall greater than 1mm given that the hour before had no rainfall) for the four model runs: CP4 (right), R25 (left), current climate (top) and future climate (bottom). All error bars are standard error.

Now comparing the observational composites to the two current model composites in Figure 10, the most obvious result is the opposite relationship modelled MCSs have with wind speed. It indicates a decrease in wind speed for all 4 model runs when a rainfall event occurs. This questions the accuracy of the model dynamics which consequently could mean that the models wrongly forecast the way MCSs evolve in the future. This is dangerous as severe wind gusts are a possible deadly impact of MCSs (Schumacher and Rasmussen, 2020). Moreover, when comparing both current models with the observations, R25 indicates a temperature anomaly emerging approximately 8 hours prior to the rainfall event, while CP4 demonstrates a closer alignment with the observational data, indicating that the temperature anomaly begins only about an hour before the rainfall event. One explanation for the incorrect prediction by the R25 model could be that convection permitting models are more accurate at predicting the diurnal cycle of rainfall (Fitzpatrick et al., 2020b). The northern Sahel experiences a nocturnal daily precipitation maximum during monsoon periods (Vizy and Cook, 2017). The misjudgement by R25 could directly affect temperature anomalies due to the region's strong diurnal temperature cycle.

3.3 Intensity-duration analysis of wet spells

3.3.1 Peak rainfall

To further analyse flood risk, the duration of rainfall events along with the intensity are analysed as Elagib et al. (2021) suggests the main causes of floods in the region are extreme rainfall, consecutive daily rainfall and high river water. In fact, intensity and duration are key factors in flooding everywhere (Kendon et al., 2014).

Comparing the two current models to the observations in Figure 11, R25 has wet spell durations over double the length of the observations, which is an expected symptom of displaying the Markov property (methods 2.3). CP4 only has one duration above the observational maximum of 5 days indicating it doesn't exhibit the same precipitation bias as R25, as was also found by Stratton et al. (2018). Due to the log scale and different number of bins in each it is hard to compare exact distributions however it is clear that the modal peak rainfall of 1 day wet spells in the observational datasets was above 10. In both current model runs this metric was below ten, indicating that both models don't accurately predict the intensity of peak daily rainfall in a wet spell. However the models do accurately follow the shape of the observation PDFs, this shape shows that, given that the wet spell has a longer duration, it is less likely to have a low peak rainfall. Therefore these longer duration events (above 3 days) are more likely to cause flooding as they have a consistently high (above 10mm) peak rainfall value and a longer duration time by definition.

Comparing the current models to future projections in Figure 11 it is clear that the R25 future run estimates longer wet spells than the current run, but the previous comparisons show that this model does not accurately represent wet spell intensities and durations. The future convection permitting model looks very similar to the current model, with no obvious

increase in wet spell length or duration. Klutse et al. (2018) suggests that wet spell lengths in Niamey will not increase with temperature, therefore agreeing with CP4. It does qualitatively appear that the extreme peak rainfalls seem more common in a future climate, which is confirmed in the rest of the paper.

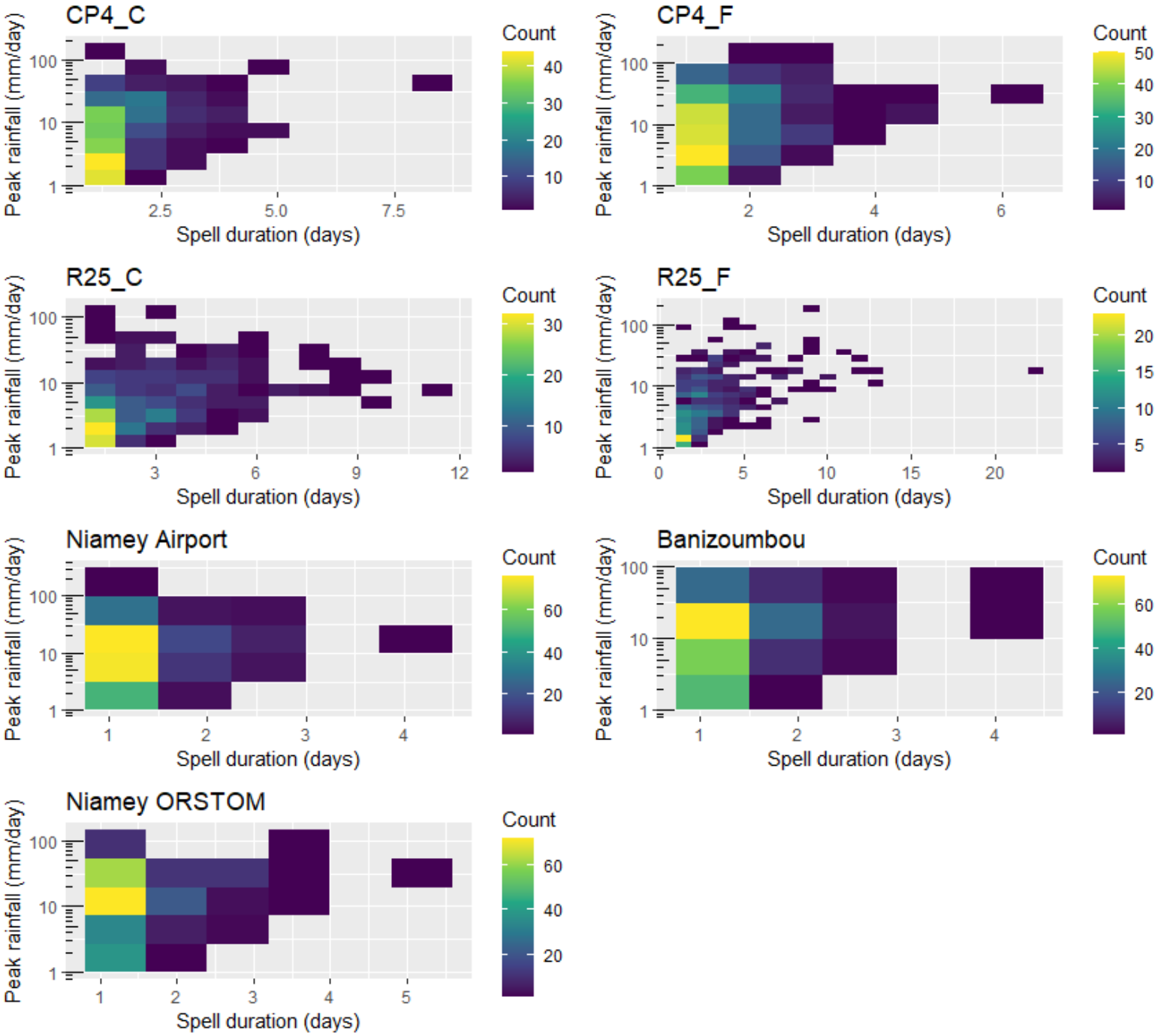

Figure 7: Intensity duration plots of wet spells for all four model datasets and the three observational datasets. Each plot has the peak daily rainfall of the wet spell plotted on a log 10 axis against the duration of the wet spell. The maximum wet spell length for each dataset is L and the number of bins is L × L+1.

3.3.2 Cumulative rainfall

Cumulative rainfall over consecutive days (a wet spell) in the Sahel has been shown to cause flooding (Taylor et al., 2017). Table 3 shows the number of wet events at least 3 days in

duration and the average cumulative rainfall of those events. Both current models over estimate the number of wet spells of duration 3 days or more, CP4 overestimates this metric by roughly double and the parameterisation model by between 5 and 12 times. There is significant variability in the lengths of wet spells between the models so averaging spatially could be used to minimise this variability.

Table 3: The data being observed in this table is with duration greater than or equal to three days. Cumulative rainfall is the sum of all rainfall over each wet spell, the mean and standard error of the 10 year long datasets are included.

	Count	**Mean Cumulative Rainfall**	**Mean Cumulative Rainfall SE**
CP4_C	36	29.96	3.725
CP4_F	35	43.50	7.305
R25_C	122	22.58	1.806
R25_F	139	32.14	2.968
Niamey Airport	15	33.58	5.873
Banizoumbou	10	40.62	7.896
Niamey ORSTOM	20	40.36	4.177

Despite Table 3 indicating that CP4 underestimates the mean cumulative rainfall, Table 4 shows that the CP4 does not display significant differences to 2/3 of the observational datasets, and the only one it does do is only at a 10% confidence level. This may be due to it not being different however it could be due to a small sample size. The parameterisation model is significantly different from all observations at a 10% confidence level and is significantly different from one of them at a 5% confidence level. Showing again that it significantly underestimates the intensity of precipitation.

For CP4, Table 3 qualitatively shows that the mean cumulative rainfall of wet spells is increasing. As it is shown in Figure 11 that the length of wet spells is not increasing, it has to be that the daily precipitation amounts are increasing, indicating more extreme rainfall. Table 4 shows that this is not a significant increase but it very nearly is at a 10% threshold, again possibly due to sample size. Finally the number of wet spells from current to future in CP4 stays roughly the same, indicating that minimally 3 day wet spells don't change in frequency in the future despite Fitzpatrick et al. (2020a) indicating that wet events will become less frequent in a future climate. For R25 the future model does show a significant increase in mean cumulated rainfall, but due to this models lack of accuracy in modelling wet spells this is not a strong result. Increased cumulative rainfall is a direct flood risk as higher rainfall amounts over multiple days means that more rainfall is falling on saturated ground increasing the surface run-off.

Table 4: Results from Welch two sample t-tests comparing the mean cumulative rainfall of wet spells (longer than or equal to 3 days) for both current models to that of the observations and the corresponding future model version. Rejecting the null hypothesis confirms that there is a significant (at the 5% or 10% threshold) difference between the mean cumulative rainfalls.

		Niamey Airport	Banizoumbou	Niamey ORSTOM	Future
CP4 Current					
	P-Value	0.6079	0.2434	0.0696	0.1050
	Reject at 10% confidence	No	No	Yes	No
	Reject at 5% confidence	No	No	No	No
R25 Current					
	P-Value	0.0916	0.0501	0.0006	0.0064
	Reject at 10% confidence	Yes	Yes	Yes	Yes
	Reject at 5% confidence	No	No	Yes	Yes

3.4 Extreme value analysis

To evaluate the previously stated hypothesis that extreme rainfall events in the Sahel are becoming more extreme, as is also suggested in Biasutti (2019), an extreme value distribution of daily precipitation maxima is fit for each simulated dataset. From this, return periods of rainfall events can be computed (methods 3.5), as shown in Figure 12 (a). Firstly, comparing the model Markov chains to the observational Markov chain (Niamey Airport), both current models underestimate the observational predicted extremes for all return periods. This is true for both the daily and hourly Markov chains, but more clearly true for the hourly Markov chains (right) where the observational dataset predicts higher extreme values of daily rainfall than all of the model runs, even future versions. Therefore this brings into question the accuracy of the values for each return period predicted by the models.. The current convection permitting model appears a better comparator in both cases but this cannot be deduced with certainty. However we can still analyse the differences between future and current climate. Figure 12 (b) shows both models see future increases for both methods of simulating rainfall data for all precipitation return levels from 2 to 100 years. However the convection permitting model shows a greater increase for both methods. Simulating data using a daily Markov chain exhibits higher percentage increases from current to future climates (Figure 12 (b)) but lower actual precipitation daily return levels than using an hourly Markov chain. The Clausius-Clapeyron relationship indicates that the atmospheres water-holding capacity should increase by between 24% and 28% for a 4°C temperature increase (6-7% increase per °C (Taylor et al., 2017)), therefore extreme rainfall should increase by a similar scaling and both hourly Markov chains seem to be closer to this scaling from current to future climates.

Comparing the return levels to the yearly maxima in the actual datasets, the hourly Markov chain appears more accurate as the daily Markov chain seems to underestimate the extremes. The shape of the CP4 graphs in Figure 12 (b) show that the higher the return period (less likely an event is to occur) the greater the percentage increase of extreme precipitation. This qualitatively indicates the relationship between increased sea surface temperatures and extreme rainfall return periods is non-linear for the convection permitting model simulated datasets and linear for the parameterisation model simulated dataset.

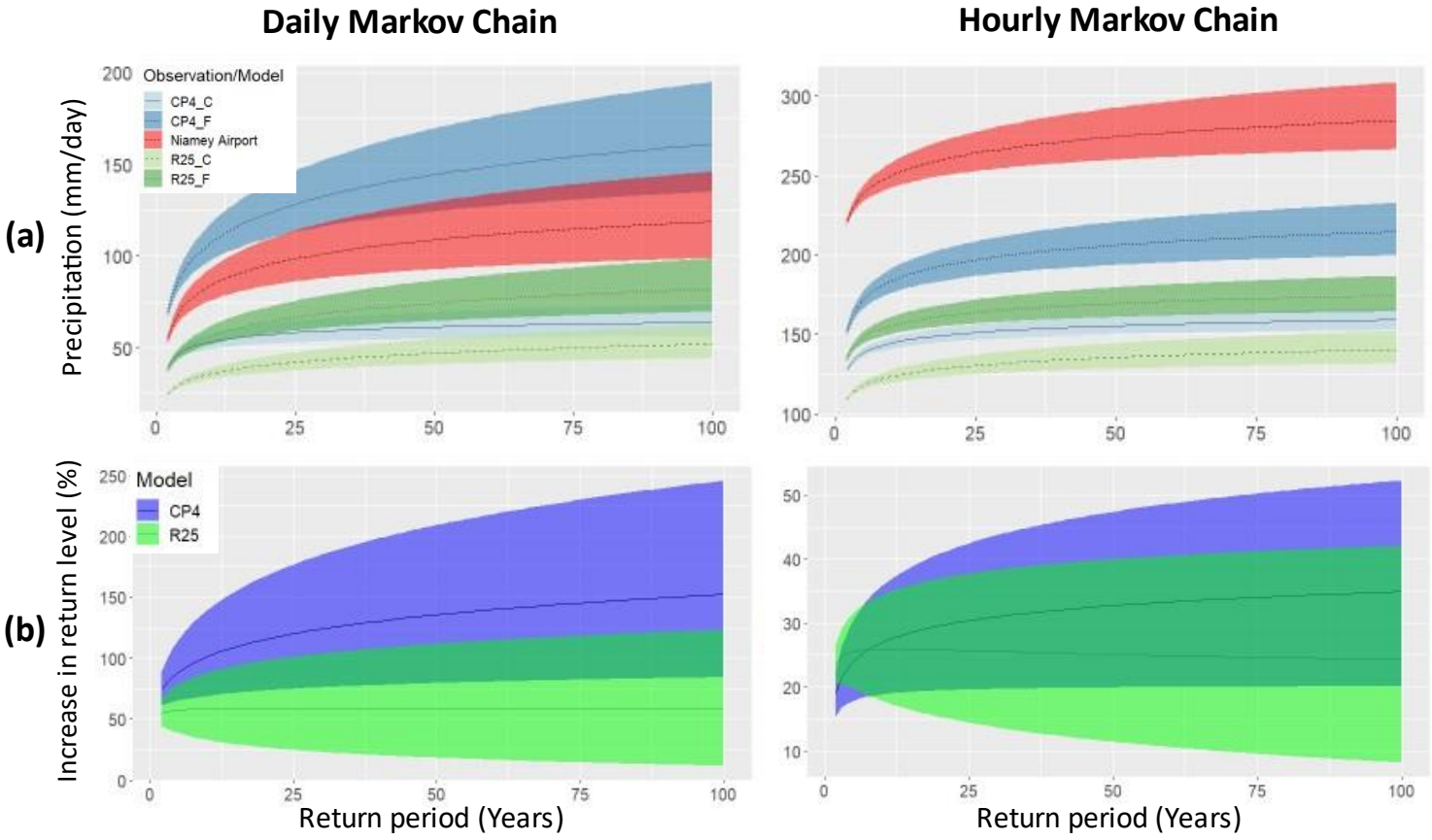

Figure 82: Data: 100 year synthesized Markov chains for each model run (CP4 future and current, R25 future and current) and Niamey (methods 2.3), a daily Markov chain (left) and hourly Markov chain in the wet season only (right). Yearly maximal values for each dataset fit to the general extreme value distribution. Return values of extreme rainfall computed from the cumulative distribution function for the return periods of 2-100 years (methods 2.5).(a) Return period in years of extreme daily precipitation. (b) Percentage increase in extreme daily precipitation between current and future climates for each return level.

4 Discussion

Despite Berthou et al. (2019) saying that generally parameterise models accurately predict the characteristics of a Sahelian storm, we have shown that wind speed is not accurately predicted in either model in Niamey. The increase in frequency of early-evening storms in the Sahel in recent times (Taylor et al., 2017) is as a result of increased midday wind shear (Fitzpatrick et al., 2020a) and displaying the wrong relationship between precipitation events and wind speed could wrongly model future changes in MCSs. The CP4 model performs better at modelling the temperature change around a rainfall event, and it has been shown

that it is accurate at modelling the duration of wet spells in Niamey. Despite, both models underestimating the intensity of extreme rainfall events and general wet spells, CP4 is better at modelling these extremes as was shown in Prein et al. (2015). Berthou et al. (2019) also showed this underestimation for Niamey, but that CP4 generally overestimates rainfall in the Sahel region. Our findings show that on a local scale, an underestimation can strongly affect the extreme rainfall predictions, therefore the models can not be used with as much confidence when evaluating precipitation risks on a single locality.

Both of our models show that the sea surface temperature increase programmed into the future models will cause more extreme rainfall events in the future climate for Niamey, matching the analysis on a regional scale done by Kendon et al. (2019). One reason for this increase could be the Clausius Clapeyron relationship which indicates that rising temperatures should increase the amount of water-vapor holding potential of clouds, which in turn will increase the total column water of MCSs (Fitzpatrick et al., 2020a; Kendon et al., 2019). Warming also causes an increase in convective available potential energy (CAPE) (Muller and Takayabu, 2020) which is shown to scale roughly at the same rate as the Clausius Clapeyron relationship. Furthermore, the intensification of MCSs could stem from amplified warming in the Sahara compared to the surrounding oceans, as predicted to be a result of greenhouse warming (Haarsma et al., 2005). This amplified warming would lead to an increase in the north-south temperature gradient, resulting in enhanced wind shear over the Sahel (Fitzpatrick et al., 2020a), a crucial element in the formation of MCSs.

The findings predict that extreme events are set to become more intense in a future climate. As previously discussed, the intensity of extreme rainfall events is a key element in the occurrence of floods (Kendon et al., 2014). An increase in extreme rainfall intensity was responsible for the extensive urban flooding in the Sahel in 2017 (Adegoke et al., 2019). Furthermore the findings of this paper show that the cumulative rainfall of wet events over 3 days is also set to increase, another cause of floods (Taylor et al., 2017). Another key factor in flood risk is the ground cover (Fiorillo et al., 2018). From 1988 to 2030 the built up areas in the city are expected to increase by 130% (Manu et al., 2014), which would mean more soil cover and more surface runoff water, resulting in a further increased flood risk.

Africa as a continent has an absence of learning tools that encourage adaptive processes (Tschakert and Dietrich, 2010). In Niamey, the state run flood response is done by three organisations: the Regional Device for Food Crisis Prevention and Management (DRPGCA); the General Directorate of Civil Protection (DGPC); and the HFA/GC (Mahamane, Oumarou and Piñeira Mantiñán, 2023). However Mahamane, Oumarou and Piñeira Mantiñán (2023) suggests that the response that they provide is insufficient because of a lack of human resource, insufficient flood planning and the inhabitants of Niamey not cooperating. Because of this, NGOs have had to try to fill the gap in support to those affected (Mahamane, Oumarou and Piñeira Mantiñán, 2023). Despite the city declaring some areas of Niamey undevelopable because they are at a high risk of floods, those areas continue to be

urbanised (Mahamane, Oumarou and Piñeira Mantiñán, 2023) because of the demand for housing caused by the massive population growth (Manu et al., 2014). This puts at risk the health of the most vulnerable, for whom feel that this is the only place they can live, along with the infrastructure that they build.

When examining extreme rainfall events, it is important to consider the global scale drivers as well. Ficchì and Stephens (2019) discovered that the timing of floods in Africa is sensitive to the Indian Ocean Dipole (IOD) and the El Nino Southern Oscillation (ENSO). ENSO and IOD are climate modes driven by anomalous sea surface temperatures which trigger meteorological responses, such as increases or decreases in rainfall. Ficchì and Stephens (2019) also suggested that other Atlantic climate modes could also have a strong impact on West African flood timing. Being able to predict flood timing can increase the accuracy of early warning systems in order to protect buildings, agriculture and human life, so this should be taken into account if building a flood detection service.

One way to extend upon the extreme value analysis would be to repeat it for multiple locations in the Sahel region see whether this trend of extreme rainfall increasing in intensity occurs everywhere, or if there is any spatial variation. Chapman et al. (2022) performed an analysis with more locations in Eastern Africa for example. Alternatively, because of the distinct interdecadal cycles in rainfall discussed in Bigi, Pezzoli and Rosso (2018), it may be interesting to provide model runs that simulate each of these cycles.

5 Conclusion

The intensity of rainfall events in Niamey has been increasing over the last few years and is set to increase further in the future climate projection. The convection permitting model for Africa is more accurate at modelling convective processes than the previous parametrisation model, but it still underestimates the intensity of extreme and general rainfall events in Niamey. This is a risk if using it for flood prediction purposes as the intensity of peak rainfall is an important factor in the occurrence and impacts of floods. Wet spell lengths in convection permitting models are more similar to observational data as parameterisation models exhibit a precipitation bias in the wet season. Rainfall events in the wet season are dependent on what has happened in the previous hours but not the previous days. Therefore, convection permitting models are more appropriate for estimating the increase in frequency and intensity of floods in Niamey. The length of wet spells does not appear to increase in a future climate, but the intensity of rainfall does, making extreme events more extreme and increasing the cumulative rainfall of wet events. We cannot confidently say how intense these extreme events are likely to be in the future but, with some certainty, we can say that they will become more extreme than they are now. Quantifying this increase should be a key future area of research. Consequently, the effectiveness of flood protection processes in Niamey needs to be evaluated. More extreme future rainfall events along with increased urbanisation are likely to cause more frequent and more intense flooding, if changes are not made to mitigate this.

Ethical Considerations

One of the key ethical considerations in this project is ensuring the quality of the analysis being performed. Mistakes in data cleaning and manipulation can lead to spurious results and the wrong recommendations. Not understanding and being transparent about the mathematical steps behind each analysis technique could also deceive the reader. This includes showing the results that don't agree with assumed hypotheses. Throughout the data analysis stringent quality checks were implemented. Furthermore, the statistical significance of the results was checked throughout.

The ethical motivation for undertaking this project is clear. Identifying trends in future rainfall by first validating the accuracies of the models and then comparing the models to their future runs. As this model is state of the art, this analysis could provide vital information about future flood risk to save human lives, protect livelihoods and increase global awareness.

Bibliography

Adegoke, J., Sylla, M.B., Taylor, C., Klein, C., Bossa, A., Ogunjobi, K. and Adounkpe, J. 2019. On the 2017 rainy season intensity and subsequent flood events over West Africa J. Adegoke, M. B. Sylla, A. Y. Bossa, K. Ogunjobi, & J. Adounkpe, eds. *nora.nerc.ac.uk*. [Online], pp.10–14. [Accessed 27 March 2024]. Available from: https://nora.nerc.ac.uk/id/eprint/524214/.

AMMA-CATCH. 1990. AMMA-CATCH : a hydrological, meteorological and ecological observatory on West Africa. IRD, CNRS-INSU, OSUG, OMP, OREME. doi:10.17178/AMMA-CATCH.all

Berthou, S., Rowell, D.P., Kendon, E.J., Roberts, M., Stratton, R.A., Crook, J. and Wilcox, C. 2019. Improved climatological precipitation characteristics over West Africa at convection-permitting scales. *Climate Dynamics*. **53**(3-4), pp.1991–2011.

Biasutti, M. 2019. Rainfall trends in the African Sahel: Characteristics, processes, and causes. *Wiley Interdisciplinary Reviews: Climate Change*. **10**(4), p.e591.

Bigi, V., Pezzoli, A. and Rosso, M. 2018. Past and Future Precipitation Trend Analysis for the City of Niamey (Niger): An Overview. *Climate*. **6**(3), p.73.

Bjerknes, J. 1966. A possible response of the atmospheric Hadley circulation to equatorial anomalies of ocean temperature. *Tellus*. **18**(4), pp.820–829.

Chan, S.C., Kendon, E.J., Roberts, N.M., Fowler, H.J. and Blenkinsop, S. 2016. Downturn in scaling of UK extreme rainfall with temperature for future hottest days. *Nature Geoscience*. **9**(1), pp.24–28.

Chapman, S., Bacon, J., Birch, C.E., Pope, E., Marsham, J.H., Msemo, H., Nkonde, E., Sinachikupo, K. and Vanya, C. 2022. Climate change impacts on extreme rainfall in Eastern Africa in a convection permitting climate model. *Journal of Climate*. **36**(1), pp.1–39.

Cullen, A.C. and Frey, C.H. 1999. *Probabilistic Techniques in Exposure Assessment*. Springer Science & Business.

D'Amato, N. and Lebel, T. 1998. On the characteristics of the rainfall events in the Sahel with a view to the analysis of climatic variability. *International Journal of Climatology*. **18**(9), pp.955–974.

Delignette-Muller, M.-L. 2023. Help to Fit of a Parametric Distribution to Non-Censored or Censored Data. *https://cran.r-project.org/web/packages/fitdistrplus/fitdistrplus.pdf*. [Online]. [Accessed 27 March 2024]. Available from: https://cran.r-project.org/web/packages/fitdistrplus/fitdistrplus.pdf.

Elagib, N.A., Zayed, I.S.A., Saad, S.A.Gayoum., Mahmood, M.I., Basheer, M. and Fink, A.H. 2021. Debilitating floods in the Sahel are becoming frequent. *Journal of Hydrology*. **599**, p.126362.

Faraway, J., Marsaglia, G., Marsaglia, J. and Baddeley, A. 2021. Classical Goodness-of-Fit Tests for Univariate Distributions. *search.r-project.org*. [Online]. [Accessed 25 March 2024]. Available from: https://search.r-project.org/CRAN/refmans/goftest/html/00Index.html.

Ficchì, A. and Stephens, L. 2019. Climate Variability Alters Flood Timing Across Africa. *Geophysical Research Letters*. **46**(15), pp.8809–8819.

Fiorillo, E., Crisci, A., Issa, H., Maracchi, G., Morabito, M. and Tarchiani, V. 2018. Recent Changes of Floods and Related Impacts in Niger Based on the ANADIA Niger Flood Database. *Climate*. **6**(3), p.59.

Fitzpatrick, R.G.J., Parker, D.J., Marsham, J.H., Rowell, D.P., Guichard, F.M., Taylor, C.M., Cook, K.H., Vizy, E.K., Jackson, L.S., Finney, D., Crook, J., Stratton, R. and Tucker, S. 2020. What Drives the Intensification of Mesoscale Convective Systems over the West African Sahel under Climate Change? *Journal of Climate*. **33**(8), pp.3151–3172.

Fitzpatrick, R.G.J., Parker, D.J., Marsham, J.H., Rowell, D.P., Jackson, L.S., Finney, D., Deva, C., Tucker, S. and Stratton, R. 2020. How a typical West African day in the future-climate compares with current-climate conditions in a convection-permitting and parameterised convection climate model. *Climatic Change*. **163**(1), pp.267–296.

Google Maps. 2024. *Niamey, Niger 1:100*. Available from: https://www.google.com/maps/place/Niamey,+Niger/@13.525199,2.1434709,10.45z/data=!4m6!3m5!1s0x11d0756cc0ddfc65:0x81ce4bafda77b74e!8m2!3d13.5115963!4d2.1253854!16zL20vMGZxZnM?entry=ttu [Accessed 15 March 2024].

Haarsma, R.J., Selten, F.M., Weber, S.L. and Kliphuis, M. 2005. Sahel rainfall variability and response to greenhouse warming. *Geophysical Research Letters*. **32**(17).

Houze, R.A., Rutledge, S.A., Biggerstaff, M.I. and Smull, B.F. 1989. Interpretation of Doppler Weather Radar Displays of Midlatitude Mesoscale Convective Systems. *Bulletin of the American Meteorological Society*. **70**(6), pp.608–619.

Jones, L., Dougill, A., Jones, R.G., Steynor, A., Watkiss, P., Kane, C., Koelle, B., Moufouma-Okia, W., Padgham, J., Ranger, N., Roux, J.-P., Suarez, P., Tanner, T. and Vincent, K. 2015. Ensuring climate information guides long-term development. *Nature Climate Change*. **5**(9), pp.812–814.

Kendon, E.J., Roberts, N.M., Fowler, H.J., Roberts, M.J., Chan, S.C. and Senior, C.A. 2014. Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nature Climate Change*. **4**(7), pp.570–576.

Kendon, E.J., Stratton, R.A., Tucker, S., Marsham, J.H., Berthou, S., Rowell, D.P. and Senior, C.A. 2019. Enhanced future changes in wet and dry extremes over Africa at convection-permitting scale. *Nature Communications*. **10**(1), pp.1–14.

Klutse, N.A.B., Ajayi, V.O., Gbobaniyi, E.O., Egbebiyi, T.S., Kouadio, K., Nkrumah, F., Quagraine, K.A., Olusegun, C., Diasso, U., Abiodun, B.J., Lawal, K., Nikulin, G., Lennard, C. and Dosio, A. 2018. Potential impact of 1.5 °C and 2 °C global warming on consecutive dry and wet days over West Africa. *Environmental Research Letters*. **13**(5), p.055013.

Mahamane, S.O., Oumarou, A. and Piñeira Mantiñán, M.J. 2023. Improving Public Action to Mitigate River Flooding in Niamey (Niger). *Land*. **12**(8), p.1523.

Manu, A., Twumasi, Y.A., Lu Kang, S. and Coleman, T.L. 2014. Predicting urban growth of a developing country city using a statistical modeling approach. *International Journal of Geomatics and Geosciences*. **5**(4), pp.603–613.

Millard, S.P. and Kowarik, A. 2020. *Package for Environmental Statistics, Including US EPA Guidance* [Online]. Available from: https://cran.r-project.org/web/packages/EnvStats/EnvStats.pdf.

Mouhamed, L., Traore, S.B., Alhassane, A. and Sarr, B. 2013. Evolution of some observed climate extremes in the West African Sahel. *Weather and Climate Extremes*. **1**, pp.19–25.

Muller, C. and Takayabu, Y. 2020. Response of precipitation extremes to warming: what have we learned from theory and idealized cloud-resolving simulations, and what remains to be learned? *Environmental Research Letters*. **15**(3), p.035001.

Özge Karadağ and Serpil Aktaş 2016. Goodness of fit tests for generalized gamma distribution. *AIP Conference Proceedings*. **1738**(1).

Panthou, G., Vischel, T. and Lebel, T. 2014. Recent trends in the regime of extreme rainfall in the Central Sahel. *International Journal of Climatology*. **34**(15), pp.3998–4006.

Philander, S.G.H. 1983. El Niño Southern Oscillation phenomena. *Nature*. **302**(5906), pp.295–301.

Pouillot, R. and Delignette-Muller, M.L. 2010. Evaluating variability and uncertainty separately in microbial quantitative risk assessment using two R packages. *International Journal of Food Microbiology*. **142**(3), pp.330–340.

Prein, A.F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., Feser, F., Brisson, E., Kollet, S., Schmidli, J., Lipzig, N.P.M. and Leung, R. 2015. A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Reviews of Geophysics*. **53**(2), pp.323–361.

Rebora, N., Ferraris, L., von Hardenberg, J. and Provenzale, A. 2006. RainFARM: Rainfall Downscaling by a Filtered Autoregressive Model. *Journal of Hydrometeorology*. **7**(4), pp.724–738.

Reynolds, R.W., Smith, T.M., Liu, C., Chelton, D.B., Casey, K.S. and Schlax, M.G. 2007. Daily High-Resolution-Blended Analyses for Sea Surface Temperature. *Journal of Climate*. **20**(22), pp.5473–5496.

Schumacher, R.S. and Rasmussen, K.L. 2020. The formation, character and changing nature of mesoscale convective systems. *Nature Reviews Earth & Environment*. **1**(6), pp.300–314.

Stephenson, A. and Ferro, C. 2018. *Functions for Extreme Value Distributions* [Online]. Available from: https://cran.r-project.org/web/packages/evd/evd.pdf.

Stratton, R.A., Senior, C.A., Vosper, S.B., Folwell, S.S., Boutle, I.A., Earnshaw, P.D., Kendon, E., Lock, A.P., Malcolm, A., Manners, J., Morcrette, C.J., Short, C., Stirling, A.J., Taylor, C.M., Tucker, S., Webster, S. and Wilkinson, J.M. 2018. A Pan-African Convection-Permitting Regional Climate Simulation with the Met Office Unified Model: CP4-Africa. *Journal of Climate*. **31**(9), pp.3485–3508.

Taylor, C.M., Belušić, D., Guichard, F., Parker, D.J., Vischel, T., Bock, O., Harris, P.P., Janicot, S., Klein, C. and Panthou, G. 2017. Frequency of extreme Sahelian storms tripled since 1982 in satellite observations. *Nature*. **544**(7651), pp.475–478.

Tschakert, P. and Dietrich, K.A. 2010. Anticipatory Learning for Climate Change Adaptation and Resilience. *Ecology and Society*. **15**(2).

Vizy, E.K. and Cook, K.H. 2017. Mesoscale convective systems and nocturnal rainfall over the West African Sahel: role of the Inter-tropical front. *Climate Dynamics*. **50**(1-2), pp.587–614.

Wilks, D.S. 1999. Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agricultural and Forest Meteorology*. **93**(3), pp.153–169.

Wilks, D.S. 2010. *Statistical methods in the atmospheric sciences*. Amsterdam: Elsevier.

Zipser, E.J., Cecil, D.J., Liu, C., Nesbitt, S.W. and Yorty, D.P. 2006. WHERE ARE THE MOST INTENSE THUNDERSTORMS ON EARTH? *Bulletin of the American Meteorological Society*. **87**(8), pp.1057–1072.