Drawing the complexity of Colombian climate from non-extensive extreme behavior

Isabel Hoyos 1,* and Boris A. Rodríguez 1

1 Instituto de Física, Grupo de Fundamentos y Enseñanza de la Física y los Sistemas Dinámicos. Universidad de Antioquia. Medellín, Colombia
* Correspondence: isabel.hoyos@udea.edu.co; Tel.: +57 4 219 5630

Abstract: We evaluate the complexity of Colombian climate from extreme behavior of gauge temperature and precipitation, using the novel Tsallis' non-extensive entropy principle based on physical information through the q-index. We find the spatial structure of non additive universal categories (q-index) and compare with some complex systems with the potential to have some degree of dynamical affinity. Our results evidence the great dynamical variability of regional climate expressed in the large range of values of q-index, and the high degree of non-extensitivity for both temperature and precipitation.

Keywords: Colombian climate complexity, climate extremes, Tsallis' non-extensive statistical mechanics, universal categories.

1. Introduction

From a dynamic point of view, climate extremes are critical phenomena emerging from spatio-temporal multi-scale interactions. Long-term memory, high degree of information content and persistent positive feedback are some of the necessary conditions to drive the system far from equilibrium and exhibit extreme behavior [1–3]. As part of the Earth system, climate exhibits inherent complexity [4–6] and it is expected a better performance in modeling and predictability from non-Gaussian stochastic approach, via entropic characteristics [7–9].

The Tsallis' non-extensive statistical mechanics is a modern extended theory built on the fact that physical states in phase space from complex processes correspond to a more general entropy than the established by the classical Gaussian thermodynamical equilibrium, since ergodicity (and statistical equilibrium as its macroscopic manifestation) is just one of the dynamic possibilities of microscopic mixing in complex systems [10–16]. In practical manners, the generalization from Tsallis' theory introduces a non-extensive entropic functional through the index $q$ which identifies non additive universal categories and provides physically based information about the underlying dynamics, revealing crucial features about spatio-temporal long-range correlations emerging in extremes [11,15,17,18].

The Tsallis' theory is being progressively applied in complex systems. In particular relative to geophysical processes: turbulence [19], estuarine hydrodynamics [20], ozone layer [21], earthquakes [22], geopotential height [23], global climate [24], ENSO [25,26], hydrological extremes [27,28], regional climate [29–31], among others. In agreement with [17], the success of Tsallis’ theory in representing complex systems is mostly due to the extension of the physical representation of the underlying universal organizing principle through a non-extensive entropy formulation that provides a measure of dynamical organization (or information content).

In this contribution, we focus on Colombian climate complexity description from a non-extensive statistical formulation. As previously discussed in [32], [33], [34] and [35], Colombia is located over
a very active region in terms of atmospheric moisture transport across Americas. The long-term
hydrologic teleconnections, the great intra-annual variability of the atmospheric moisture contributions
and the heterogeneity of orographic interactions, are some of the factors that have been identified as
regional sources of climate variability and extreme events. However, there is still great uncertainty
related to underlying dynamics and there is not a previous measure of regional climate complexity
based on observations. We present the spatial distribution of \( q \) index for temperature and precipitation
extreme behavior, which can be interpreted as a complexity property of regional climate. This findings
provide useful information about the nature of local physical processes for regional climate modeling.

This paper is organized as follows: Section 2 gives a general description of study area and its
climatic and extreme features. Section 3 describes the statistical model applied. Results are shown in
Section 4. Finally, concluding remarks are presented in section 5.

2. Study area and data

This research is focused on Colombia (Fig 1). This country is characterized by a great landform
heterogeneity due to the splitting of the Andes mountain chain in three branches. As a result, the
country exhibits a mixed landscape that includes snow peaks, highland plateaus, deep canyons, large
rainforest areas and wide valleys, among others. Accordingly, a great ecosystem variability and a large
biodiversity are also a footprint for the Colombian terrain. Further, the country is surrounded by the
basins of rivers Amazon and Orinoco, and the Pacific ocean and Caribbean sea. A complex interplay
between these geographic particularities and regional circulation is responsible for a large variability
of rainfall patterns [32,33,35–37], compromises the atmospheric transport across Americas [32–34,38]
and leaves a great regional sensibility to global climate phenomena [32,34,39–47].

The four Colombian catchment basins Caribbean, Pacific, Orinoco and Amazon have been
becoming in classical units of hydrological and climatic analysis since each one has the proper size
for ensuring the closure of the water balance and also, annual cycle of temperature an precipitation
have qualitative similarities across each one [48–51]. In the other hand, these hydrological units have
had differentiated development paths. While the inter-Andean region has been the central pole of
the country’s development, the remaining areas have been marginalized in terms of social, cultural
and economic growth. This situation is portrayed in the national weather network. Distribution,
availability and quality of gauge stations are limited. The best sampled region in Colombia
corresponds to Caribbean Colombian catchment (especially in the inter-Andean region) and southern
Colombian Pacific catchment.

The study of extreme behavior of climate requires large enough time series for statistical analysis,
that is why we have fitted our target region to the Caribbean and Pacific Colombian catchment basins.
The Amazon and Orinoco catchment basins are not considered in this work because their limited
observational time series despite their environmental and eco-hydrological regional importance.
This lack of information is often overcome by using data in the state of art, as reconstructed fields,
reanalysis and model output. However, these datasets have limitations in the representation of
Colombian climate (especially referred to extremes), mostly because of the available spatial resolution
is still being insufficient to represent the great land-form variability and local climate processes
[33,48,52,53]. In this sense, we rather extract information from gauge observations to provide dynamic
clues that can help to determine which model configuration is better for a more realistic representation
of the Colombian climate.
Figure 1. Study area. Points correspond to locations of gauge stations across Caribbean and Pacific Colombian catchments. Orography from Global Land One-kilometer Base Elevation GLOBE [54].

3. Statistical model

As a generalization of the Boltzmann-Gibbs thermodynamic entropy ($S_{BG}$), Tsallis [55] proposed an entropy function $S_q$ that better describes complex systems whose phase space is not ergodically visited and therefore the extensive property of thermodynamic entropy is violated. For instance, in systems regulated by several spatio-temporal scales as climate. This new entropy function has the generic form:

$$S_q = k_B \frac{1}{q-1} \left[ 1 - \int_{\Omega} [f(x)]^q dx \right]; \quad q \in \mathbb{R}, \quad (1)$$

where $k_B$ is the Boltzmann’s universal constant, $\Omega$ is the state space represented by the variable $x \in \mathbb{R}^n$, and $f(x)$ is the probability density function. $q$ is an entropic index that characterizes the universality classes of non-additivity [14] and describes the deviation of Tsallis entropy from the standard Boltzmann-Gibbs entropy, more precisely, the emergence of long-range interactions, long-term memory and/or multi-fractal behavior [2]. From this new definition of entropy and applying proper constraints, a set of generalized distribution functions are obtained through the maximum entropy principle. In the limit case $q \to 1$, Tsallis’ distributions converge to classical distributions ($S_q \to S_{BG}$) [56].

The normalized $q$-exponential distribution satisfies the maximum entropy principle for $S_q$ with constant mean as constraint, which is defined in [57] as:

$$f_{q,\beta}(x) = \frac{1}{Z_{q,\beta}} e_q(-\beta x), \quad (2)$$
where $Z_{q,\beta} = 1/(\beta(2-q))$ is the partition function with $0 < q < 2$ and $e_q(-\beta x) = (1-\beta(1-q)x)^{1/(1-q)}$ is called the q-exponential function. $\beta$ is a positive scale parameter associated to the distribution mean $\mu$ through $\beta = 1/(\mu(3-2q))$ for $q < 3/2$. If $q < 1$, the q-exponential function has an upper boundary in $x = 1/(\beta(1-q))$ and is unbounded if $1 < q < 2$.

Supposing that the state of the system is described by the random variable $X$ whose excesses $Z = X(t) - u \mid X > u$ are defined over a enough high threshold $u$ so that $Z$ represents the tail distribution of $X$. In agreement with [57], if $X$ follows a q-exponential probability density $f_{q,\beta}(x)$ its excesses remain having a q-exponential distribution $f_{q',\beta'}(z)$, where:

$$q' = q \quad \text{and} \quad \beta' = \frac{\beta}{1-\beta(1-q)u}.$$  \hspace{1cm} (3)

The q-exponential distribution is particularly interesting because it provides the information of q-index from the excesses set. The definition of climate extremes as the set of excess over a high threshold allows an statistical description in terms of a Generalized Pareto (GP) distribution via asymptotic limit for heavy tails (Fréchet domain) in accordance with [58]:

$$f_{\sigma,\zeta}(z) = \frac{1}{\sigma} \left(1 - \frac{\zeta z}{\sigma}\right)^{1/\zeta - 1},$$  \hspace{1cm} (4)

$\sigma$ and $\zeta$ are the GP distribution parameters. $\sigma$ is a positive scale parameter that gives information about the variability and central value of excesses. $\zeta$ is a dimensionless shape parameter, which is referred to the shape and bound of the distribution. If $\zeta > 0$, the GP distribution has an upper boundary in $z = \sigma/\zeta$. If $\zeta \leq 0$, the distribution is unbounded. The GP distribution becomes to the exponential distribution when $\zeta \to 0$ and the uniform distribution when $\zeta \to 1$.

As both $f_{q',\beta'}(z)$ and $f_{\sigma,\zeta}(z)$ belong to Fréchet domain, they are directly linked through the relations:

$$q' = \frac{2\zeta - 1}{\zeta} = q \quad \text{and} \quad \beta' = \frac{1-\zeta}{\sigma}.$$  \hspace{1cm} (5)

This procedure permits the estimation of non-extensive parameters from the extreme behavior of the system.

4. Results

Figs 2 and 3 present the spatial layout of the q-exponential distribution function parameters for monthly gauge temperature and precipitation excesses over a non-stationary threshold, within the period 1930-2009. This data were calculated from the GP distribution parameters reported in [41]. Here we focus on excess over 90th percentile as a good agreement between the extreme behavior representation and enough sample length for statistical purpose.

The estimated parameter $q$ shows large spatial variability. In the entire region $0 < q' < 3/2$ for both temperature and precipitation, laying in the range of expected value of excesses $\mu'$ defined by the scale parameter $\beta'$ as $\mu' = 1/(\beta'(3-2q'))$. However, qualitatively differences were found in the Caribbean and Pacific Colombian basins.

For temperature (Figs 2 and 4a), values of q-index in the Caribbean basin area range from 0.22 to 1.31. 59% (39%) of gauge stations have $q' < 1$ ($q' > 1$) that define bounded (unbounded) q-exponential distributions. Only 2% of sampling data are in the canonical exponential function limit ($q' = 1$), evidencing a high degree of non-extensivity in regional temperature. In contrast, the Pacific basin has q-index values in a narrower interval, ranging from 0.75 to 1.15. Bounded and unbounded q-exponential
distributions are in equal proportion and neither sampling point with \( q' = 1 \).

\[
\begin{align*}
q' & \quad \mu' \ (°C)
\end{align*}
\]

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2}
\caption{Parameters for the \( q \)-exponential distribution function for temperature excesses over non-stationary 90th percentile. (a) \( q \)-index. (b) expected value for excesses set. The inset boxes show minimum, maximum and the three first quartiles of parameters for Caribbean and Pacific Colombian basins.}
\end{figure}

For precipitation (Figs 3 and 4c), the study area covers \( q \)-index values from 0.49 to 1.46. In the Caribbean basin ranging from 0.49 to 1.41 with 25% of stations with values < 1, while in the Pacific basin \( q \)-index ranges from 0.72 to 1.46, with 15% of bounded distributions. In both basins, the percentage of stations with \( q' = 1 \) is near to 8%, remaining a significant non-extensive character of regional precipitation.

Figure 4b compiles previous results for some systems with the potential to have some degree of dynamical affinity with our studied system. Here, we refer to dynamical affinity in terms of universality concept as coincidence or similarity in the \( q \)-index value since the connections we are looking for overtake the particular details of any specific mechanism and revel a kind of order in real-world systems. Regional temperature and precipitation cover the range of \( q \)-index reported for last glacial global climate temperature [24], the ENSO [25,26], Ozone layer dynamics [21], rainfall extremes [27] and Couette-Taylor turbulence [17]. This is a benchmark for values of \( q \)-index in typical climate-related complex systems and establishes a context that allows us to highlight two primary aspects: i) the great dynamical variability of regional climate expressed in the large range of values of \( q \)-index, and ii) the high degree of non-extensivity for both temperature and precipitation. These results are in agreement with previously obtained in [31] for daily precipitation in a smaller region and shorter period in tropical Andes.
Fig 5 compares $q$-index for stations with available information of both temperature and precipitation. The regional extreme behavior of temperature and precipitation is frequently related to global climate phenomena as ENSO [32,39,41,44,47,59]. In this sense, similar $q$-index values could be expected for both variables, however we find that kind of coincidence for just a few stations. The common case is same location has significant different $q$-index values for temperature and precipitation. This result means that dynamics related to global phenomena is expressed in a different kind of complexity for both variables, mostly because of local processes are influencing in a strong manner how the long-term phenomena are locally expressed in each variable. This result is in agreement with obtained in [32] through information transference theory.
5. Conclusions

In this contribution we evidence the non-extensive property of regional climate from the extreme behavior. A clear signature of this non-extensive character is the great portion of gauge stations with $q' \neq 1$ compared to stations with $q' = 1$, for both temperature and precipitation. This implies that regional climate dynamics is not deterministic with a (multi-)fractal phase space involving several spatio-temporal scales, where the interplay of long-term global phenomena with short-term local processes explains the great intra-regional variability. On practical matters, the q-index provides a unique physically based identifier of statistical complexity.

Tsallis’ distributions provide valuable clues about dynamical affinity and the q-index can be used as an useful tool to compare model output (or any other dataset) with observations and define a criteria of goodness climate representation which encompasses several range of interactions and processes.

This statistical approach could be a useful forecast tool risk management in a climate change context since the differences of q-index for temperature and precipitation in the same location evidence intrinsic properties of how the extreme dynamics is expressed in each variable. The bounded/unbounded character of Tsallis’ q-exponential distribution shape the nature of extreme behavior, in particular, information of return periods, range and probability for extreme events of interest can be easily obtained for estimated parameters.

Author Contributions: Conceptualization, Isabel Hoyos and Boris A. Rodríguez; Formal analysis, Isabel Hoyos; Funding acquisition, Boris A. Rodríguez; Methodology, Boris A. Rodríguez; Software, Isabel Hoyos; Visualization, Isabel Hoyos; Writing – review & editing, Isabel Hoyos and Boris A. Rodríguez.

Conflicts of Interest: “The authors declare no conflict of interest.”


49. IDEAM. *Atlas climatológico de Colombia*; IDEAM, 2005.


59. Poveda, G.; Álvarez, D.M.; Rueda, Ó.A. Hydro-climatic variability over the Andes of Colombia associated with ENSO: a review of climatic processes and their impact on one of the Earth’s most important biodiversity hotspots. *Climate Dynamics* 2011, 36, 2233–2249.