CompAMa
www.compama.co.usb.ve

Bull. Comput. Appl. Math., Vol.1, No.2, 2013
ISSN 2244-8659

Oceanic influence on extreme rainfall trends in the north central coast of Venezuela: present and future climate assessments

L. Guenni [*], *C. Nobre* [†], *J. Marengo* [‡], *G. Huerta* [§], *B. Sansó* [¶]

CompAMa Vol.1, No.2, pp.7-45, 2013 - Accepted October 1, 2013

Abstract

Extreme events are an important part of climate variability and their intensity and persistence are often modulated by large scale climatic patterns which might act as forcing drivers affecting their probability of occurrence. When the North Tropical Atlantic (NTA) and the Equatorial Pacific (Niño 3 region) sea surface temperature (SST) anomalies are of opposite signs and the first one is positive while the second one is negative, the rainfall response is stronger in the northern coast of Venezuela as well as in the Pacific coast of Central America during the Nov-Feb period. The difference between these two SST anomaly time series (NTA-Niño3) is used in this analysis and it is called the Atlantic-Pacific Index or API. By fitting a dynamic generalized extreme value (GEV) model to station based daily rainfall at different locations and to the Xie and Arkin dataset for the Vargas state, we found the API index to be an adequate index to explain the probabilistic nature

[*]Departamento de Cómputo Científico y Estadística, Universidad Simón Bolívar, Caracas, Venezuela

[†]Centro de Ciencia do Sistema Terrestre, Instituto Nacional de Pesquisas Espaciais, São Paulo, Brazil

[‡]Centro de Ciencia do Sistema Terrestre, Instituto Nacional de Pesquisas Espaciais, São Paulo, Brazil

[§]Department of Mathematics and Statistics, University of New Mexico, Albuquerque, USA

[¶]Department of Applied Mathematics and Statistics, University of California, Santa Cruz, USA

of rainfall extremes in the northern Venezuelan coast for the months Nov-Feb. Dependence between the Atlantic-Pacific index and the probabilistic behavior of extreme rainfall was also explored for simulations from two global coupled General Circulation Models for the 20th century climate (20C3M experiment) and the 21st century climate (SRES A2 experiment): the Echam5 model and the HadCM3 model. A significant dependence of extreme rainfall on the Atlantic-Pacific index is well described by the GEV dynamic model for the Echam5 20C3M experiment model outputs. When looking at future climates under the SRES A2 experiment, the dependence of extreme rainfall from the API index is still significant for the middle part of the 21st century ($2046 - 2064$), while this dependence fades off for the latest part of the century ($2081 - 2099$).

Keywords: Extreme Rainfall, North Central Coast of Venezuela, Atlantic-Pacific Index, Generalized Extreme Value distribution, Echam5 model, HadCM3 model

1 Introduction

Extreme events are an important part of climate variability ([12]) and their intensity and persistence are often modulated by large scale climatic patterns which might act as forcing drivers affecting their probability of occurrence. From the probabilistic point of view these events are attached to values located at the tail of the probability distribution of the climatic variable of interest (e.g. temperature or precipitation), and can be associated with the upper or lower tail of the distribution. Drought events are considered as extreme events associated with the lower tail of the probability distribution; however their impacts are the result of successive occurrences of such events which make up cumulative seasonal anomalies with severe consequences for society and the environment.

Sea surface temperature (SST) anomalies measure how much is the temperature departure in degrees Celsius from normal conditions, for a particular time of the year. SST anomalies associated with ENSO (El Niño-Southern Oscillation) play an important role in defining the probability of occurrence of such events. The Amazonian drought of 2005 was analyzed in [13], where they concluded that the causes of the drought were not related to El Niño, but to an anomalously warm Tropical Atlantic. By examining total and extreme rainfall for South America during the period $1960 - 2000$, [6] concluded that the patterns of change of extreme events agreed with the patterns of change

for total rainfall and most of the observed coherence in the spatial variability of rainfall trends could be explained by ENSO conditions. More recently [16] examined the shift in the seasonal histograms of daily rainfall over South America conditional on the ENSO phase and their results suggest a broader impact on daily values than on the mean values of rainfall. In [5], they also show that there is more sensitivity to ENSO in the extreme range of daily rainfall than in the ranges of moderate to light rainfall. Therefore applying adequate tools to separate SST anomalies impacts over the middle range of the rainfall probability distribution from its impacts over the distribution tail might be more effective to assess ENSO impacts on extremes.

Exceptional rainfall events occurring during mid-December 1999 produced floods and landslides along the north central coast of Venezuela with over $10,000$ fatalities reported and economic looses estimated at over \$1.8 million ([11],[10]). Similar events occurred also in February 1951, February 2005 and November 2010. [19] also reported that many of these severe events documented in the region have occurred during the period November-February. An analysis of the atmospheric circulation and oceanic conditions was presented by [11], which gave place to the exceptional rainfall of December 1999. They found that anomalous tropical Atlantic conditions with warmer temperatures than normal, jointly with local conditions enhancing the convection of humid air from the Atlantic, were the main causes of the development of a strong convective activity in the region. At the end of their analysis, the following scientific questions were brought up: Are similar events likely to occur in the near future? Will they occur in the same region or in any other region? Is it possible to foresee such kind of events? Would it be possible to make provisions such that catastrophic impacts on the population could be reduced or avoided?

In terms of predictability, [4] suggest that Caribbean rainfall is influenced by both Atlantic and Pacific oceans but the strength of the rainfall response depends on how the SST anomalies in the tropical Atlantic and the eastern Pacific interact. The strongest response, they conclude, is when the tropical Atlantic is in the configuration of a meridional dipole (antisimmetric across the ITCZ) and the eastern tropical Pacific is of opposite sign to the tropical North Atlantic. Later [17] conclude that there seems to be a robust relationship between a late season Caribbean rainfall (August-September-October) and an east-west gradient of SST anomalies between the two equatorial ocean basins, with a cold equatorial Pacific and a warm equatorial Atlantic likely linked to an enhanced late season rainfall. On another hand,

the enhanced September-October-November rainfall season was analyzed by [10] as the main precursor conditions for December 1999 events. Negative SST anomalies in the tropical Pacific due to an evolving La Niña event and positive SST anomalies over much of the Caribbean and the central tropical Atlantic were associated with a northward shift in the ITCZ additionally to the anomalous low-level (925hPa) westerlies along Northern Venezuela ([10]). The tropospheric bridge effect resulting from the high degree of boreal winter connectivity between the two oceanic basins, affects rainfall variations in Central America and the Caribbean ([4]). These atmospheric teleconnections between the tropical Atlantic and the tropical Pacific ocean have also been demonstrated by [20].

Along these lines suggested by [4] and [17], the aim of this analysis is to find common patterns of oceanic and/or atmospheric variables which might be used as explanatory signals to extreme trend detection under present and future climates. The key approach is to look at the daily rainfall extreme value distribution and to model these data by using appropriate large scale climatic indices. To estimate the risk of Amazonian drought on future climates, [3] used the 21st century projections of the north-south SST gradient across the equatorial Atlantic (ANSG index), based on the results from the HadCM3LC General Circulation Model (GCM). This index correlates well with the reduction of dry season rainfall in western Amazonia and future projections of the ANSG index could be used to estimate probabilities of exceeding the 2005 value. In this research the oceanic influence on extreme rainfall trends for the north central coast of Venezuela are estimated by using different data sets. Station rainfall data is used to analyze local trend features. Large scale gridded rainfall data sets (CMAP data) are used for comparisons with 20^{th} century GCM simulations (20CM3 experiment) and future climate projections for the SRES A2 experiment from the IPCC AR4 scenarios ([14]). Daily precipitation and monthly sea surface temperature projections from the ECHAM5/MPI-OM and the UKMO HadCM3 GCMs were used for this purpose.

The organization of the paper is as follows: In section 2, the different types of data sets and the general data analysis methodology is described. In section 3, the extreme value analysis (EVA) and trend detection methods using non-stationary models are explained and applied to observed rainfall data sets. These methods are also applied to 20^{th} century $(1961 - 2000)$ modeled climate data from two global models; and the equivalent analysis for future climates projected by the `ECHAM5/MPI-OM` and HadCM3 GCMs

are explained in section 4. Discussion and conclusions about this work and future research is also provided at the end of the paper.

2 Data description and methodology

2.1 Data

Daily rainfall data from eight climatological stations in Vargas state, located at the north central coast of Venezuela were used in the first phase of the analysis. A list of the locations and their available period of records is presented in Table 1. A map with the station locations is presented in Figure 1. All stations lie in a horizontal strip of about 80 kilometers along the north central Venezuelan coast. Most locations are part of the climatological network operated by the Ministry of Environment and Natural Resources (MARN). The location Maiquetía -Aeropuerto is operated by the Venezuelan Air Force (FAV).

Tab. 1: Rainfall stations in Vargas state.

Name of Station	Longitude	Latitude	Elevation	Period of records
Maquetia-Aeropuerto	66° 59'	10° 36'	43	1961 – 1999
Catia La Mar	67° 02'	10° 37'	7	1970 – 1983
Maiquetia	66° 57'	10° 36'	75	1948 – 1983
Macuto	66° 53' 47"	10° 36' 01"	53	1951 – 1999
Mamo	67° 03' 13"	10° 35' 22"	16	1956 – 1983
Puerto Cruz	67°21'	10° 32'	5	1951 – 2006
La Guitarrita	67° 16'	10° 28'	-	1948 – 2006
Los Caracas	67° 34' 42"	10° 28'	15	1967 – 2006

Data from NCEP/Reanalysis II were also downloaded from the site `http://www.cdc.noaa.gov/cdc/data.ncep.reanalysis.html`. Several variables were investigated: sea surface temperature, wind fields at 300 and 850 hPa levels, relative and specific humidity at 300 and 850 hPa levels, geopotential heights at 300 and 850hPa levels, sea level pressure and outgoing long wave radiation.

Satellite derived pentad rainfall from the CMAP (CPC Merged Analysis of Precipitation) data set found in `http://www.cdc.noaa.gov/cdc/data.cmap.html` ([21]) were also used in

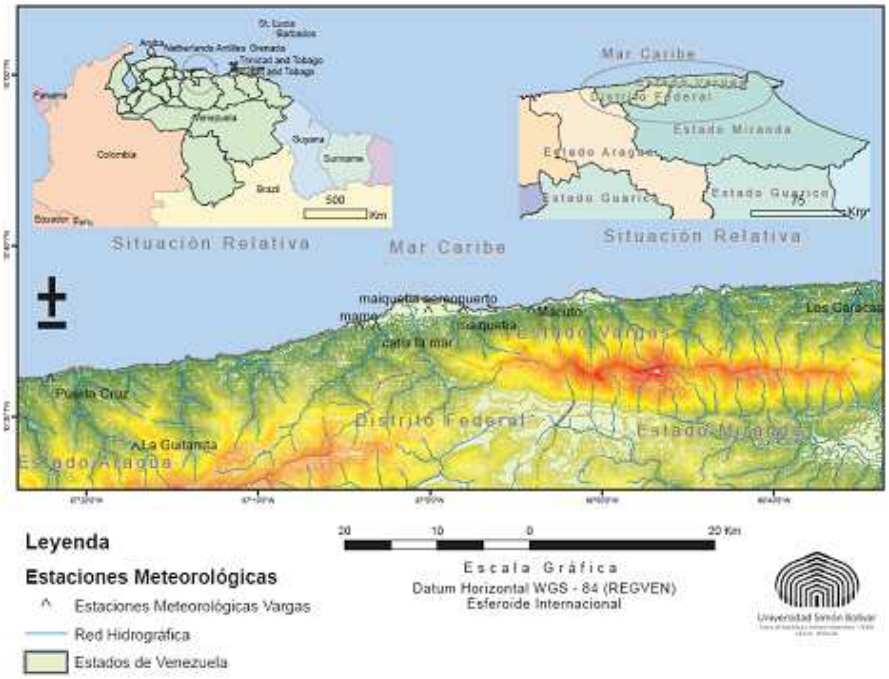

Fig. 1: Study area and location of the eight rainfall stations in Vargas state, northern central coast of Venezuela.

the analysis. The CMAP data provides pentad and monthly analyzes of global precipitation in which observations from rain gauges are merged with precipitation estimates from several satellite-based algorithms (infrared and microwave). The resolution of the data set is of 2.5 degrees, and it is available from January 1979 until present. The period $01/01/1979 - 02/28/2005$ was used in this analysis.

Daily rainfall and monthly sea surface temperature from the `ECHAM5/MPI-OM` model ([9]) outputs from the CMIP3 World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project phase 3 multi-model) for the 20th century IPCC experiment (20C3M, $1961 - 1999$) and the middle section ($2046 - 2064$) and final section ($2081 - 2099$) of the 21st century IPCC experiment SRES A2 were also used in the analysis. Daily rainfall and monthly sea surface temperature from the UKMO HadCM3 model ([8]) were available from $1961 - 1989$ (20CM3 experiment), and for the period $2070 - 2099$ for the SRES A2 IPCC experiment.

2.2 Methods for data analysis

Daily rainfall data from eight meteorological stations from Vargas state were available for analysis. Some quality control was required in order to amend problems with missing data, accumulated records and outliers. Monthly boxplots for the mean daily rainfall at each location for the base climatological period $1961-1990$ are shown in Figure 2. The median values represented by the horizontal lines in the boxplot diagrams give an indication of the annual cycle, with maximum values during the months of June, July and August. These are the central months comprising which is regionally called the rainy season. [17] discriminates about an early rainfall season (May-June-July) and a late rainfall season (August-September-October) for the Caribbean. The dry season comprises the months from November to April, with the lowest values observed for the months of January, February and March. The station *La Guitarrita* presents a different pattern from the rest of the stations, which is probably due to its location which does not lie exactly on the coast. The station *Los Caracas* presents very high inter annual variability with high inter-quartile range observed specially for the months of November-December, as indicated for the heights of the boxplots. Our period of interest in this study is November to February which comprises the first four months of the dry season. Hereafter, for short we will call November-February as the dry period and June-July-August as the wet period.

Several rainfall statistics were calculated from the daily rainfall data at the eight locations for all available years:

1. Monthly maximum daily rainfall for the four months comprising the dry period.

2. Total rainfall during the wet period June-August; total rainfall during the dry period and total rainfall dry period anomalies using the base period $1961-1990$.

3. Annual ratio of the dry period to the wet period.

4. Monthly mean daily rainfall in mm/day and monthly mean daily rainfall anomalies using the base period $1961-1990$ for the dry-period.

From the annual ratio of the dry period to wet period, a composite of years in where this ratio is greater than 3 (dry period total rainfall is at least three times greater than wet period total rainfall) and lower than 0.5

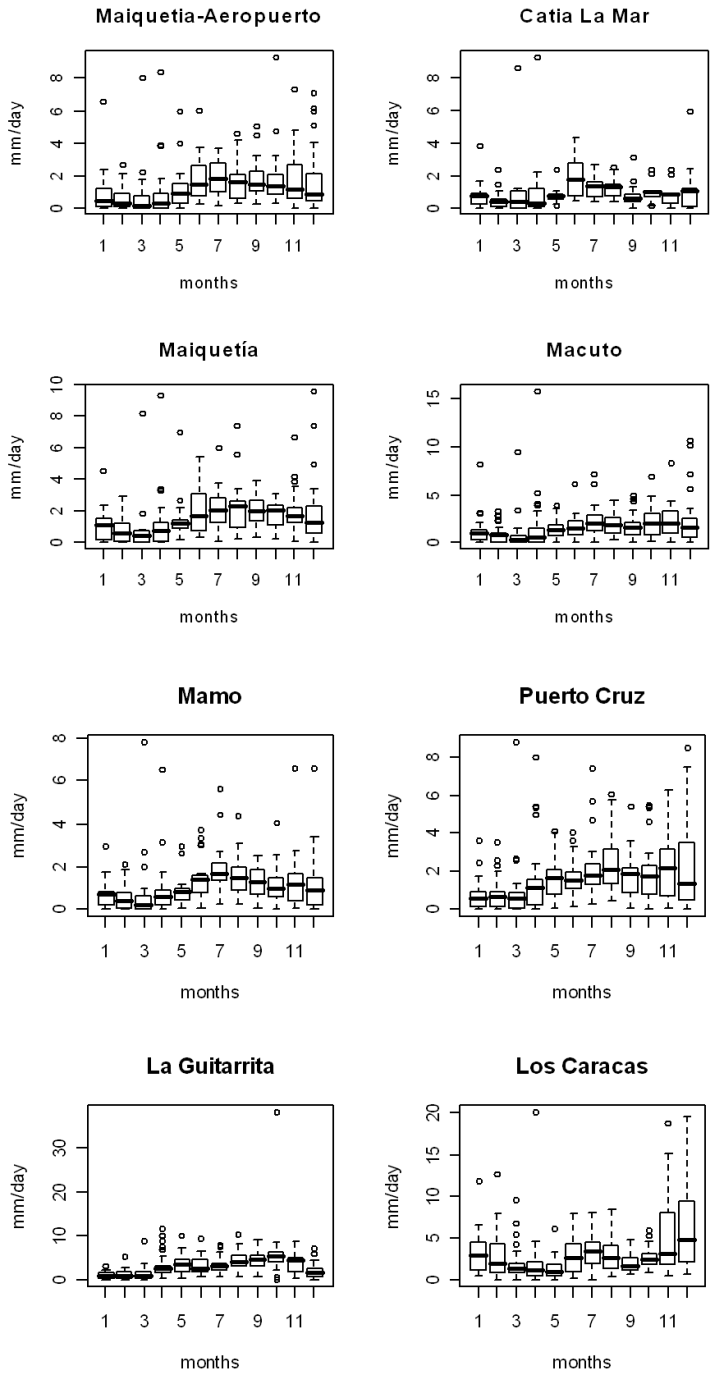

Fig. 2: Monthly boxplots for the mean daily rainfall (mm/day) at eight rainfall stations in the Vargas state, Venezuela for the period $1961-1990$

.

(wet period total rainfall is at least twice the dry period total rainfall) were selected. The ending year of the period November-February is used as an indicator. The set of wet years comprising the period of records is: 1951, 1976, 1986, 1992, 2000, and 2005. The set of dry years is 1960, 1964, 1971, 1973, 1978, 1984, 1987, 1994, 1995, 1996, 1998, 2003, and 2007.

A list of daily precipitation values close to or greater than 100 mm and their dates of occurrence during the selected wet years is given in Table 2. The events of the year 1951, 1999 and 2005 were mentioned in the introduction as very important events in terms of the damage they caused; and they became included within the list of wet years.

Tab. 2: Extreme daily rainfall amounts observed during the wet years for the dry season (Nov-Feb).

Year	Location	Amount	Date
$1950-1951$	Maiquetia	193.0	16/02/1951
$1975-1976$	Los Caracas	100.6	17/02/1976
$1985-1986$	Macuto	130.21	04/12/1985
	Los Caracas	106.8	01/11/1985
	Los Caracas	159.0	07/12/1985
	Maiquetía-Aerop.	118.7	04/12/1985
$1991-1992$	Macuto	152.95	26/11/1991
	Los Caracas	157.0	19/11/1991
	Maiquetía-Aerop.	98.7	24/11/1991
$1999-2000$	Maiquetía-Aerop.	410.4	15/12/1999
$2004-2005$	Puerto Cruz	130.9	09/02/2005
	La Guitarrita	103.6	07/02/2005
	Los Caracas	249.5	08/01/2005
	Los Caracas	162.6	08/02/2005

The following meteorological fields were calculated as composite fields anomalies during the dry season for the wet and dry years: geopotential heights at 300 and 850 hPa pressure levels; wind speed and vector wind at 300 and 850 hPa pressure levels; outgoing long wave radiation; sea surface temperature anomalies and precipitation anomalies (mm/day) from the Xie and Arkin pentad rainfall data set. Only the figures for precipitation anomalies and SST anomalies are shown.

Figure 3 shows the precipitation anomalies during November-February in

the dry (bottom) and wet (top) years. It is evident that the northern coast of Venezuela and the Pacific coast of Central America become affected by anomalous rainfall during these years. In Figure 4 sea surface temperature (SST) anomalies are presented and in Figure 5 a wider window is used to monitor the behavior of the equatorial Pacific ocean.

When the north tropical Atlantic and the equatorial Pacific SST anomalies are of opposite signs and the first one is positive while the second one is negative, the rainfall response is stronger in the northern coast of Venezuela as well as in the Pacific coast of Central America. These results agree with the results presented by [4]. The authors also discussed the benefits of this situation for predictive efforts in comparison with a situation in which both oceans have the same sign anomalies. [17] assure that a cold equatorial Pacific or a warm equatorial Atlantic are linked to an enhanced late season rainfall. They make a distinction between an early rainfall season comprising the months May-June-July, and a late season comprising the months August-September-October.

Although their study is not covering specifically the period November-February, an enhanced late rainfall season might well be the precursor for disastrous pre-conditions leading to catastrophic landslides and floods in a similar situation as the one described by [10] for the 1999 event. [17] also find evidences of the influence of an east-west gradient of SST anomalies in the tropical Pacific and Atlantic on Caribbean Rainfall, with a tendency of a warm Atlantic-cool Pacific to favor Caribbean rainfall. However they stress the fact that this mechanism is only true for late Caribbean rainfall season.

2.3 Atlantic-Pacific index

The antisymmetric configuration between the Niño 3 region in the Equatorial Pacific (5° S-5° N, 90° W- 150° W) and the North Tropical Atlantic (NTA) region (60° W-20° W, 6° N -18° N and 20° W -10° W, 6° N -10° N) is proposed as the large scale climate signal to be investigated in this study, and its possible influence on extreme rainfall in the north central Venezuelan coast is the center of concern. The difference between these two SST anomalies series: NTA-Niño3 is used in the analysis and it is called the *Atl-Pac Index* or API. The mean values of this time series are shown in Figure 6.

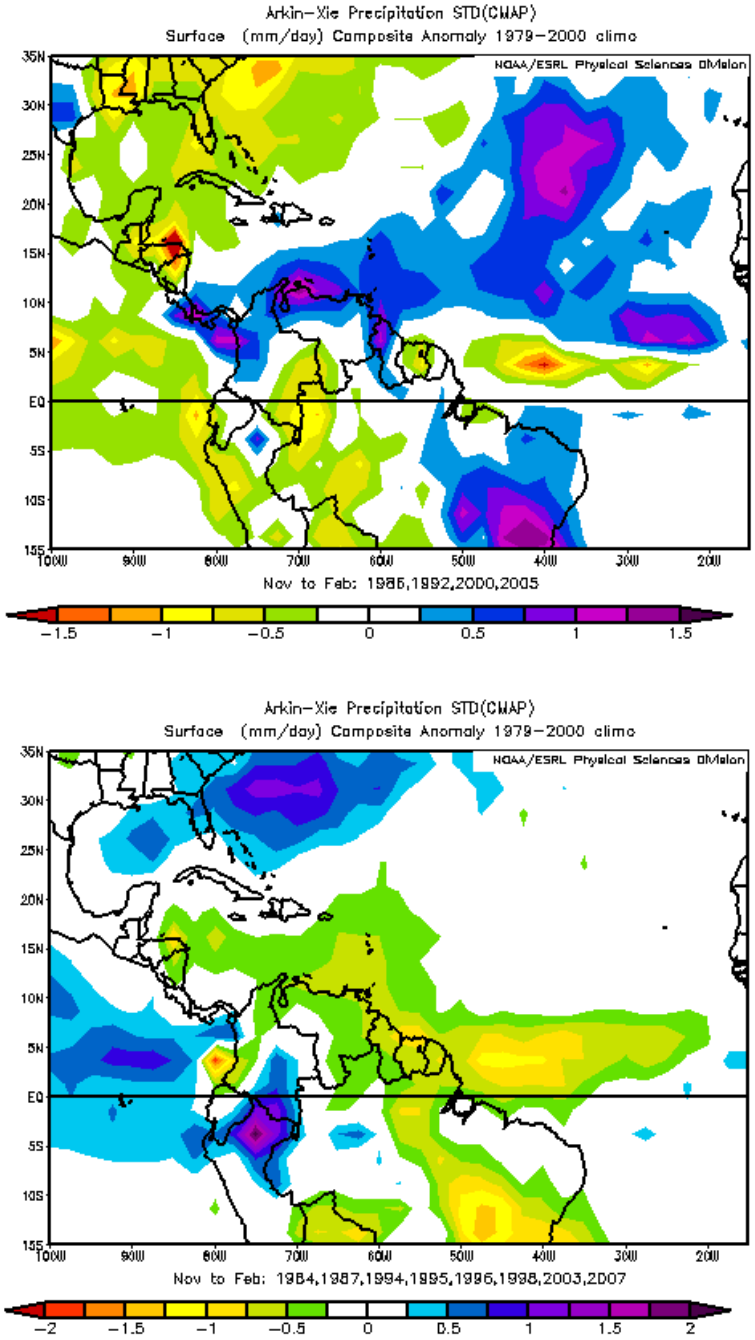

Fig. 3: Precipitation anomalies (mm/day) during November-February for wet (top) and dry (bottom) years.

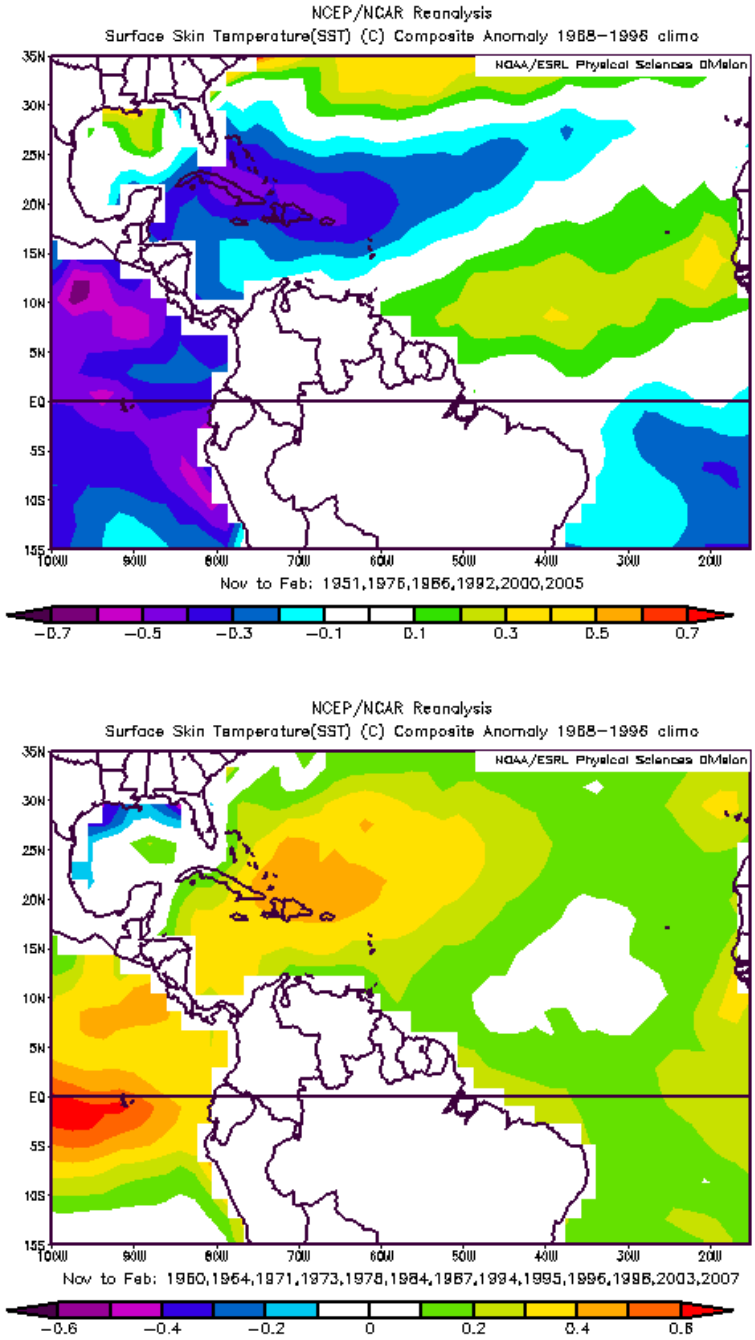

Fig. 4: SST anomalies (C) during November-February for wet (top) and dry (bottom) years.

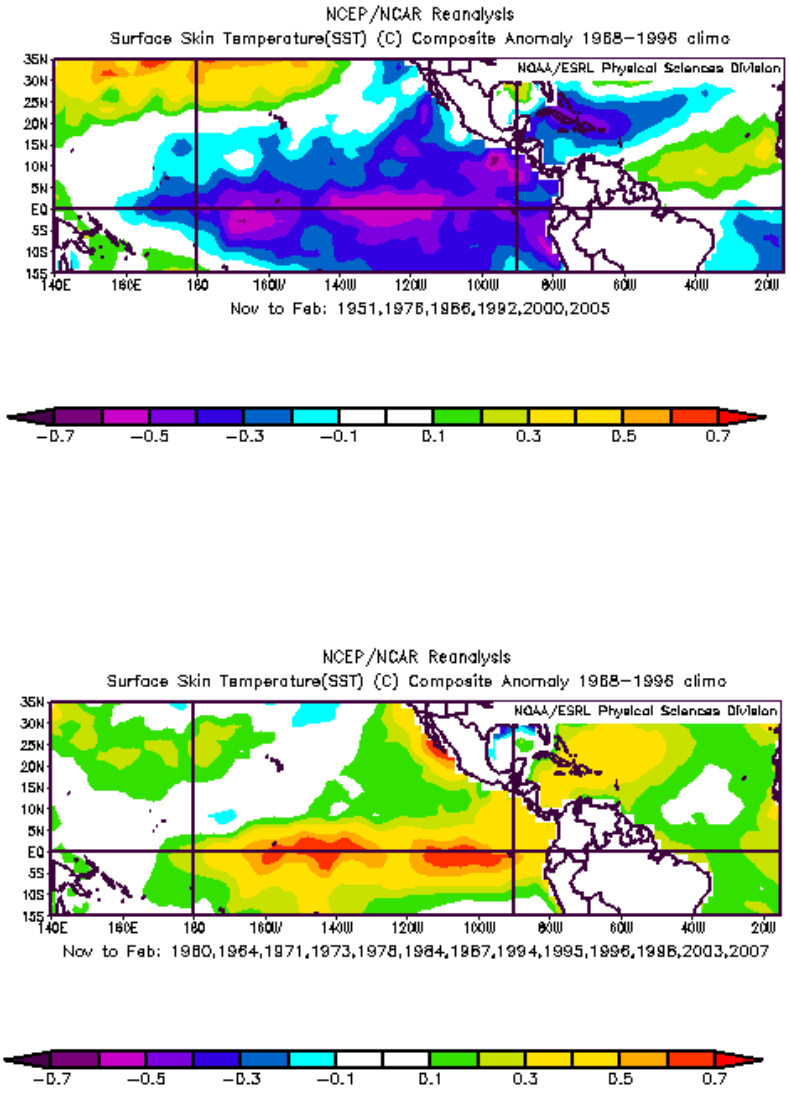

Fig. 5: SST anomalies (C) during November-February for wet (top) and dry (bottom) years including the Equatorial Pacific.

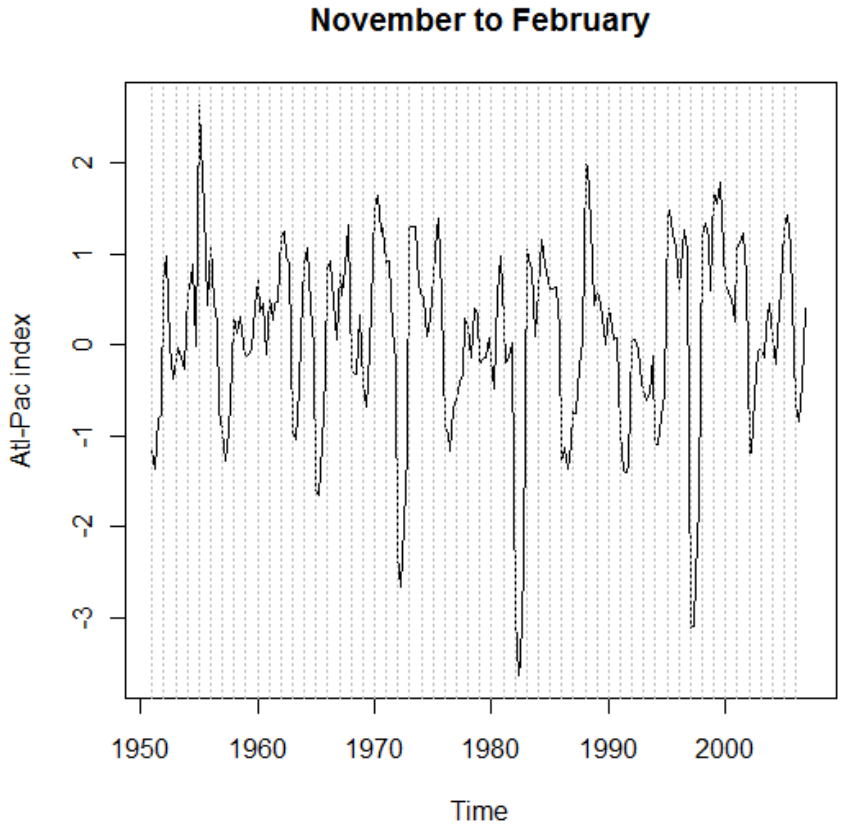

Fig. 6: Monthly time series of the SST (°C) anomalies differences (Atl-Pac Index) between the NTA index region and the Niño 3 region for the dry period months (November to February) during 1951 – 2007

3 Results

3.1 Seasonal total rainfall Analysis

The question remains on whether positive values of the Atl-Pac index are effectively associated with a stronger rainfall response in all locations. Figure 7 shows how the highest total rainfall anomalies from the base period $1961-1990$ during November-February are mostly associated with high values of the mean November-February Atl-Pac Index (API), while the lowest anomalies are linked with the most negative values of the API. The least square fitted lines have been overlapped to all figures.

The Spearman correlation coefficient between the total rainfall anomalies in November-February and the Atl-Pac Index are shown in Table 3 for the eight locations in Vargas state.

Tab. 3: Spearman correlation coefficient between the Total Rainfall anomaly in Nov-Feb and the Atl-Pac Index.

Name of the station	Spearman Correlation Coefficient (ρ_s)	p-Value
Maquetia-Aeropuerto	0.42	0.00814
Catia La Mar	0.57	0.03174
Maiquetia	0.39	0.02306
Macuto	0.34	0.01697
Mamo	0.21	0.27350(*)
Puerto Cruz	0.30	0.02425
La Guitarrita	0.43	0.00079
Los Caracas	0.37	0.01977

(*) (ρ_s) is not significant at the 95% confidence level

Most of the correlations as suggested in Figure 7 are statistically significant at a 95% level (p-values < 0.05), except for location Mamo. However we are more interested in the influence of API index on the rainfall extremes instead of total or mean rainfall values. We now describe how we can detect a possible signal in the rainfall extremes in relation to the evolution of the Atl-Pac Index time series.

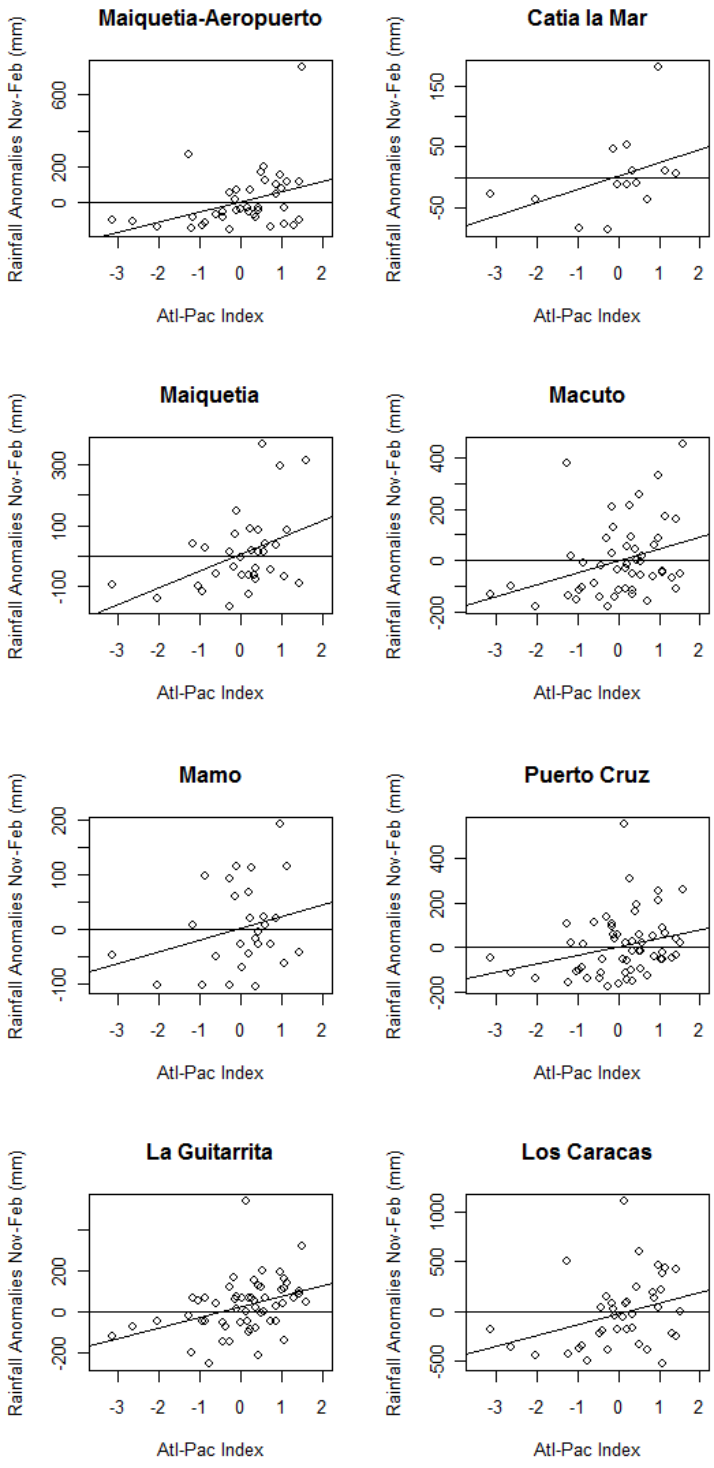

Fig. 7: Total Nov-Feb rainfall anomalies vs. mean Nov-Feb Atl-Pac Index. Rainfall anomalies are calculated with respect to the base period 1961-1990. Least-squares fitted lines have been overlaid.

3.2 Extreme rainfall analysis for 20*th* century observations

Classical extreme value theory looks for the probabilistic behavior of values which are defined as the maximum values of a sequence of observations measured on a regular time-scale (hourly, daily, etc.) during a fixed time length observation interval (weeks, months, years, etc.), which is known as the block maxima approach. For example if daily rainfall is available for several years, we consider the sequence of annual maxima. A different approach is to select all values above a specified high threshold. In this case an observation is characterized as an extreme value if the observation exceeds the threshold (threshold excesses approach). More details on both approaches and their generalizations can be investigated in [1].

In order to analyze rainfall extremes and their frequency characteristics we selected the block maxima approach and used the Generalized Extreme Value (GEV) probability distribution to assess the probabilistic behavior of the annual daily maxima. This model is fitted to the different rainfall data sets, point and gridded, including the observed values for the present climate, the GCMs simulated values for the $20th$ century climate (experiment 20C3M) and the GCM simulated values for the future climates in the $21st$ century (SRES A2 experiment). However within each time fitting period the model parameters might be kept fixed or might change with time, making up the stationary or non-stationary case. In the non-stationary case, one option is to assume a constant parameter change with time or to go further and to assume a changing dependence with time. GEV model fitting procedures under the stationary and non-stationary cases will be described as follows.

3.2.1 Stationary extremes

We consider the maximum daily rainfall per month in mm/day for the eight locations during the period Nov-Feb for all available years. For the Xie and Arkin data set which is in pentads, we take the maximum pentad value (in mm/day) for each month in Nov-Feb of all available years. We also calculate the average rainfall in the window covering the northern central coast (7.5°N-12.5°N, 72°W-60°W) of Venezuela, given the coarse resolution of the data set ($2.5° \times 2.5°$).

We consider time blocks of fixed lengths (months) and take the maximum value for each block (block maxima). The data sequence $z_1, z_2, \ldots, z_m$

is assumed to be conditionally independent and it can be modeled by the Generalized Extreme Value (GEV) distribution. The GEV probability distribution function is given by the equation:

$$F(z) = \exp\{-[1 + \xi(\frac{z-\mu}{\sigma})]_+^{-1/\xi}\} \quad (1)$$

where μ is a location parameter, σ is a scale parameter and ξ is a shape parameter. This distribution is flexible enough to encompass the Fréchet class ($\xi > 0$), the Weibull class ($\xi < 0$) and the Gumbel class ($\xi = 0$). The $+$ sign denotes the positive part of the argument. An extensive discussion about this and related models is given in [1]. The distribution parameters can be estimated by maximum likelihood by assuming independence of the data conditional on μ, σ, ξ and maximizing the likelihood of the sample maxima by using (1). A Bayesian estimation procedure can also be carried out by providing prior distributions for μ, σ, ξ and calculating the posterior probability distribution of conditional on the sample maxima as in [2].

Given an extreme value y with probability distribution given by (11) for known values of μ, σ, ξ , the quantity $H(y) = 1 - F(y)$ is the probability of exceeding the value of y, and the expression $N(y) = 1/H(y)$ is called the"return period" of y, since the level y will be exceeded on average once each $N(y)$ years ([1]). Changes in the location parameter μ with time would lead to changes in the exceeding probabilities $H(y)$ for a fixed level y. This in turn indicates that the recurrence or return period for a given precipitation level would also change with time.

3.2.2 Non-stationary extremes

Under non-stationary conditions, as for example, under climate change conditions, rainfall extremes might change with time. This situation can be modeled by assuming that the GEV model parameters might also be time dependent. [1] proposes several alternatives, for example, $\mu_t = \alpha + \beta t$, if the location parameter changes linearly with time; or $\mu_t = \beta_0 + \beta_1 t + \beta_2 t^2$ if the change is quadratic. The time change in the model parameters might also be modulated by external variables or co-variables. In this case we might have an expression of the form $\mu_t = \alpha + \beta X_t$, which indicates that the variable extreme behavior depends on the variable X_t. Changes in the scale and shape parameters might also be included in the non-stationary formulation.

For the time varying location models the sample maxima $z_1, z_2, \ldots, z_m$ are assumed to follow a GEV distribution with parameters μ, σ, ξ of the form:

$$F(z) = \exp\{-[1 + \xi(\frac{z_t - \mu_t}{\sigma})]_+^{-1/\xi}\} \quad (2)$$

Tab. 4: Non-stationary GEV model parameters with location parameter expressed as $\mu_t = \alpha + \beta X_t$ and X_t = Atl-Pac Index. Scale and shape parameters are kept constant. Values in parenthesis are the standard error estimates.

Name of Station	α (intercept)	β (slope)	σ (scale)	ξ (shape)	Loglikehood ratio test (p-value)
Maquetia-Aeropuerto	6.47	0.49	7.66	0.80	0.077
	(0.78)	(0.30)	(0.88)	(0.12)	
Catia La Mar	4.38	0.19	4.66	0.57	0.60(*)
	(0.79)	(0.38)	(0.77)	(0.18)	
Maiquetia	6.21	0.32	7.52	0.68	0
	(0.86)	(0.47)	(0.89)	(0.14)	
Macuto	7.61	0.59	8.76	0.68	0.073
	(0.79)	(0.36)	(0.83)	(0.11)	
Mamo	4.54	0.36	5.50	0.74	0.245(*)
	(0.64)	(0.33)	(0.70)	(0.13)	
Puerto Cruz	5.35	0.48	6.44	0.79	0.192(*)
	(0.52)	(0.37)	(0.59)	(0.09)	
La Guitarrita	7.85	0.98	7.70	0.47	0
	(0.60)	(0.41)	(0.57)	(0.07)	
Los Caracas	17.14	1.06	17.15	17.15	
	(1.61)	(0.76)	(1.57)	(0.09)	0.142(*)
Xie and Arkin	1.74	0.28	1.27	0.28	0.007
	(0.15)	(0.11)	(0.13)	(0.12)	

(*): There is not a significant difference between the stationary and the non-stationary model at a 90% confidence level.

In Table 4 results from fitting the GEV distribution by assuming a location parameter depending on the API Index as a covariable are presented. The library *ismev* from the R package ([15]) was used in the analysis which

uses the maximum likelihood method for parameter estimation. The last column is the p-value corresponding to a Chi-square distribution with one degree of freedom, when comparing the stationary model (location parameter does not depend on the API Index) with the non-stationary model (location parameter does depend on the API Index) using the Loglikehood ratio test.

There are four locations out of eight in where the dependence of the location parameter of the GEV distribution on the API Index is not significant, while in the other four there is a significant difference at a 90% confidence level ($p < 0.1$) between the two models. For the Xie and Arkin data set (last row in Table 4) the location parameter of the GEV distributions results significantly dependent on the API index, which favors the selection of the non-stationary model.

3.2.3 Time varying non-stationary extremes

In the models proposed above the intercept (baseline) and slope (rate of change) of the location parameters are fixed for the whole period of study. A further extension of the problem is to consider these two parameters as random variables which might change also with time. This formulation is called the Dynamic Linear Model (DLM) approach proposed by [18]. The general methodology applied to the GEV distribution is described in [7]. This approach is also used in this study to model extreme rainfall for all available data sets. In a DLM formulation with a time varying location parameters, a system of equations models the changes in the μ_t parameter with time:

$$\begin{aligned} \mu_t &= \theta_{1t} + \theta_{2t} X_t + \varepsilon_t; \quad \varepsilon_t \sim N(\mathbf{0}, V) \\ \theta_t &= G_t \theta_{t-1} + \omega_t; \quad \omega_t \sim N(\mathbf{0}, W) \end{aligned} \tag{3}$$

The first equation is the observation equation and the second equation is the evolution equation. In (3) $\theta_t = \begin{pmatrix} \theta_{1t} \\ \theta_{2t} \end{pmatrix}$; G_t is the 2×2 evolution matrix; V is the observation variance and W is the 2 evolution covariance matrix. t varies from $t = 1, \ldots, m$ where m is the number of time periods (months). m is of the form $m = 4 * k$ where 4 represents the number of months in the dry period (Nov-Feb) and k is the number of years of data. The parameters θ_{1t} and θ_{2t} represent the baseline (intercept) and rate of change (slope) of the GEV distribution location parameter.

A Bayesian approach is used to estimate the model parameters. Under the Bayesian paradigm all parameters become random variables and a full poste-

rior probability distributions can be obtained for each parameter conditioned on the sample and the prior knowledge provided by the prior distribution. A Markov Chain Monte Carlo (MCMC) methodology is used to get samples from the posterior probability distributions of $\theta = (\theta_1, \theta_2, \ldots, \theta_m), \mu = (\mu_1, \mu_2, \ldots, \mu_m), \sigma, \xi$, as proposed in [7].

In our analysis we propose a location parameter μ changing with time and depending on the evolution of the Atl-Pac Index. This hypothesis can be tested by analyzing the posterior probability intervals of $\theta_t = (\theta_{1t}, \theta_{2t})$ for all time periods $(t = 1, \ldots, m)$. The evolution in time of these parameters given by the mean, the median and the 10th and 90th percentiles are given in Figures 8, 9 and 10 for the Xie and Arkin data set and the locations of Maiquetía-aeropuerto and Mamo. They both showed one of the highest and the lowest level of significance respectively (given by the 1-p value) in their total rainfall anomalies vs. Atl-Pac Index correlation (See Table 3). Ten thousand samples were simulated form the posterior probability distributions of the whole parameter set, $\theta = (\theta_1, \theta_2, \ldots, \theta_m), \mu = (\mu_1, \mu_2, \ldots, \mu_m), \sigma, \xi$, and the first $5,000$ samples were discarded as *burn-in* sample for posterior statistic calculations.

Model fitting to the Xie and Arkin data set demonstrates that the dependence of the GEV distribution location parameter on the Atl-Pac Index is statistically significant. The 80% probability intervals (lines between the 10th and 90th percentiles) show that the model slope is entirely positive and does not contain the zero value. This departure from zero is less evident for Mamo (Figure 10) while the signal is stronger for Maiquetia-aeropuerto, where the slope seems to be increasing with time. In both cases the baseline parameter clearly increases with time indicating that the location parameter of the GEV distribution is increasing.

3.3 Extreme rainfall analysis from Global Models

Data from the Echam5/MPI-OM and UKMO_HadCM3 model outputs were used in this part of the analysis. As a first step global model outputs from the 20_{th} century experiments are compared with contemporary data sets taking care of the temporal and spatial resolutions of the data sets to be compared. Sea surface temperature data from the 20C3M IPCC experiment $(1961 - 1999)$ were used to calculate the Atl-Pac Index. A comparison of the calculated index from the global models ECHAM5 and HadCM3 with the NCEP data is presented in Figure 11. Although we do not expect similar

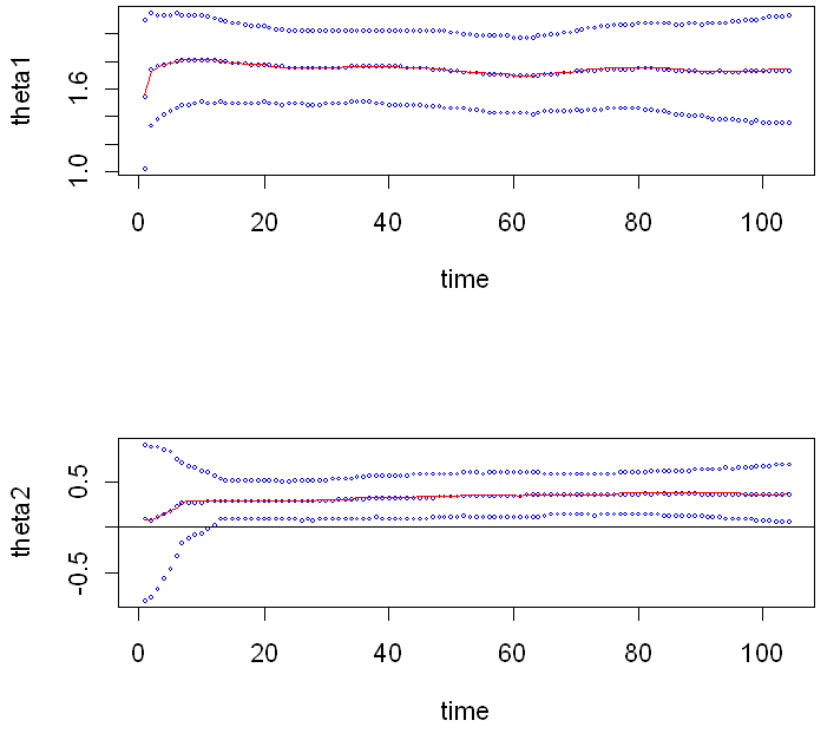

Fig. 8: Baseline (top) and slope (bottom) of the linear model relating the GEV location parameter with the API index. Values are given for each time interval (Nov-Feb months) for the period 01/01/1979 – 28/02/2005 (108 months) for the Xie and Arkin data. Dotted lines are the $90th$, $50th$ (median) and $10th$ percentiles. Continuous line corresponds to the simulated mean values.

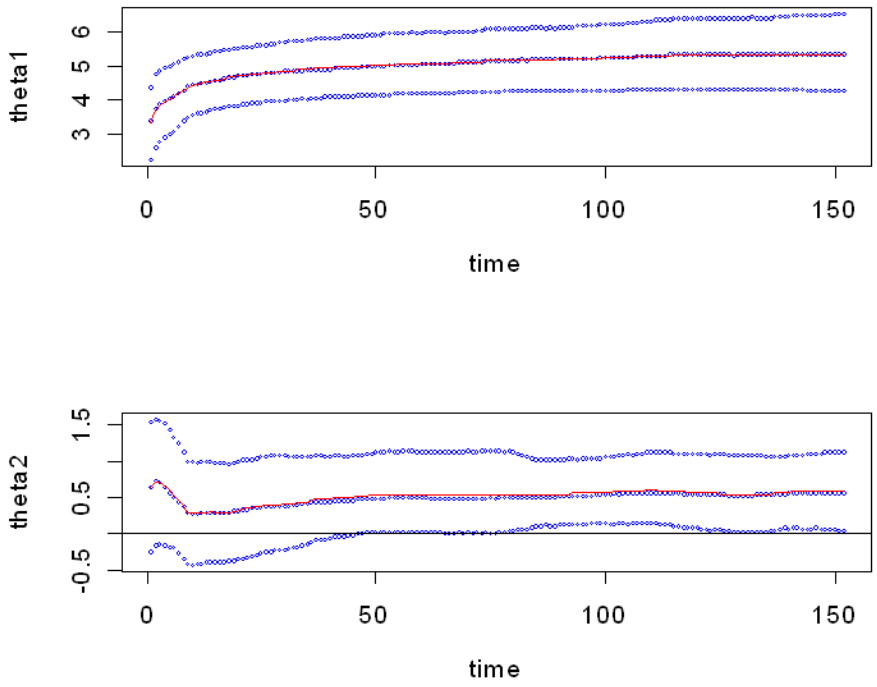

Fig. 9: Baseline (top) and slope (bottom) of the linear model relating the GEV location parameter with the API index. Values are given for each time interval (Nov-Feb months) for the period 01/01/1961 – 31/12/1999 (154 months) for Maiquetia-aeropuerto station. Dotted lines are the $90th$, $50th$ (median) and $10th$ percentiles. Continuous line corresponds to the simulated mean values.

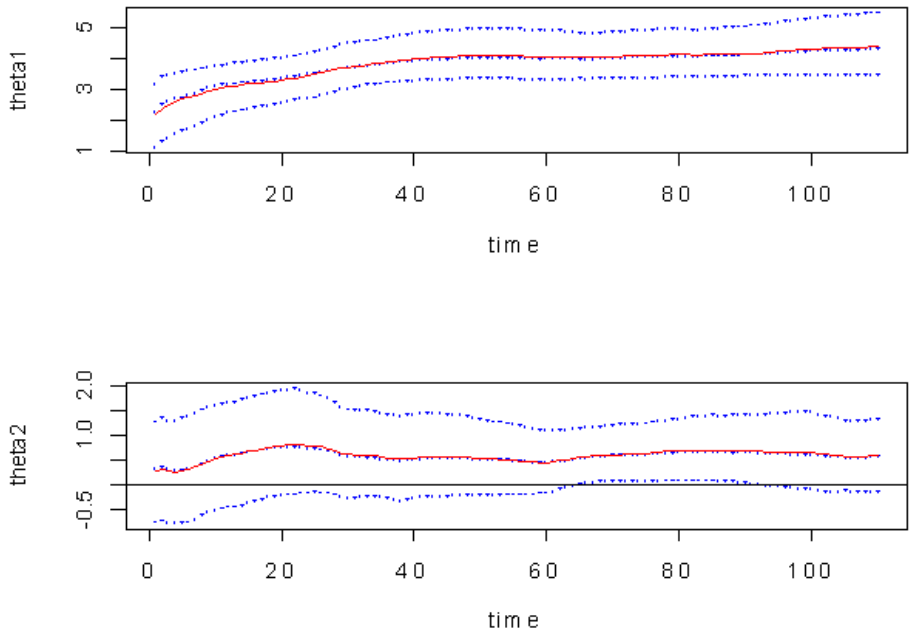

Fig. 10: Baseline (top) and slope (bottom) of the linear model relating the GEV location parameter with the API index. Values are given at each time interval (Nov-Feb months) for the period 01/01/1956 – 31/12/1983 (110 months) for Mamo station. Dotted lines are the $90th$, $50th$(median) and $10th$ percentiles. Continuous line corresponds to the mean values.

values at the same dates, the figure shows the contrast between the modeled and the observed values.

Monthly maximum pentad values (five-days accumulated rainfall) for the November-February period were calculated for the Echam5 and HadCM3 20C3M IPCC experiment and compared with the Xie and Arkin data set. To make them comparable, these rainfall values were converted to pentad averages in the domain 7.5°N-12.5°N and 72°W-60°W for the common period of records. A comparison of both data sets is presented in Figure 12. As observed from the Figure, the peak timing for both data sets is practically the same, but the rainfall intensities are overestimated by the Echam5 model in the period Nov-Feb $1979-1999$ and underestimated by the HadCM3 model in the period Nov-Feb $1979-1989$. These two periods comprise the common periods for each data set respectively.

From the analysis of the Xie and Arkin data the maximum pentad value registered for the whole period of records ($01/01/1979-28/02/2005$) was in November 1999. This maximum level was of $y = 12.85$ mm/day. However one must remember that this is a five day average value spatially averaged over a rather large geographic window of 7.5°N-12.5°N and 72°W-60°W, which is far from the observed 410 mm/day observed at Maiquetia station on the 15th of December 1999.

One of the aims of the analysis is to investigate whether the relationship observed between the oceanic variables and the extreme rainfall holds for the GCMs model simulations for present and future climates.

The relationship between the monthly maximum pentad precipitation for November-February and the Atl-Pac Index is presented in Figure 13 for the Echam5 data (top) and the Xie and Arkin data (bottom). Similar patterns can be observed in both figures. In Figure 14 similar graphs are presented for the HadCM3 data (top) and the the Xie and Arkin data (bottom). The main feature of Figure 13 is that the proportion of values above for example, 4 mm/day, are higher for API values above -1, while for low values of the API index the probability of exceeding such a rainfall level is very small. This is true for the Xie and Arkin data set and for the Echam5 data set. However this relationship is not followed by the HadCM3 data set as shown in Figure 14. In this case observed and modeled data show a different dependence structure, since moderate extreme rainfall values have a higher probability of occurrence for API value less than -1 in the HadCM3 data set, when compared with the Xie and Arkin data set.

This behavior is reflected when the dynamic GEV model is fitted to both

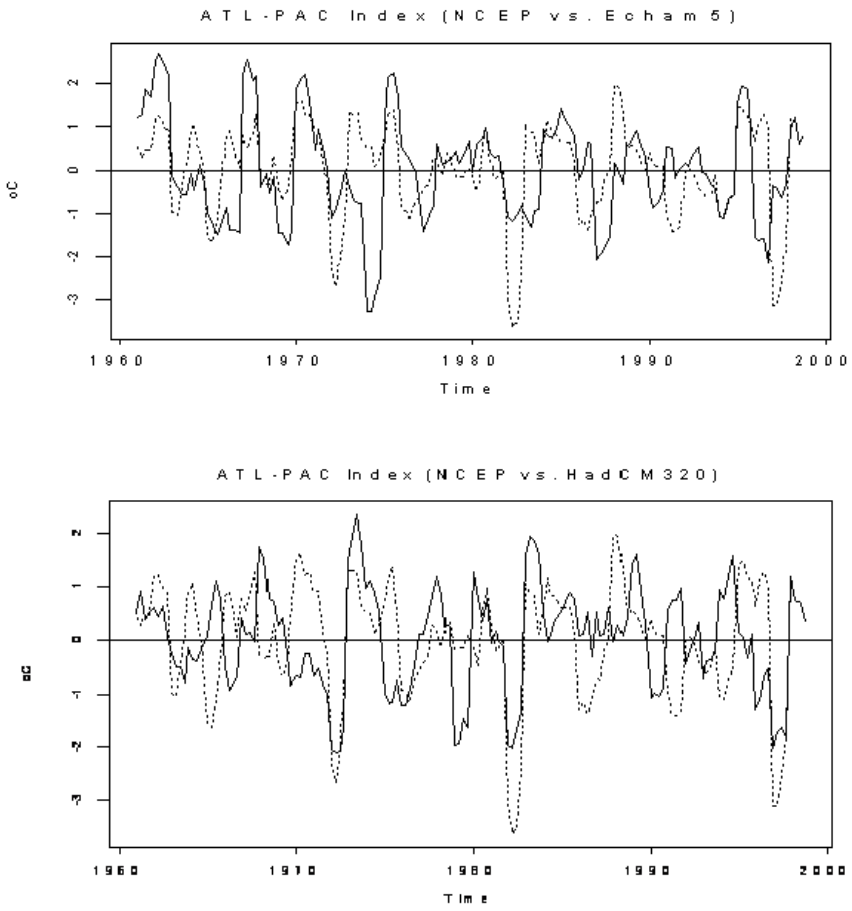

Fig. 11: Atl-Pac Index comparison between the Echam5 (top) and HadCM3 (bottom) (20C3M IPCC experiment) (solid lines) and the NCEP data (dotted line) for period $1961 - 1999$.

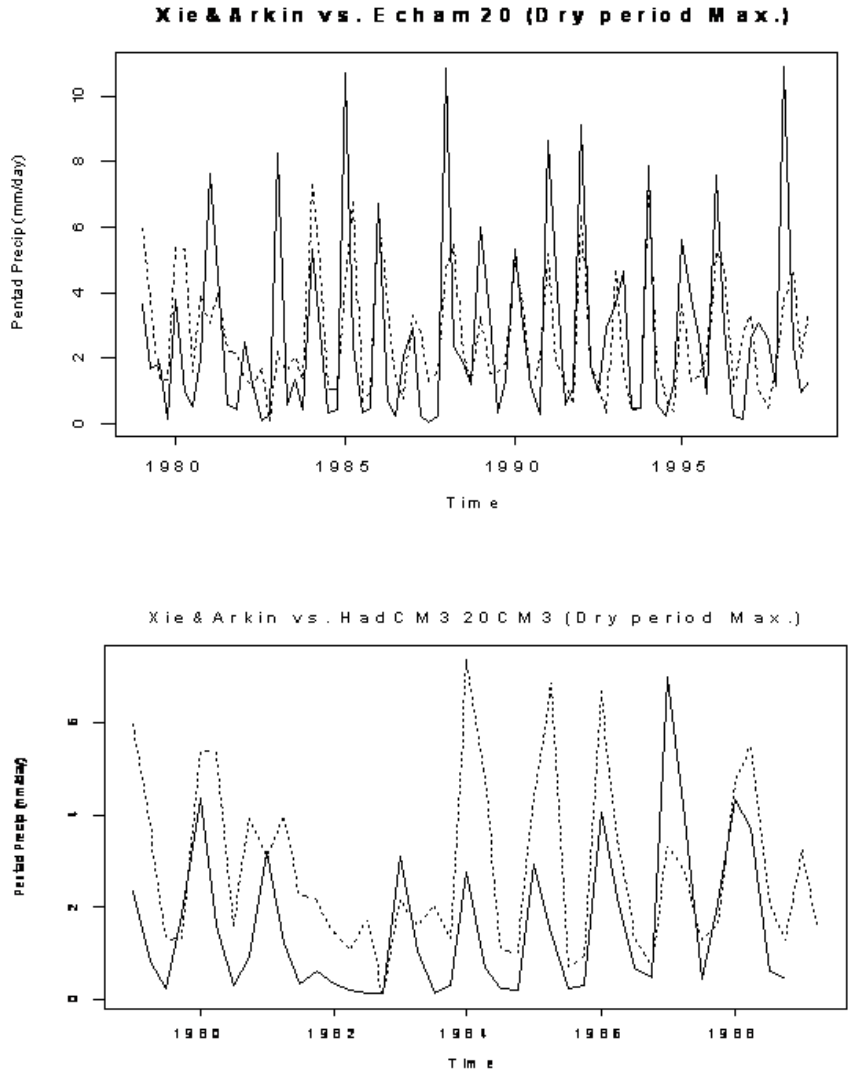

Fig. 12: Comparison of monthly maximum pentad rainfall in November-December between the Echam5 model (top) and the HadCM3 model (bottom) (20C3M IPCC experiment) (solid lines) and the Xie and Arkin data set (dotted lines).

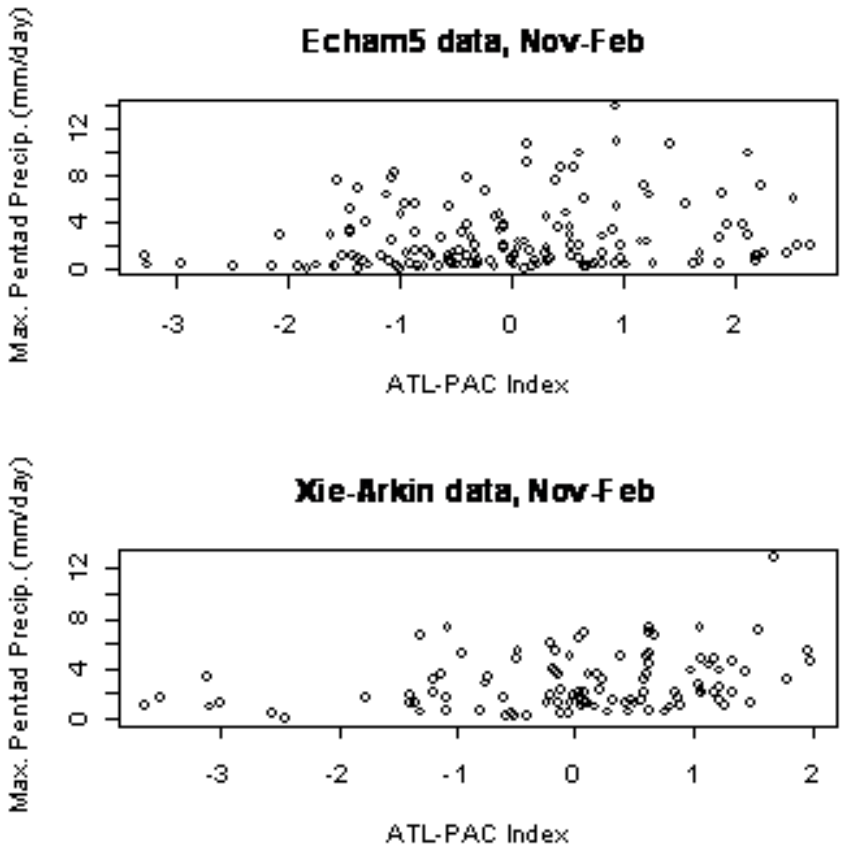

Fig. 13: Monthly maximum pentad versus the Atl-Pac Index for the Echam5 (top), and the Xie-Arkin rainfall data set for the common period of records.

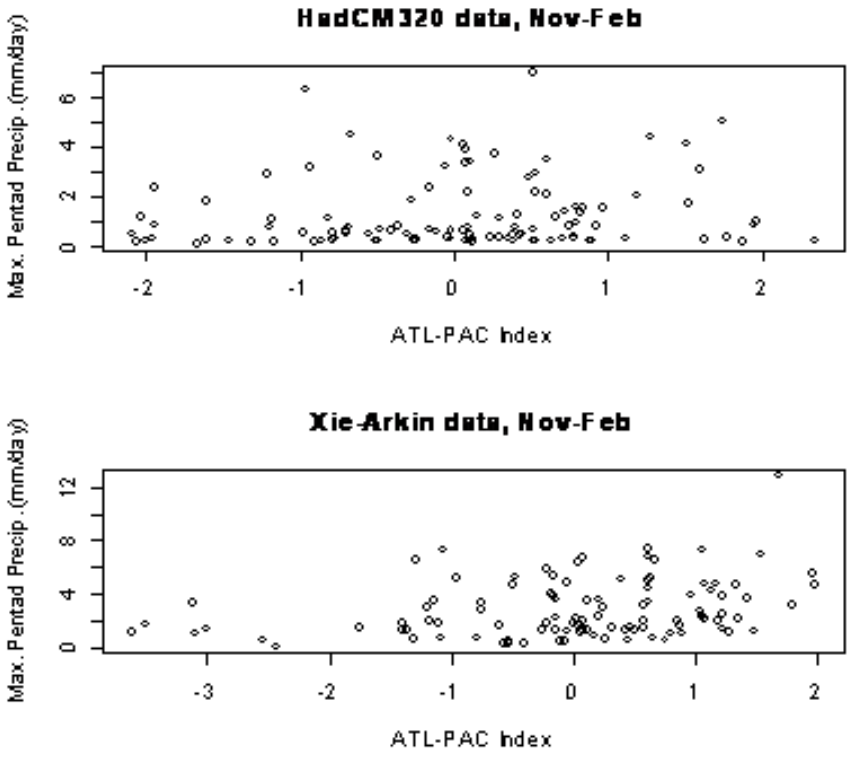

Fig. 14: Monthly maximum pentad versus the Atl-Pac Index for the HadCM3 (top), and the Xie-Arkin rainfall data set for the common period of records.

data sets as shown in Figures 15 and 16. In Figure 15 a significantly varying dependence between extreme rainfall and the API index is reflected in the 80% probability intervals for θ_{2t} (bottom) which are above the zero line, while in Figure 16 the zero line for the slope parameter (θ_{2t}) is completely included in the probability interval.

For the Echam5 model data (SRES A2 experiment), the dependence of extreme rainfall from the API index is still significant for the middle part of the $21st$ century (2046-2064), while this dependence fades off for the latest part of the century (2081-2099) (figures not shown) as shown by results from the GEV dynamic model.

With the sea surface temperature data from the Echam5 model projections we have also calculated the time varying probability of exceeding the API index value for the December 1999. A normal dynamic linear regression model was fitted to the $21st$ century API index data set estimated with the Echam5 model outputs. In this case the API index value is assumed normally distributed with location parameter μ varying with time in the sense of the state-space mode defined in equation 3. In this case the covariable used is the time index. The baseline (θ_{1t}) and slope (θ_{2t}) with their posterior 80% probability intervals are shown in Figure 17. In this case both parameters does not seem to change with time, although the baseline shows a higher uncertainty as the projection period approaches the end of the $21st$ century.

Ten thousands simulations from model 3 were obtained using the *Forward Filtering Backward Simulation* algorithm described in [7]. The simulated μ_t were used as input parameters to get samples from a normal probability distribution with time-varying mean and constant variance. Half of the samples were discarded and the remaining values were assumed as samples from the posterior predictive distribution of the dynamic normal linear regression model.

From these values the proportion of data points exceeding the December 1999 API value of approximately 1.7°C were calculated to estimate the posterior probability of exceeding the December 1999 value. These calculations were made for the NCEP data set; the Echam5 model (20CM3 experiment, 1961-1999) and the Echam5 model (SRES A2 experiment, $2046-2099$). The results are shown in Figure 18, in where the probability line of 10% has been drawn for all data sets.

From Figure 18, there is not apparent trend in the probability of exceeding the December 1999 value. In [3] they observed an increasing probability of exceeding the July-November anomalous 2005 ANSG index, which can be

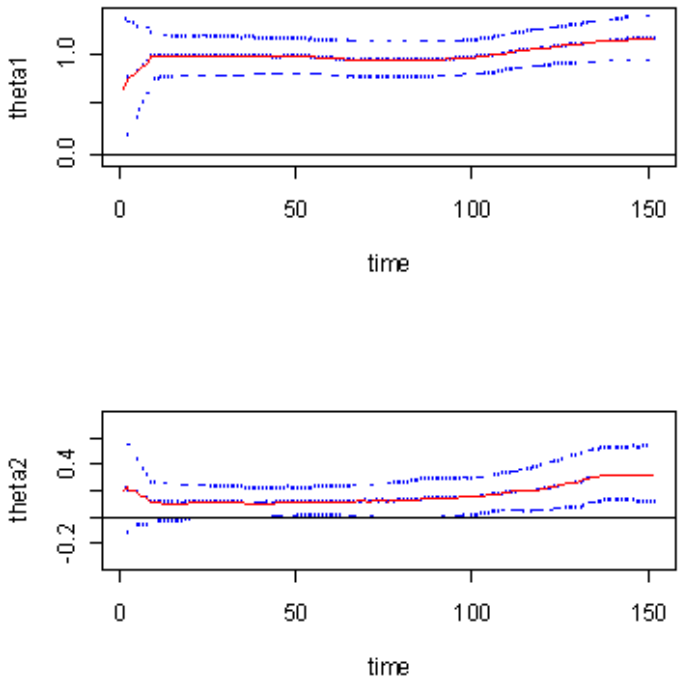

Fig. 15: Baseline (top) and slope (bottom) of the linear model relating the GEV location parameter with the API index. Values are given at each time interval (Nov-Feb months) for the period 01/01/1979 – 31/12/1999 (86 months) for the Echam5 data (20C3M experiment). Dotted lines are the 90th, 50th (median) and 10th percentiles. Continuous line corresponds to the simulated mean values.

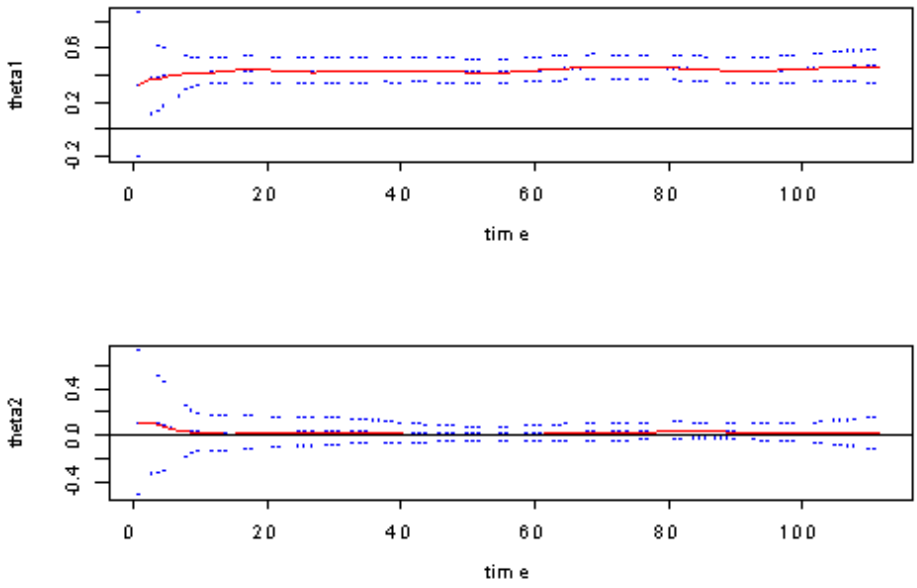

Fig. 16: Baseline (top) and slope (bottom) of the linear model relating the GEV location parameter with the API index. Values are given at each time interval (Nov-Feb months) for the period 01/01/1979 – 30/12/1989 (46 months) for the HadCM3 data (20CM3 experiment). Dotted lines are the 90th, 50th (median) and 10th percentiles. Continuous line corresponds to the simulated mean values.

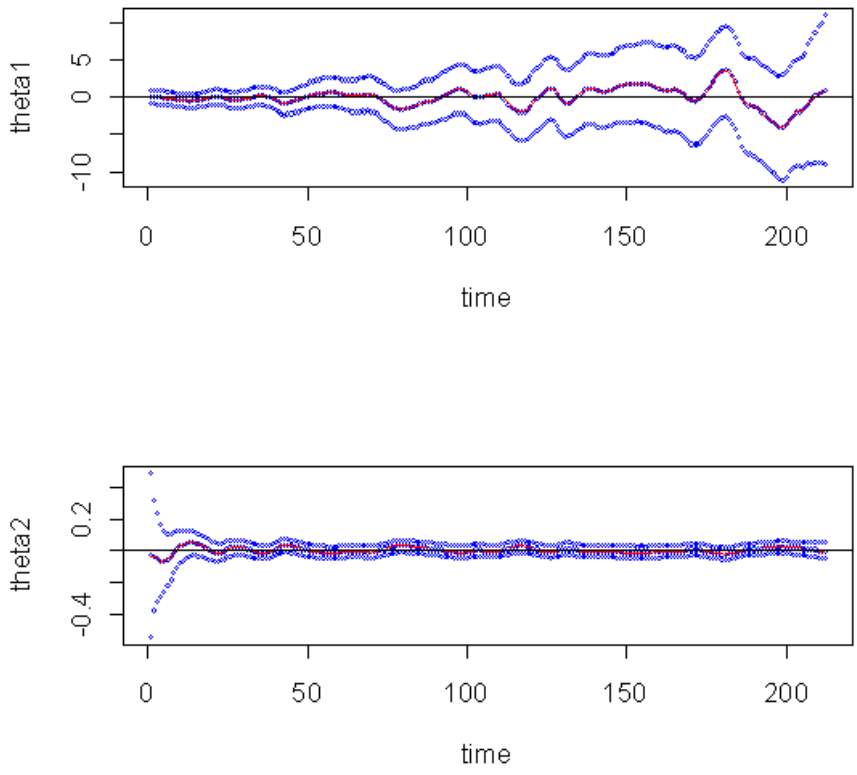

Fig. 17: Baseline (top) and slope (bottom) of the linear model relating the GEV location parameter with the API index. Values are given at each time interval (Nov-Feb months) for the period $2046 - 2099$ for the Echam5 data (SRES A2 experiment). Dotted lines are the $90th$, $50th$ (median) and $10th$ percentiles. Continuous line corresponds to the simulated mean.

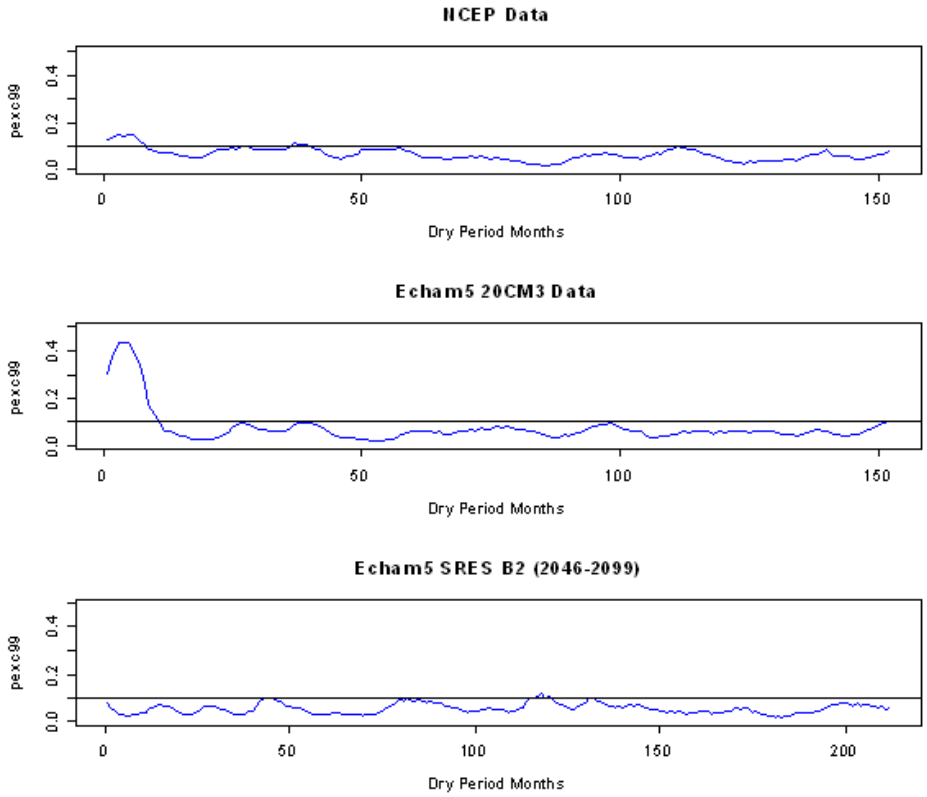

Fig. 18: Exceedence probability of the December 1999 value of the ATL-PAC index during the Dry period months (Nov-Feb) for the NCEP data set, the Echam5 20CM3 experiment data set and the Echam5 SRES B2 ($2046 - 2099$) data set.

attributed to faster warming of the Northern Atlantic with respect to the South Atlantic in the future. This behavior is not apparent in this case, and the equatorial tropical Pacific seems to be heating at the same rate as the northern tropical Atlantic (figures not shown), bringing as a consequence that there is not a significant trend in the mean values of the API index as shown in Figure 17, and anomalous values associated with a stronger rainfall response in the northern Venezuelan coast, might not be changing their frequency with time. However, the relationship between the API index and extreme rainfall as projected for future climates does not seem to have the same dependence structure as for present climate, which might limits its predictive potential for future extreme rainfall events.

4 Conclusions

By performing an extensive exploratory analysis of point rainfall data for different locations as well as using globally gridded data sets (the Xie and Arkin data set), the combined Atlantic-Pacific sea surface temperature anomalies were investigated as possible co-variables to explain the anomalous behavior of rainfall in the central coast of Venezuela during the dry period (Nov-Feb). Following the significance analysis achieved by fitting a dynamic generalized extreme value model to station based daily rainfall at different locations and to the Xie and Arkin data set for the Vargas state, the Atlantic-Pacific index (API) described as the difference between the SST anomalies of the tropical north Atlantic and the region Niño3 in the equatorial Pacific, seems to be an adequate index to explain the probabilistic nature of rainfall extremes in the northern Venezuelan coast for the months Nov- Feb. For most locations and for the Xie and Arkin data set, it was found that there is a significant dependence of the GEV distribution location parameter on the Atl-Pac Index and this dependence varies with time.

One of the aims of the analysis was also to investigate whether the relationship observed between the oceanic variables, in particular the Atlantic-Pacific index, and the extreme rainfall observed values holds for the GCMs model simulations for present and future climates. A similar dependence between the Atlantic-Pacific index and the probabilistic behavior of observed extreme rainfall is found from simulations of global models for the 20th century climate (20C3M experiment). Specifically for the Echam5 model, a significant dependence of the location parameter of the GEV model on the

Atlantic-Pacific index is well described by the GEV dynamic model for the 20C3M experiment. These results are not replicated for the HadCM3 model in where this dependence is weaker and does not resemble the behavior found with the $20th$ century observations. When looking at future climates we considered global model simulations for the period $2046 - 2099$ and the experiment SRES A2. For the Echam5 model data, the dependence of extreme rainfall from the API index is still significant for the middle part of the 21st century ($2046 - 2064$), while this dependence fades off for the latest part of the century ($2081 - 2099$) as shown by results from fitting the GEV dynamic model. The relationship between the API index and extreme rainfall as projected for future climates does not seem to have the same dependence structure as for present climate, which might limits its predictive potential for future extreme rainfall events based on the projections of this index.

The proportion of data points exceeding the December 1999 API value of approximately 1.7°C were calculated to estimate the posterior probability of exceeding the December 1999 value. From these results it can be concluded that there is not apparent trend in the probability of exceeding the December 1999 value. However since the dependence of extremes values on the API index is projected to change in future climates this conclusion should be taken with a lot of care. As suggested from this research, the combined effect of the Atlantic and Pacific SSTs may set the background conditions to trigger extreme events during the Nov-Feb dry season, however other features as the topography influence and the propagation of frontal systems coming from northern latitudes into the Caribbean, are key to understand the physical mechanisms causing these events. Synoptic scale circulation features associated with extreme rainfall events in northern Venezuela should be further investigated. Differences in global model results also require a more extensive model comparison, as for example, on how does topography in the HadCM3 compare with Echam5 in this region.

To all uncertainties discussed in this analysis one must not forget uncertainties related with the spatial resolution used in analyzing global model outputs. Vulnerability analysis should be carried out at local and regional scales, in order to provide to policy makers more precise information on how extreme events are expected to affect people lives and livelihoods. Thereof adaptation measures must be adopted without delaying timely decisions for the communities' wellbeing.

Acknowledgements: This research was carried out while the first author was a visiting researcher of the Instituto Nacional do Pesquisas Espacias (INPE) with funding support from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) Project No. 2007/57860-0 and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

References

[1] Coles S.G. *An introduction to statistical modeling of extreme values.* Springer, London, 2001.

[2] Tawn J.A. and Coles, S.G. A bayesian analysis to extreme rainfall. *Applied Statistics*, 45:463–478, 1996.

[3] Harris P.P., Huntingford C., Betts R.A., Collins M., Jones C.D., Jupp T.E., Marengo J.A., Nobre C.A. and Cox P.M. Increasing risk of amazonian drought due to decreasing aerosol pollution. *Nature*, 2008. DOI: 10.1038.

[4] Alfaro E. and Enfield D. The dependence of caribbean rainfall on the interaction of the tropical atlantic and pacific oceans. *J. Climate*, 12:2093–2103, 1999.

[5] Tedeschi R.G. and Grimm A.L. Enso and extreme rainfall events in south america. *J. Climate*, 2009.

[6] Peterson T., Abreu de Sousa J.R., Alves L.M., Ambrizzi T., Baez J., Barbosa de Brito J.I., Barros V.R., Berlato M.A., Bidegain M., Coronel G., Corradi V., Garcia VJ., Grimm A.M., Jaildodos Anjos R., Karoly D., Marengo J.A., Marino M.B., Meira P.R., Miranda G.C., Molion L., Muncunil D.F., Nechet D., Ontaneda G., Quintana J., Ramirez E., Rebello E., Rusticucci M., Santos J.L., Varillas I.T., Vincent L., Yumiko M. and Haylock M.R. Trends in total and extreme south american rainfall 1960-2000 and links with sea surface temperature. *J. Climate*, 19:1490–1512, 2006.

[7] Sansó B. and Huerta G. Time-varying models for extreme values. *Environ. Ecol. Stats.*, 14:285–299, 2007.

[8] Gregory J.M., Ingram W.J., Johnson C.E., Jones A., Lowe J.A., Mitchell J.F.B., Roberts D.L., Sexton B.M.H., Stevenson D.S., Tett S.F.B., Woodage M.J. and Johns T.C. Anthropogenic climate change for 1860 to 2100 simulated with the hadcm3 model under updated emissions scenarios. *Clim. Dyn.*, 20:583–612, 2003.

[9] Keenlyside N., Botzet M., Haak H., Luo J.J., Latif M., Marotzke J., Mikolajewicz U., Roeckner E. and Jungclaus, H. Ocean circulation and tropical variability in the coupled model echam5/mpi-om. *J. Climate*, 19:3952–3972, 2006.

[10] Lyon B. Enhanced seasonal rainfall in northern venezuela and the extreme events of december 1999. *J. Climate*, 16:2302–2306, 2003.

[11] Bustamante J.F., Polanco M.I. and Marengo, J.A. Availação dos eventos de chuva no norte da venezuela de 13-17 de dezembro de 1999. estudos observacionais. *Revista Brasileira de Metorología*, 11(1):43–60, 2003.

[12] Jones R., Alves L.N., Valverde M.C. and Marengo, J.A. Future change of temperature and precipitation extremes in south america as derived from the precis regional climate modeling system. *Int. J. Climatol.*, 2009. DOI: 10.1002/joc.1863.

[13] Nobre C.A., Tomasella J., Sampaio G., De Oliveira R., Camargo H., Alves L.M., Brown I.F. and Marengo, J.A. The drought of amazonia 2005. *J. Climate*, 2008. DOI: 10.1175/2007JCLI1600.1.

[14] Pachauri R.K. and A.(eds.) Reisinger. Climate change 2007: Synthesis report. contribution of working groups i, ii and iii to the fourth assessment report of the intergovernmental panel on climate change. Technical report, IPCC, 2007. Geneva, Switzerland, 104 pp.

[15] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008. ISBN 3-900051-07-0.

[16] Bell M.A. and Ropelewski C.F. Shifts in the statistics of daily rainfall in south america conditional on enso phase. *J. Climate*, 21:849–865, 2008.

[17] Enfield D.B., Chen A.A. and Taylor, M.A. Influence of the tropical atlantic versus the tropical pacific on the caribbean rainfall. *J. Geophys. Res.*, 107(C9):10–1–10–14, 2002.

[18] Harrinson J. and West M. *Bayesian forecasting and dynamic models. 2nd edn.* Springer Verlag, New York, 1997.

[19] Larsen M.C., Eaton L.S., Morgan B.A., Blair J.L. and Wieczorek, G.F. Debris-flow and flooding hazards associated with the december 1999 storm in coastal venezuela and strategies for mitigation. Technical report, U.S. Geological Survey, 2001. Open File Rep. OFR-01-144, 40 pp.

[20] He F., Liu Z., Li Ch. and Wu L. Atmospheric teleconnections of tropical atlantic variability: Interhemispheric, tropical extratropical, and cross basin interactions. *J. Climate*, 20:856–870, 2007.

[21] Arkin P.P. and Xie P. Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.*, 78:2539–2558, 1997.