UNIVERSIDAD DE GUADALAJARA

CENTRO UNIVERSITARIO DE CIENCIAS EXACTAS E INGENIERÍAS

DIVISIÓN DE CIENCIAS BÁSICAS

“Analysis of climate models with complex systems tools: assessing the possible impacts of global warming on tropical cyclone activity”

T E S I S

Q U E P A R A O B T E N E R E L T Í T U L O D E

F Í S I C O

P R E S E N T A

ERICK ALEJANDRO MADRIGAL SOLÍS

DIRECTOR: DR. GERARDO GARCÍA NAUMIS

ASESOR: MTRO. ELIO ROCA FLORES

Guadalajara, Jalisco. Marzo de 2023

UNIVERSIDAD DE GUADALAJARA

CENTRO UNIVERSITARIO DE CIENCIAS EXACTAS E INGENIERÍAS

DIVISIÓN DE CIENCIAS BÁSICAS

"Analysis of climate models with complex systems tools: assessing the possible impacts of global warming on tropical cyclone activity"

T E S I S

Q U E P A R A O B T E N E R E L T Í T U L O D E

F Í S I C O

P R E S E N T A

ERICK ALEJANDRO MADRIGAL SOLÍS

DIRECTOR: DR. GERARDO GARCÍA NAUMIS

ASESOR: MTRO. ELIO ROCA FLORES

Guadalajara, Jalisco. Marzo de 2023

CUCEI/SAC/CDFIS/030/2022
Código 212580894

C. Erick Alejandro Madrigal Solis
Egresado de la carrera de Licenciado en Física
Presente

Hacemos de su conocimiento el resultado en el Dictamen emitido por el Comité de Titulación de la Licenciatura en Física, con relación a su solicitud de aprobación de modalidad y opción de titulación, conforme al Reglamento General de Titulación de la Universidad de Guadalajara:

Artículo 14, **Tesis, Tesina e Informes**
Opción I, **Tesis**
Con el título: **Analysis of climate models with complex systems tools: assessing the possible impacts of global warming on tropical cyclone activity**

El Comité emite el siguiente resolutivo:

PROPUESTA APROBADA

Quedando asentada en el acta de la sesión con fecha 20 de septiembre de 2022, con el folio 61, que éste Comité designa al Dr. Federico Angel Velázquez Muñoz como Presidente del jurado para realizar la ceremonia. Asimismo, se le otorga el plazo de un año a partir de la fecha de aprobación para concluir su proceso de titulación.

El presente Dictamen deberá aparecer en el trabajo de titulación antes mencionado.

ATENTAMENTE
"Piensa y Trabaja"
"2022, Guadalajara, hogar de la Feria Internacional del Libro y Capital Mundial del Libro"
Guadalajara, Jal. a miércoles 16 de noviembre de 2022

Dra. Gloria Arlette Méndez Maldonado
Presidente del Comité de Titulación

COMITÉ DE TITULACIÓN
LICENCIATURA EN FÍSICA

GAMM/egg*
C.c.p. Archivo.

CUCEI/SAC/CDFIS/009/2023

Comprobante Académico

El Comité de Titulación de la carrera de **Licenciado en Física**, hace constar que:

Erick Alejandro Madrigal Solis
Código 212580894

Egresado del plan semestral modular, ha cumplido con los requisitos académicos para obtener el grado de Físico y acreditó el dominio de lecto comprensión del idioma inglés, correspondiente al nivel A2 del Marco Común Europeo de referencia para las lenguas, o su equivalente, como lo marca el resolutivo Décimo Segundo del Dictamen I/2012/387. Se adjunta documento original equivalente al NIVEL B1 que le da derecho a solicitar ceremonia de titulación. La modalidad y opción que le han sido aprobadas, se indican en la siguiente tabla:

Artículo 9. Desempeño Académico Sobresaliente		I. Excelencia Académica (**calif**)
		II. Titulación por Promedio (**calif**)
Artículo 10. Exámenes		I. Examen Global Teórico-Práctico (**calif**)
		III. Examen General de Certificación Profesional (**EGEL-calif**)
		IV. Examen de Capacitación Profesional (**calif**)
Artículo 11. Producción de Materiales Educativos		I. Guías Comentadas o Ilustradas
		II. Paquete Didáctico
		III. Propuesta Pedagógica
Artículo 12. Investigación y Estudios de Posgrado		I.Cursos de Maestría o Doctorado en IES de reconocido prestigio
		III. Seminario de Investigación
		IV. Seminario de Titulación
		V. Diseño o Rediseño de Equipos, Aparatos, Maquinaria, Proceso o Sistema de computación y/o Informática
Artículo 14. Tesis, Tesina e Informes	**x**	I. Tesis
		II.Tesina
		III. Informe de Prácticas Profesionales

El Comité de Titulación ha designado al Dr. Diego Armando Pantoja González, Dr. Federico Angel Velázquez Muñoz y Dra. Iryna Tereschenko como Presidente, Secretario y vocal del jurado, respectivamente, para realizar la ceremonia.

ATENTAMENTE
"Piensa y Trabaja"
"2023, Año del fomento a la formación Integral con una Red de Centros y Sistemas Multitemáticos"
Guadalajara, Jal. a viernes 17 de marzo de 2023

Dra. Gloria Arlette Méndez Maldonado
Presidente del Comité de Titulación

COORDINACIÓN DE PROGRAMA DOCENTE EN LICENCIATURA EN FÍSICA
C.U.C.E.I.

GAMM/egg.**
C.c.p. Archivo.

Comité de Titulación
Licenciatura en Física
Centro Universitario de Ciencias Exactas e Ingenierías
Universidad de Guadalajara

Estimados Miembros del Comité

Por medio de la presente hago de su conocimiento que, después de revisar la tesis profesional del pasante de la Licenciatura en Física C. ERICK ALEJANDRO MADRIGAL SOLÍS, que lleva por título "ANALYSIS OF CLIMATE MODELS WITH COMPLEX SYSTEMS TOOLS: ASSESSING THE POSSIBLE IMPACTS OF GLOBAL WARMING ON TROPICAL CYCLONE ACTIVITY", considero que la misma cumple con los objetivos presentados en el Art. 13 del reglamento de titulación, así como las formas exigidas por tan respetable Comité, por lo que otorgo mi autorización para su impresión y defensa.

A T E N T A M E N T E

DR. GERARDO GARCÍA NAUMIS
DIRECTOR DE TRABAJO DE TITULACIÓN

Agradecimientos

Quiero expresar mi profundo agradecimiento y admiración a mi director de tesis, Dr. Gerardo García Naumis y a mi asesor, Mtro. Elio Roca Flores, quienes hicieron que disfrutara este proceso y que estuviera lleno de aprendizajes. Me encuentro muy agradecido con Gerardo por brindarme la oportunidad de ser parte del grupo de investigación y de hacerme sentir como un elemento importante del equipo. Asimismo, me encuentro en deuda con Elio por su asesoría y la dedicación con la que me ayudó a lo largo del camino. También le agradezco por el apoyo moral que me brindó en cada etapa del proceso de investigación y redacción del presente trabajo.

También quiero mencionar que, gracias al apoyo económico brindado por el CONACyT en la modalidad de ayudante de investigador nacional nivel III o emérito, tuve la oportunidad de dedicarme de lleno a este trabajo de investigación.

A mis profesores y mi universidad, la Universidad de Guadalajara, les debo la oportunidad de haberme formado profesionalmente y de haber despertado mi curiosidad por el aprendizaje y la investigación.

Por último, y tal vez más importante que nada, le agradezco a mi familia y mis seres queridos por motivarme a seguir adelante y brindarme coraje para seguir mis sueños y buscar mi lugar en el mundo.

Erick Madrigal

Abstract

Tropical cyclones are one of the most fascinating, organized atmospheric systems. They are, together with the climate, a benchmark of complex systems. This means that they are composed of many subsystems and every element that forms the whole interacts with each other in different ways. The aim of this thesis was to analyze annual tropical cyclone activity in the Northwest Pacific using the accumulated cyclone energy index. Methods stemming both from statistics and nonlinear dynamics were used to perform analyses of time series coming from chaotic systems and stochastic processes as well as observational data on the accumulated cyclone energy of the Northwest Pacific basin. More specifically, the statistical tools used were: probability density function, autocorrelation function and rank-ordering representation. The nonlinear approach studied the Hurst exponent, reconstructed the phase space and estimated the fractal dimension of the underlying attractor. The main results were that the data seem to have two dynamical states and that they present antipersistent behavior similar to a stochastic process. The tools used to analyze the observational data delivered inhomogeneous results. This may be due to the length of the data sets as statistical and nonlinear dynamics tools usually require large amounts of data points to deliver reliable results. Even with the limitations introduced by the length of the data set, the conclusions that can be reached provide an insight into the general state of tropical cyclone activity. A natural progression of this work is to extend these analyses to different basins and, ultimately, to global tropical cyclone activity.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The climate system is extremely complicated as it is composed of a myriad of elements and it functions over multiple spatial and temporal scales. Many interconnected subsystems make up the climate system, among which are the atmosphere, hydrosphere, cryosphere and others. These subsystems, in turn, interact very strongly with each other, presenting new characteristics at different scales [1]. The climate, along with its subsystems, is driven by the action of external agents, such as solar radiation and the rotation of the Earth [2]. The number of components, together with their interactions and external forcings, make the study and understanding of the climate and its phenomena a challenging task [3], [4].

In addition to the scientific understanding of the climate system, there is a variety of phenomena that make this study a pressing issue. Tropical cyclones are one of such phenomena. They are one of the most devastating natural phenomena on the planet, costing lives and wreaking havoc[5], [6]. A further pressing issue, which makes it yet more important to understand the climate system, is global warming. The exact effects that anthropogenic activity and emissions have or will have on the climate system and its associated phenomena are not yet known [7], [8]. In particular, tropical cyclone intensities are expected to increase globally [9], as well as the proportion of tropical cyclones that reach very intense levels [10], [11]. Additionally, hazards that are connected to tropical cyclone activity are projected to become more severe, such as sea level rise, tropical cyclone rainfall rates, among others [12].

The vastness of components mentioned earlier, together with their interactions and novel behavior, mean that the climate system is not only complicated, it is *complex* [1], [13]–[15]. These reasons make it clear that a new approach to tackle the study of such a complex system as the climate is needed. This new way of thinking is called complexity science [16], [17].

The standard approach to study the climate system and its associated phenomena has been to deal with it as a complicated system, rather than a complex one [18]. Traditionally, the aim of studying the climate system has been to predict its state in a future time. This is carried out with the use of numerical models and statistical analyses [12]. The first method is based on models that reproduce the features observed in experimental settings. They are complicated models that use thousands of variables and aim at predicting the general state of the climate system given a list of initial conditions [3]. The second approach to predict the future state of the system is done performing statistical analyses on historical data. This method does not seek a deeper understanding of the dynamics of the system. Instead, it aims at predicting future behavior based on present and past observations [19]. Approaching the study of the climate system through the lens of complexity science offers the benefits of using dynamical models to study the most elementary features of the system while considering the big picture and taking the multiple interactions into account [18].

The aim of this project was to characterize the state of the annual tropical cyclone activity in the Northwest Pacific basin using tools that have been recently used and developed to study complex systems. A parallel aim of this project was to assess the performance of these tools on the study of short, univariate, time series, both on synthetic data obtained from chaotic and stochastic models, and observational data retrieved from standardized databases.

The results of this research, along with previous work, were presented at the National Physics Conference in Zacatecas, Mexico [20]. Additionally, a paper has been published in an indexed journal: Roca-Flores, E., Naumis, G. G., Madrigal-Solis, E., & Fraedrich, K. (2022). Typhoon complexity: Northwest Pacific tropical cyclone season complex systems analysis. *International*

Journal of Modern Physics C. The cover page is reproduced in App. (A). An additional research paper has been accepted for publication and is currently under publishing process: Roca-Flores, E., Naumis, G. G., Madrigal-Solis, E., Fraedrich, K., & Torres, E. (2023). Hurricane season complexity: the case of North-Altantic tropical cyclones.

This text is organized as follows: the following sections give background theory to tackle the research work, namely the theory on tropical cyclones and complex systems. The next chapter (Ch. 2) presents the tools used to study the time series used for this research work. Chapter 3 covers the results on the observational data obtained with the use of dynamical and statistical tools and gives a brief explanation of them. Chapter 4 goes over the results again and draws interpretations from them using the approach of complex systems and puts them into context. Finally, Ch. 5 gives insights on the meaning of the results. This chapter reflects on the process and significance of the results and gives possible directions for further study and development.

1.1 Tropical Cyclones

Tropical cyclones are one of the most important phenomena for both mankind and the climate. On the one hand, they account for countless deaths [21] and enormous economic losses [22]. On the other hand, they play a major role in the dynamics of the ocean and atmosphere, effectively regulating global climate [23]–[25]. Thus, studying the dynamics of tropical cyclones can help improve forecasts to warn the population under threat, as well as gain a deeper understanding of their role in the general state of the climate [2], [5].

Tropical cyclones are rapidly rotating storm systems that form over tropical oceans. They are characterized by a low-pressure center and a spiral arrangement of thunderstorms [26]. They are normally between 200 and 500 km in diameter but there have been instances when they have reached a diameter of more than 1000 km [27]. These storms are always originated by pre-existing disturbances in the atmosphere [28]. Given such a storm-generating disturbance, tropical cyclones will form when the warm water in the tropical sea evaporates and then condenses into clouds. As the hot air rises, a low-pressure area is formed and the surrounding air flows inward. Because of the Coriolis effect, this uprising air starts to rotate (counterclockwise in the Northern Hemisphere) until a well-defined system emerges [5].

Following the classification of the National Weather Service (NWS), tropical cyclones, in their formative stage, will have maximum wind speeds close to 17 ms^{-1}. In this stage, they are called tropical depressions. When their wind speeds reach between 18 and 32 ms^{-1}, they are called tropical storms and they are assigned a name. Tropical cyclones with wind speeds exceeding 33 ms^{-1} are referred to as severe tropical cyclones.

The following subsections give an account of the various characteristics of Tropical Cyclones that are important to describe them. These include the description of tropical cyclone activity over large time spans (Subsec. 1.1.1), their description in terms of energy budget (Subsec. 1.1.2) and the conditions needed to spawn them (Subsec. 1.1.3). Finally, the variable used to study and quantify the intensity and energy used by tropical cyclones is presented in Subsec. 1.1.4.

1.1.1 Climatology

Although the term tropical cyclone encompasses every cyclone that originates over tropical oceans, these systems receive different names depending on the region. Strong tropical cyclones, for instance, will be known as Hurricanes in the Atlantic ocean, cyclones in the Indian Ocean and Typhoons in the Western Pacific [29].

Tropical Cyclones have been found to form and move over defined regions. These regions are called basins and they have been exactly defined to study tropical cyclones more closely [30]. These basins are shown in Fig. (1.1) and are as follows: North Indian (NI), Northwest Pacific (WP), Northeast Pacific (EP), North Atlantic (NA), South Indian (SI), South Pacific (SP) and South Atlantic (SA).

Around 80 tropical cyclones form globally every year. This quantity has remained approximately constant since reliable records began [26], [31]. Figure (1.2) shows the total number of global tropical cyclones recorded from 1980 to the year 2021. Of these 80 tropical cyclones, the

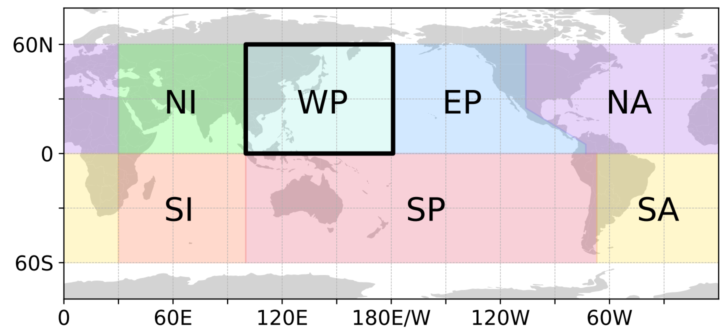


Figure 1.1: Tropical cyclone basins (From upper-left to right): North Indian (NI), Northwest Pacific (WP), Northeast Pacific (EP), North Atlantic (NA), South Indian (SI), South Pacific (SP) and South Atlantic (SA).

northern hemisphere experiences 70 %, in contrast with 30 % in the southern hemisphere. The most active single basin is the Northwest Pacific, which accounts for 31 % of the total annual tropical cyclone counts. This makes it the ideal subject of study to understand global tropical cyclone activity. This basin is followed by the Northeast Pacific (19 %) and the North Atlantic (16 %).

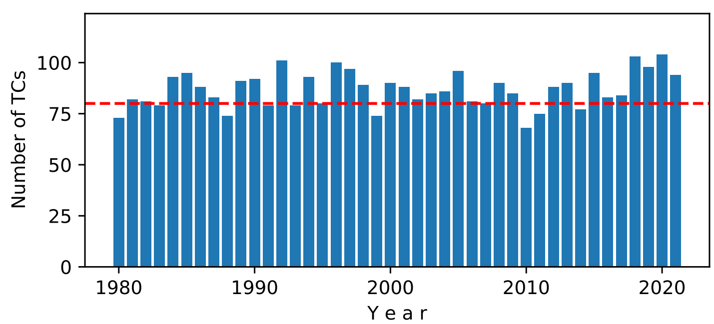


Figure 1.2: Bar plot of annual global tropical cyclone counts (data taken from [30]).

Since the start of reliable global records, no discernible trend in the global count of tropical cyclones has been identified [31]. Nevertheless, single tropical cyclone basins have seen significant changes in annual tropical cyclone counts. The North Atlantic basin, for instance, has had a significant, upward trend [32]–[34]. The Northwest Pacific, in turn, has seen a downward trend that coincides with the observed increase in the North Atlantic [31], [32], [34]. The inter-annual variability of tropical cyclone counts in the different basins is driven by natural modes of variability, with El Niño Southern Oscillation (ENSO) having the biggest influence across multiple basins [35]–[37].

1.1.2 Energy budget

Although the total number of tropical cyclones that form globally or in different basins in a given year can give us information on the general state of tropical cyclone activity, this information does not shed light on the state of the energy budget of the system. Since what drives the formation and life cycle of tropical cyclones is the available energy from the warm waters, which in turn receive it from the incoming radiation from the sun and other sources, it becomes natural to define a way to describe tropical cyclone activity in terms of the energetics of the system.

As first noted by Emanuel [38], the energy cycle of mature tropical cyclones can be idealized as that of a Carnot heat engine. In the case of tropical cyclones, the working fluid is a combination of dry air, water vapor and condensed water. The power source of this cycle is the heat transfer that results from thermodynamic disequilibrium between the sea and atmosphere. This heat transfer is done mainly through water evaporation.

Figure (1.3) shows a depiction of the thermodynamic cycle of tropical cyclones. The cycle starts at point A, at around 100 km from the eyewall. Here, air starts to flow in due to the low pressure at the storm center. As it flows in, air undergoes (approximately) isothermal expansion while in contact with the surface of the ocean and its entropy increases. At the eyewall (B), the flow turns upward toward lower pressure and maintains its entropy. This ascent is nearly adiabatic (BC). When it is far from the storm center (C), air descends losing the entropy it gained as electromagnetic radiation to space. This part of the cycle is nearly isothermal (CD). The cycle is completed as air undergoes adiabatic compression by cooling down (DA) while approximately maintaining its entropy. As already known, the energy in the Carnot cycle is used to do work on its environment. In tropical cyclones, on the other hand, work is used in turbulent dissipation, where it is turned back into heat. One more difference is that the energy cycle of tropical cyclones is in fact open, but the differences in energy are small in comparison to the energy used, so we can idealize the cycle as closed [26].

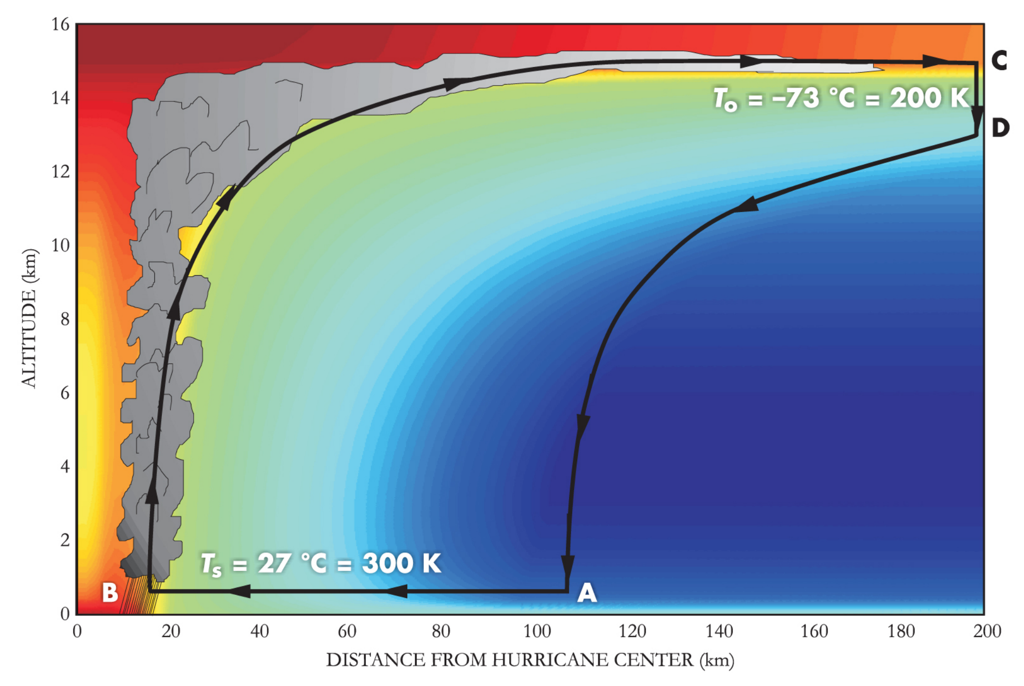


Figure 1.3: Thermodynamic cycle of tropical cyclones. Colors show a measure of the total specific entropy content of the air (cooler colors indicate lower entropy) (taken from [39])

A meaningful quantity to measure the destructiveness of a tropical cyclone is the rate of generation of kinetic energy, which, in a steady storm, equals the rate of dissipation [40]. The corresponding dissipation rate per unit area, D, is given by

$$D = \rho\, C_D v^3 \,, \tag{1.1}$$

where ρ is the density of air, v is a characteristic near-surface wind and C_D is the surface drag coefficient. If we take T_s to be the Sea Surface Temperature and T_0 the mean temperature of the outflow, then, according to the Carnot theorem, the net production of mechanical energy over the first stage of the cycle will be

$$P = 2\pi \frac{T_S - T_0}{T_S} \int_a^b [C_k \rho\, v(k_0^* - k) + C_D \rho v^3] r \,\mathrm{d}r \,, \tag{1.2}$$

where k is the specific enthalpy of the near-surface air, k_0^* is the enthalpy of the air that is in contact with the ocean and C_k is the transfer coefficient of enthalpy. We integrate r, which is the radius measured from the eyewall. The net energy dissipation of the tropical cyclone, on the other hand, will be

$$D = 2\pi \int_a^b C_D \rho v^3 r \, \mathrm{d}r. \tag{1.3}$$

As mentioned earlier, for a steady storm, the rate of generation of kinetic energy is the same as the rate of dissipation. Thus, equating Eqs. (1.2 and 1.3) and integrating near the radius of maximum wind speed, we have the expression

$$v_{\max}^2 \approx \frac{C_k}{C_D} \frac{T_s - T_o}{T_o} (k_0^* - k)\,. \tag{1.4}$$

We take $E \equiv C_k(k_0^* - k)/C_D$, which we call the thermodynamic disequilibrium between the ocean and the atmosphere. We now have the expression

$$v_{\max}^2 = \frac{T_s - T_o}{T_o} E\,. \tag{1.5}$$

This expression, following the considerations made thus far, stands as an equivalent of the kinetic energy used by each tropical cyclone.

1.1.3 Cyclogenesis

Since the earliest studies of tropical cyclone formation, it has been recognized that these systems form from a disturbance that is independent of the regional state of the atmosphere [41]. Later, general conditions that are necessary for tropical cyclone formation were identified by [42] and are as follows:

1. Sea surface temperature above 26 °C.
2. Small vertical shear of the horizontal wind.
3. low-level relative vorticity.
4. Larger values of the Coriolis parameter.

5. Relative humidity of the middle troposphere.

It is important to note that these conditions are necessary, albeit not sufficient, for the development of tropical cyclone-intensity systems.

These necessary conditions account for the formation of individual tropical cyclones. In a larger temporal and spatial scale, some elements have been identified that determine annual and decadal tropical cyclone activity. Sea Surface Temperature, for example, has been found to be strongly correlated with tropical cyclone activity [43]. More recently, Sea Level Pressure (SLP) has been identified as a good predictor of tropical cyclone seasons [44].

It has been noted that the main input of energy into the atmosphere is the sun through the greenhouse effect. This makes solar radiation the main determiner of tropical cyclone activity. Nevertheless, the exact effects of solar activity on tropical cyclones are not known [45], [46]. An additional natural radiative forcing that affects the climate in different ways are stratospheric aerosols (volcanic activity), which act as covers and prevent the incoming radiation from the sun to transfer energy to the atmosphere [47]. More recently, it has been hypothesized that human activity may have an impact on tropical cyclone activity through forcings such as the increase of global temperature and, thus an inrease, albeit lagged, of Sea Surface Temperature. Additionally, anthropogenic aerosols have an influence similar to that of volcanic activity, which can cause different effects on global tropical cyclone activity [9], [12], [48], [49].

1.1.4 Accumulated cyclone energy

Although the total number of tropical cyclones recorded in a given year provides an insight into tropical cyclone activity, studying the mechanical energy released every season can give a more accurate picture of the state of the system. This description is achieved with the Accumulated Cyclone Energy (ACE). This metric, first proposed by [50] as the Hurricane Destruction Potential, is a quantity that varies as the square of velocity, thus being proportional to the kinetic energy of the system.

The ACE index is calculated by summing the squares of the maximum one-minute sustained wind speeds (in knots; $1\,\mathrm{kt} = 0.5144\,\mathrm{ms}^{-1}$) of every storm at six-hour intervals while they are at least at tropical storm strength (> 35 knots) [51]. This is expressed as

$$ACE = \frac{\sum v_{\max}^2}{10,000}\,, \tag{1.6}$$

which has units $10^4\,\mathrm{kt}^2$. This metric can be calculated for individual tropical cyclones or it can be summed to describe whole seasons.

Since the ACE is a sum of squared velocities, it is used as a proxy for the kinetic energy of individual tropical cyclones and seasons [52]. As mentioned earlier (Subsec. 1.1.2), we can idealize tropical cyclones as being thermal engines, which means that the ACE index stands for the mechanical energy produced by these engines [2]. This means that, by studying this metric, we are indirectly studying the right-hand side of eq. (1.5) [53].

As previously stated, tropical cyclones are phenomena that pertain to the atmosphere and climate system, which are themselves perfect examples of complex systems. In the next section we give an account of the main characteristics of complex systems, as well as the tools for their mathematical description and relevant concepts to understand their behavior.

1.2 Complex systems

Tropical cyclones, like many other natural and social systems, exhibit complex behavior that cannot be inferred from the understanding of their isolated individual components alone. The main feature of complex systems is the arising of emergent phenomena due to the interactions between the many components of the system [54]. This idea can be expressed by Aristotle's quote "The whole is more than the sum of its parts". The word complex, which comes from the Latin *plectere* (meaning to weave, to entwine), entails the interconnectedness of the constituent parts of a system. Nevertheless, this interconnectedness does not fully represent what complexity science observes and studies [55].

There is currently no agreed upon definition of the term complex system or complexity science. Nevertheless, we could give a very general definition that encompasses any system that we consider to be complex: a system consisting of many parts that interact with each other in different levels and which showcases nontrivial emergent behavior [56]. Many disciplines define complex systems in different ways based on their specific foci and analyses, but there are some properties that all systems which are considered complex share. These properties include, among others:

- **Nonlinearity.** This property means that the response of a system to a change is not directly proportional to the applied change. Instead, the response can follow a complicated rule that connects the system's constituents.

- **Emergence.** When a system presents behavior or characteristics that its elements do not individually posses it is said to show emergent behavior. This global outcome results from the interactions of the constituent elements and cannot be predicted with their understanding alone.

- **Self-organization.** This term refers to the spontaneous emergence of new levels of order or complexity within a system. It is worth mentioning that this new level of organization is achieved in the absence of external inputs.

Complex systems science approaches the study of nature in a comprehensive and integrative way and attempts to build bridges between the different scientific disciplines to form a more complete understanding of these natural phenomena. This is in contrast to the reductionist approach, which seeks to describe the universe in terms of its basic constituents and their interactions [57]. As mentioned earlier, the most important aspect of complexity science is its attempt at explaining the behavior of complex systems in terms of the interactions between their components.

In spite of the myriad of elements comprising the climate system, including tropical cyclones, it is an example of a complex system whose elements follow fixed physical laws. This means that we can describe the climate system with a mathematical function that describes at least the most basic features of its evolution along time. This mathematical description pertains to the theory of dynamical systems, which will be described in the next section.

1.2.1 Dynamical systems

The theory of dynamical systems studies how systems change over time. We can define a dynamical system as a system whose state is uniquely specified by a set of variables and whose behavior is described by predefined rules. This mathematical model represents the basic behavior of the observed qualities of a system [54].

We can categorize dynamical systems into two main categories depending on their use of time: continuous dynamical systems take time to be continuous and discrete dynamical systems consider discrete time steps for their description. Continuous-time systems are described using differential equations, such as:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = F(x,t)\,. \tag{1.7}$$

Discrete-time systems, on the other hand, use difference equations, also known as iterative maps:

$$x_{t+1} = F(x_t, t)\,. \tag{1.8}$$

The second characteristic of a dynamical system is whether it is linear or nonlinear. In terms of their mathematical description, we say that a linear dynamical system's rule involves a linear combination of state variables. Any system whose state is is not described by a linear combination is known as nonlinear. Lastly, the number of independent variables involved in determining the state of the system is known as the number of degrees of freedom of the system. The number of degrees of freedom of a system determines the dimension of its associated state space, which is the set of all possible configurations of the system [58].

Some dynamical systems, whose rules for evolution along time are completely deterministic, exhibit seemingly erratic behavior. This type of behavior is now known as chaotic behavior and it is the result of nonlinearities in the system. Chaos, as opposed to periodic and quasiperiodic motion, is considered to be irregular in time. This means that chaotic motion cannot be the result of the superposition of periodic motions [59]. Two chaotic systems that are initialized at very close, but not equal states, will follow a close trajectory up to a certain point, where they will diverge and present completely different trajectories. This feature of chaotic systems is known as sensitive dependence on initial conditions. This feature results in the impossibility of exactly predicting the state of the system in the long term. This means that, although chaotic systems are completely deterministic, the range in which we can predict their behavior is exponentially short [60].

As mentioned earlier, chaotic systems exhibit seemingly random behavior and motion in real space. These systems, nevertheless are clearly-organized in their phase space. After transient motion, the trajectories in phase space will settle in a defined region, which is known as the attractor of the system. In the case of chaotic systems, this attractor tends to be fractal in nature, which is known as a strange attractor. [61].

The approach of dynamical systems theory to study natural phenomena analytically assumes they follow a deterministic path of evolution. This assumption works well for simple systems, or when only the essence interests their study. In the case of complex systems, or any real system for that matter, this basic description proves insufficient to describe minor qualities or behavior that does not align with the deterministic rules defined beforehand. Some of these qualities, among others, can be introduced to the analysis assigning random qualities to the processes being studied. This introduces the use of statistics and statistical concepts in the study of complex systems, enriching the deterministic, dynamical systems approach.

1.2.2 Determinism and randomness

The deterministic view of nature that was established before the advent of quantum mechanics and the understanding of chaos was thought to give a complete description of nature and that every phenomenon could be predicted if we had the sufficient information on its variables. The "uncertainty principle" of quantum mechanics discovered by Heisenberg in 1927 pinned down this idea and introduced the notion of intrinsic indeterminism in nature [62]. Additionally, the discovery of chaotic behavior in atmospheric systems by Lorenz in 1960 introduced the idea of unpredictability in deterministic systems because of their strong sensitive dependence to initial conditions [63].

The description of nature of quantum mechanics introduced the possibility of the existence of real randomness in nature. Apart from the discussion of whether real randomness exists, many phenomena and systems behave in a way that is indistinguishable from mathematical randomness. This may come from an incomplete understanding of their inner dynamics, measurement noise or real erratic behavior. Additionally, chaotic systems, as mentioned before, exhibit behavior that can be practically indistinguishable from random motion, rendering a deterministic description of their observables practically impossible.

The impossibility to perfectly describe many systems with the tools developed for deterministic behavior leads to find other ways to study them. An alternative way to describe this kind of phenomena, be they deterministic, random or highly complex in nature, is to consider them as following the rules of random behavior. With this approach we can use probability theory and statistics to describe these systems and gain insight into how they behave [64]. This apparent dichotomy between the dynamical systems approach and the study of phenomena using concepts from probability and statistics does no other than enrich the discussion. Studying a phenomenon using concepts and tools stemming both from deterministic and random theory offers a broader understanding of their inner workings, as well as their macroscopic behavior.

In these introductory sections we defined the aims of this thesis work and put the present study into context. We then laid out the theoretical background necessary to undertake the research ahead. In the following chapter we present the analysis tools that will be used to undertake the research. These tools will first be used to describe synthetic data and then, in Ch. 3, they will describe real, observational data from global tropical cyclone activity in the Northwest Pacific.

Chapter 2

Time series analysis tools

In this thesis we will analyze observations of the Accumulated Cyclone Energy index in the Northwest Pacific over many years. These observations are presented in the form of a time series. By definition, a time series is a set of regular observations of a quantitative characteristic of a phenomenon taken at successive points in time. There are two main schools of thought to analyze time series. The first one views time series as samples from stochastic processes and applies traditional statistical tools and assumptions to study them. The second school of thought views the data as coming from deterministic dynamical systems that can be plagued with noise. This approach aims to reconstruct the underlying system from which these data were obtained [64].

In the following sections we will present the tools to analyze time series which fall into the two categories mentioned above, namely statistical and dynamical analysis tools. In order to illustrate the results they produce, we will study two kinds of time series. The time series from the first kind come from deterministic systems, namely one- and three-dimensional chaotic maps. The second kind comprises time series resulting from stochastic processes with different characteristics.

2.1 Synthetic data

Time series can be obtained from virtually anything that can be quantified at different points in time. The two main categories in which a process can fall is two-fold: deterministic and stochastic processes. In the following subsections we will present models that generate time series and their main characteristics will be mentioned.

2.1.1 Chaotic systems

A deterministic system is one in which future behavior depends directly on its present state. A deterministic model will always produce the same result given predefined parameters. The following systems exhibit seemingly erratic behavior and thus may lead to believe they are produced by random processes, but they are fully deterministic. These systems are called chaotic.

Tent map

The best examples of complex behavior are systems that exhibit chaotic behavior. A simple one-dimensional map that exhibits this type of behavior is given by the expression:

$$X_{n+1} = \frac{1}{|X_n|^{1/4}} - \frac{1}{2} - |X_n| \,. \tag{2.1}$$

Although being completely deterministic, this system has an approximately Gaussian distribution [65]. Figure 2.1a) shows a time series generated by the tent map.

Logistic map

The most basic and paradigmatic dynamical system that exhibits chaotic behavior is the logistic map. This discrete-time map was first proposed by [66] as a discrete-time alternative to the logistic equation, which is in turn a model of population growth. This system is given by:

$$X_{n+1} = rX_n(1 - X_n)\,, \tag{2.2}$$

where $X_n \in [0,1]$ represents the ratio of the population and r, usually taken from the range $(0,4]$, is called the control parameter. For low values of r, i.e. $r \in (0,3)$, the solutions will have one period. After $r > 3$ the period will double until, we reach $r \approx 3.56995$. This is when the number of periods becomes infinite and the system exhibits chaotic behavior [67]. An implementation of the logistic map can be seen in Fig (2.1c).

Lorenz system

After the logistic map, which is the simplest one-dimensional system that exhibits chaotic behavior comes the Lorenz system, which is three-dimensional and continuous. This system of ordinary differential equations is a simplified mathematical model for atmospheric convection [68]. This model is given by the next system of differential equations:

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} =& \sigma(y - x)\,, \\ \frac{\mathrm{d}y}{\mathrm{d}t} =& x(\rho - z) - y\,, \\ \frac{\mathrm{d}z}{\mathrm{d}t} =& xy - \beta z\,. \end{aligned} \tag{2.3}$$

Here, x is the rate of convection, y is the horizontal temperature variation and z stands for the vertical temperature variation. The constants σ, ρ and β are positive system parameters [69]. It has been noted that some combinations of these parameters yield chaotic behavior. Figure (2.1e) shows the time series produced by the Lorenz system.

2.1.2 Stochastic processes

We can model systems and phenomena that vary in an apparently random manner with the help of stochastic processes. The simplest way to obtain a time series of a stochastic process is to take a sequence of variables produced by a random number generator. The following expression produces a random sequence of uncorrelated values:

$$X_{n+1} = \eta_n \, , \tag{2.4}$$

where $\eta \in [0, 1]$ is a random number generated by a specified probability distribution. A random sequence generated by a uniform distribution is shown in Fig. (2.1b).

Random walk

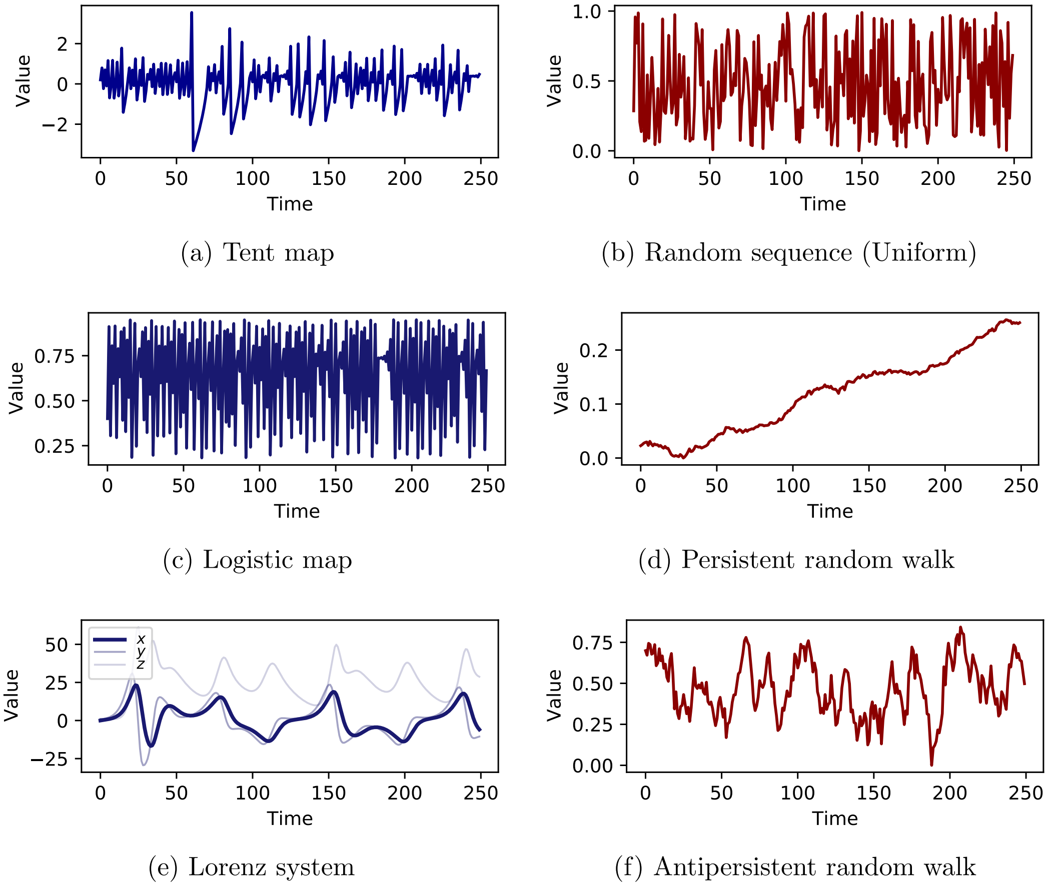
A random walk is a stochastic sequence given by:

$$X_{n+1} = X_n + \eta_n \tag{2.5}$$

where η_n stands for a random number at time n. Note that, while η_n is an independent random variable, the actual value of X_n is not, as it describes the path followed by the random increments given by η_n. Figures (2.1d, 2.1f) depict realizations of a random walk that exhibit different types of behavior. Figure (2.1d), for instance, exhibits persistent behavior, which means that the values will tend to follow the local (immediate) trend. Figure (2.1f), in turn, is an example of antipersistent behavior, which means that following values will tend to oscillate around the local mean.

2.2 Statistical analysis

The highly nonlinear nature characteristic of complex systems and the frequent presence of chaos mean we cannot exactly predict their future behavior. Furthermore, there are inherent uncertainties in the components and interactions of these systems. This means that, even if the underlying rules controlling the dynamics are deterministic, the evolution of these systems is practically stochastic. So, according to this school of thought, we can study and model them



(a) Tent map

(b) Random sequence (Uniform)

(c) Logistic map

(d) Persistent random walk

(e) Lorenz system

(f) Antipersistent random walk

Figure 2.1: Time series of chaotic systems (left) and stochastic processes (right).

as if they were the result of stochastic processes [70].

In light of this, the first thing to do when presented with time series data is to plot it and perform a visual inspection. With this we can directly tell if the data follow a trend and seasonalities are present. This preliminary inspection can yield clear outliers or extreme values.

2.2.1 Probability Density Function

Once the most evident features of a time series have been pointed out we can turn our attention to the distribution of the data. The most common way to represent the outcomes of an experiment is through the frequency distribution, which shows the frequency of occurrence of each possible outcome. A representation of this characteristic from observed data is called a histogram. We build a histogram by grouping the values of the observations into equally-sized intervals. We then plot a bar graph whose height represents the number of observations present in every interval. This graph gives an immediate impression of the Probability Density Function (PDF) that produced the observations, which is a function that assigns a probability of occurrence to each value of a random variable, which corresponds to every data point in the case of a time series [70].

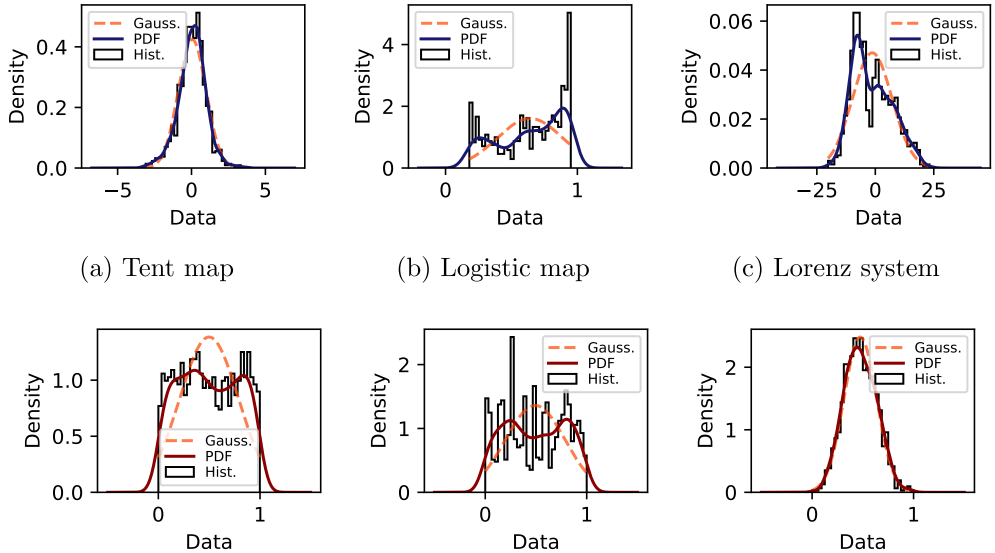
We can describe the shape of a function (in this case the PDF) with its moments. For the present analysis, only the first two moments of the data will be obtained. The first moment, known as the mean is the expected value and it is a measure of the central tendency of the data. We represent this expected value with $\mu \equiv E[X]$. The second moment is the variance, which is the square of the standard deviation. The standard deviation, in turn, is given by $\sigma \equiv (E[(x-\mu)^2])^{\frac{1}{2}}$ and it measures the amount of variation within the data.

Figure (2.2) presents the histograms, Probability Density Function and Gaussian fit of our systems for comparison. As mentioned earlier, the tent map has an approximately Gaussian distribution (2.2a), as well as the antipersistent random walk (2.2f). Both the time series obtained from a uniform distribution and the persistent random walk display an apparent uniform histogram (2.2d and 2.2e respectively). This means that every value has the same probability to be observed. The Lorenz system shows an approximately bimodal distribution (2.2c), which means the values tend to be clustered around two centers (this is reminiscent of the actual butterfly-shaped attractor). The logistic map, in turn, appears to have a trimodal distribution (2.2b).

2.2.2 Autocorrelation function

A given time series may have an underlying structure, but the presence of noise can make it difficult to identify it. The autocorrelation of a time series can help us identify randomness, whether it is stationary or it has a trend and detect the seasonalities present in it.

$$ACF(k) \equiv \sum_{n=1}^{N-k} \frac{(X_n - \langle X \rangle)(X_{n+k} - \langle X \rangle)}{\sum_{n=1}^{N-k}(X_n - \langle X \rangle)^2}\,, \tag{2.6}$$



(a) Tent map (b) Logistic map (c) Lorenz system

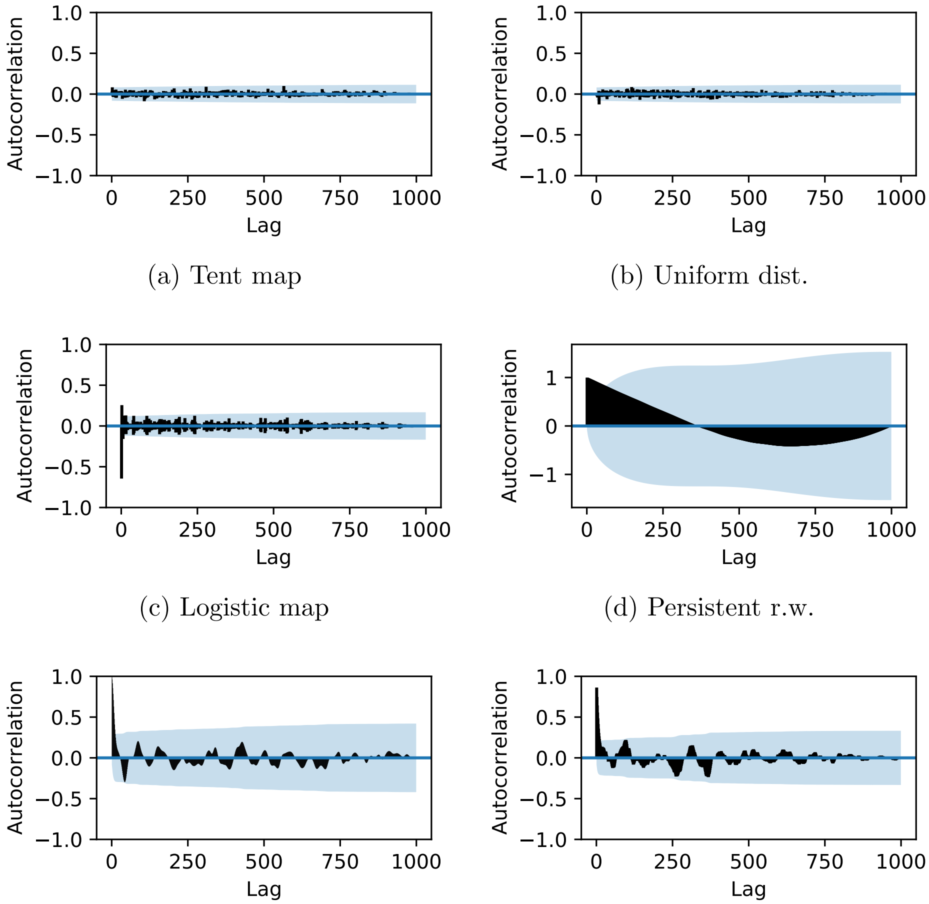
(d) Uniform dist. (e) Persistent r.w. (f) Antipersistent r.w.

Figure 2.2: Histogram, Probability Density Function and Gaussian fit of some chaotic systems (upper panels) and stochastic processes (lower panels).

where $\langle X \rangle$ is the mean value of X and k is the lag. This function is the ratio of the autocovariance to the variance of the data [65].

The behavior of ACF as k increases can give insight into the nature of the data. For instance, when X_n comprises a random process, $ACF(k)$ will be near zero for all k. When the system is nearly periodic, the function will be a decaying oscillation. The autocorrelation decays rapidly to zero when the time series is stationary and it drops slowly for a non-stationary time series [71]. In general, the value of ACF is 1 at $k = 1$ and falls to zero for large k. The value of k at which ACF falls to $1/e \simeq 37\%$ is called the correlation time τ_c and it can give some insight into the memory of the system. An estimate of the rate at which predictability is lost is given by $1/2\tau_c$ [65].

Figure (2.3) shows the autocorrelation plot for the chaotic systems and stochastic processes we are comparing. As mentioned earlier, an intrinsically random process will show low correlation values for any k, such as Fig. (2.3b) shows. Figures (2.3a) and (2.3c), however, are completely deterministic yet show low correlation as well. Figure (2.3d), for instance, drops slowly to zero, which shows that this system has a trend. From these figures, we can see that



(a) Tent map

(b) Uniform dist.

(c) Logistic map

(d) Persistent r.w.

(e) Lorenz system

(f) Antipersistent r.w.

Figure 2.3: Autocorrelation plots of chaotic systems (left) and stochastic processes (right). The shaded region marks the 95% confidence interval.

ACF is useful to identify persistence in a given time series.

2.2.3 Rank-ordering statistics

Many physical and natural systems' PDFs are characterized by *heavy tails*, which means they differ from the well-behaved Gaussian and Gaussian-like distributions. These heavy tails mean that large events have larger probabilities than they would have were they well-behaved distributions. In a heavy-tail distribution, neither the mean nor the variance can give information on the behavior of the phenomenon we are studying since there is no characteristic scale in which it happens. If the sample gets bigger, both the mean and variance can grow indefinitely [70].

Apart from these applicability problems, there are some practical difficulties that arise when using the histogram representation. The first difficulty of representing these systems with a histogram is that, by definition, the statistics of rare events is limited as they become undersampled due to their being in the tail [72]. Another limitation of the histogram representation is that the selection of number of bins is arbitrary and different numbers can lead to varying conclusions [70].

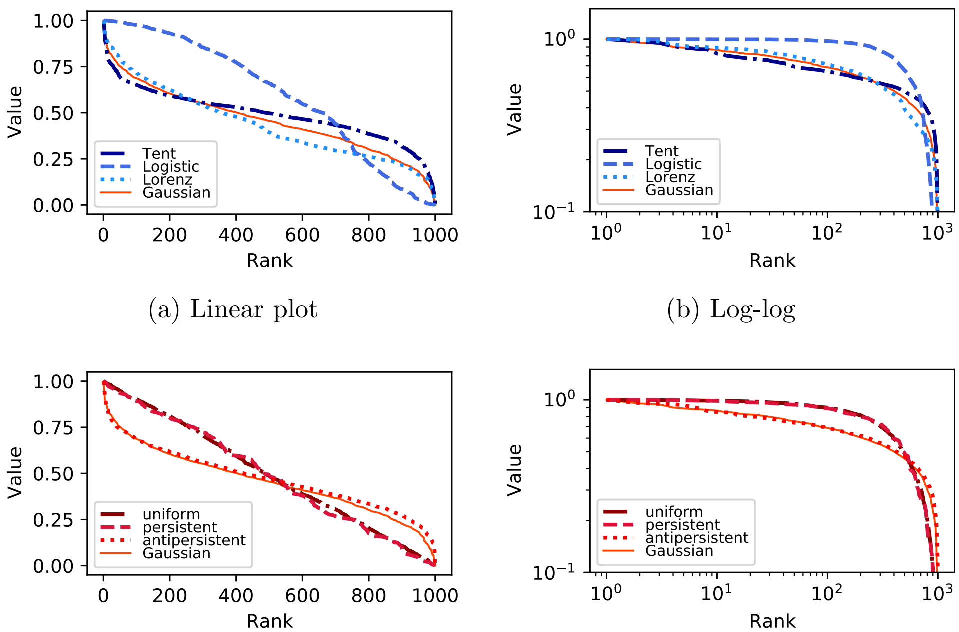
An alternative approach to study the distribution of data is the rank-ordering technique first introduced in this context by [73]. This technique presents a function that quantifies the dependence of the observations according to their rank. This approach introduces the advantage of constraining the fit of empirical distributions by the large events, rather than by the large majority of small and intermediate event sizes, which is the usual fit of density and cumulative distributions [74].

To construct the rank-ordering representation we do as follows. We have a time series consisting of N observations. We reorder these observations by decreasing values:

$$v_1 \geq v_2 \geq \ldots \geq v_n \geq \ldots \geq v_N \, , \tag{2.7}$$

where v_1 is the largest value, v_2 is the second, and so on. This method consists in quantifying the dependence of v_n as a function of the rank n.

Figure (2.4) shows the linear and log-log plots of the ranking distribution of our reference systems, namely chaotic (blue lines) and stochastic processes (red lines). The linear plots show the general shape of the distributions, while the log-log plot zooms in on the larger values, which means we can more closely inspect outliers. A Gaussian distribution is plotted as a reference (orange lines) to be compared. Two of the chaotic systems, namely the tent map and Lorenz system, have a distribution that closely resembles a Gaussian process. Regarding the stochastic processes, the antipersistent random walk resembles the Gaussian. This process is ultimately a correlated Gaussian noise. The sequence taken from a uniform distribution and the persistent process resemble a straight line. This means that the values are uniformly distributed



(a) Linear plot

(b) Log-log

(c) Linear plot

(d) Log-log

Figure 2.4: Ranking plots of chaotic systems (upper half panels) and stochastic processes (lower half panels).

throughout the data.

The ranking representation is useful to identify the presence of fractal behavior in a process or system, which means a specific mechanism dominates the behavior across scales. This can be identified as a straight line in the log-log representation. The fact that our reference systems show no fractal qualities in the log-log plot means there is no self-similarity in their behavior.

2.3 Dynamical Analysis

As mentioned at the start of this chapter, the second approach to study time series tries to elucidate the inner dynamics of the systems that produced them. Chaotic systems, for instance, often produce outputs that are apparently random, but studying their dynamics unveils an often simple structure. The first tool proposed here studies the presence of long-term dependence in the system. The next tools work on the state space of the system that produced the data.

2.3.1 Long-term memory: Hurst exponent

It has been observed that many natural systems, including the atmosphere, exhibit long-memory dependence [75]. This means that correlations between observations with increasing time interval or spatial distance decay much slower than would be expected from classical stochastic processes [76].

The presence of long-memory dependence or persistence in a time series can be characterized by the Hurst exponent H [77]. This index is defined in the interval $0 < H < 1$ and it quantifies the presence of persistent (the tendency to cluster in a direction) or anti-persistent (the relative tendency to regress strongly to the mean) behavior in a time series. For values $0 < H < 0.5$, H indicates anti-persistent behavior. Conversely, H indicates persistent behavior if it is in the range $0.5 < H < 1$. For $H = 0.5$, it means the time series is similar to a random walk. A value of H close to 0.5 may be an evidence of chaotic behavior.

The most popular tool to estimate H is the rescaled range analysis (R/S). This method computes the ratio R/S, where R is the range of the input data and S is its standard deviation. We compute these quantities as follows [78]. The difference between two consecutive data points will be the input data:

$$I'_i = I_i - I_{i+1}\,, \tag{2.8}$$

where I_i is the record at time i. Over a period τ, the mean will be

$$\langle I' \rangle_\tau = \frac{1}{\tau}\sum_{i=1}^{\tau} I'_i\,. \tag{2.9}$$

We define the difference between I'_i and $\langle I' \rangle_\tau$:

$$X(i,\tau) = \sum_{u=1}^{i}[I'_u - \langle I' \rangle_\tau]\,. \tag{2.10}$$

Finally, the range will be

$$R(\tau) = \max X(i,\tau) - \min X(i,\tau)\,, \tag{2.11}$$

and the standard deviation will be given by:

$$S(\tau) = \left\{ \frac{1}{\tau} \sum_{i=1}^{\tau} [I'_\tau - \langle I' \rangle_\tau] \right\} . \tag{2.12}$$

We finally have the relation proposed by Hurst [77]:

$$\frac{R(\tau)}{S(\tau)} = C \cdot \tau^H , \tag{2.13}$$

where C is a positive constant and H is the Hurst exponent . For every τ we get a value R/S that follows a power law. The Hurst exponent H, then, will be the log-log slope of Eq. (2.13). Table (2.1) shows the calculated Hurst exponent for our reference systems. Note that the algorithm used gave an error for the Lorenz system and the persistent and antipersistent random walks.

Table 2.1: Hurst exponent with linear and nonlinear fit of reference systems together with the expected result from theory or known results

	H (linear fit)	H (nonlinear fit)	Expected
Tent	0.539	0.516 ± 0.008	
Logistic	0.329	0.321 ± 0.011	
Lorenz	NaN	NaN	≈ 0.5 [79]
Uniform	0.571	0.579 ± 0.037	0.5
Persistent	NaN	NaN	0.7
Antipersistent	NaN	NaN	0.3

2.3.2 Phase space

As mentioned earlier, in dynamical systems theory, the evolution of a system is defined in some phase space. This phase space represents the space of all its possible states. Thus, a specific point in phase space represents a specific state of the system. Every degree of freedom of the system is represented as an axis in phase space. Thus, the number of degrees of freedom of the system determines the dimension of the phase space. The evolution of a dissipative system will tend to a subset of its phase space once the transient response due to the initial conditions settles. This subset is called the attractor of the system, and it is invariant under its evolution [80].

Phase space reconstruction

As mentioned earlier, the evolution of a system is fully determined by its associated differential equations, in the case of a continuous system, or by a map for discrete-time systems. However, in experimental or observational settings, there is no access to the underlying expressions that govern the dynamics of the system [80]. Furthermore, in the case of complex systems, there may be times in which no underlying governing rules exist [81]. Thus, in an experimental or observational setting, what we observe is not the phase space of the system itself, but imperfect observations that take the form of a time series [82].

According to Takens' theorem (also known as embedding theorem), it is possible to reconstruct the attractor of the system from the time series alone [83]. This reconstructed attractor will be topologically equivalent to the unknown attractor. This reconstruction is achieved through the method of delays, which works as follows.

We have a one-dimensional time series of length n comprised of observations s_n with a sampling time Δt. Every observation will be given by

$$s_n = \psi(\mathbf{x}(n\Delta t))\,, \tag{2.14}$$

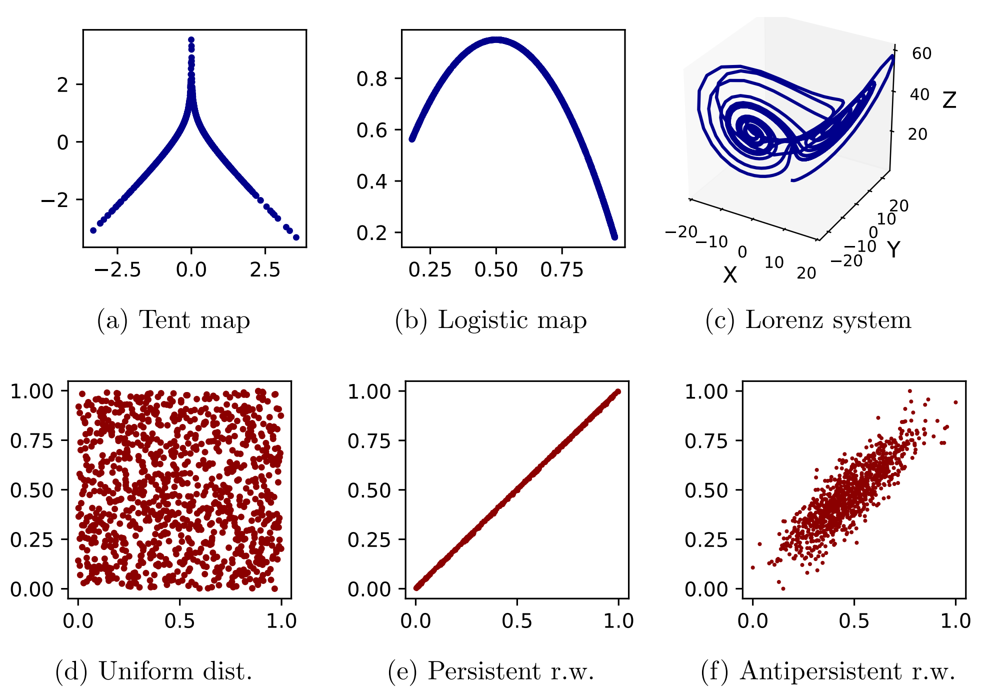
where ψ is an observable measurement function. We neglect the effect of measurement noise for this discussion. An m-dimensional reconstruction of the $\mathbf{s}_n$ state vectors will be given by

$$\mathbf{s}_n = (s_{n-(m-1)\tau}, \ldots, s_{n-\tau}, s_n) \qquad \tau \in \mathcal{Z}^+\,. \tag{2.15}$$

The time interval between adjacent coordinates of the state vector is $\tau\Delta t$, and it is known as the lag or delay time.

The optimal selection of the parameters for reconstructing the phase space is ultimately a question of optimization and accuracy [82]. For most systems, and as an initial approximation, we can set a time lag $\tau = 1$. The embedding dimension m can be set to two or three dimensions in order to elucidate some sign of structure.

The phase spaces of our reference systems are shown in Fig. (2.5). Looking at these phase portraits we can immediately identify the two types of systems: chaotic (plotted in blue) and stochastic (plotted in red). The upper half panels show a clear structure. The first two panels (2.5a and 2.5b), for instance, show a very simple two-dimensional attractor. Figure (2.5c), in turn, comprises three-dimensional trajectories following a clear attractor, known as the Lorenz attractor. The panels on the lower half, on the other hand, show something different. Figure (2.5d), for instance fills the space uniformly, as expected from the distribution from which the data points were taken. Figures (2.5e and 2.5f) show clear correlation (we can identify this for the data points follow a straight line of slope $m = 1$). The antipersistent time series, in contrast, is more scattered around the line, which means the variance is higher.



(a) Tent map (b) Logistic map (c) Lorenz system

(d) Uniform dist. (e) Persistent r.w. (f) Antipersistent r.w.

Figure 2.5: Phase space of chaotic systems (upper half) and stochastic processes (lower half).

2.3.3 Complexity: Attractor dimension

The fractal dimension, D, of a profile or surface is a measure of its roughness. We can use fractal dimension to describe physical objects, which exist in ordinary physical space, and strange attractors, which exist in the state space of chaotic dynamical systems [84]. Higher values of D indicate rougher surfaces in the case of real-world objects and a more complex nature in the case of objects in phase space.

When quantifying the complexity of a system through its time series, we can do so by capturing either its spatial or temporal dimensionality [15]. By quantifying the temporal dimensionality of a time series, we are estimating the complexity of the signal treating the time series itself as a geometric figure [15]. In the case of the spatial dimensionality of a system we are actually determining the fractal dimension of the attractor contained in its state space. Estimating the fractal dimension of the attractor reconstructed from a recorded time series can give an estimate of the active Degrees of Freedom of the system [84].

There are many ways to estimate the fractal dimension of a system from its time series. The most robust and closest estimation of the attractor dimension of a given system is given by its Correlation Dimension (CD), which was introduced in [85].

This so-called correlation dimension (CD) is obtained by considering correlations between points of a time series on the attractor and it is computed as follows. We define the correlation integral

$$C(r) = \lim_{N\to\infty} \frac{1}{N^2} \sum_{i,j=1}^{N} \theta(r - |\mathbf{X}_i - \mathbf{X}_j|) \equiv \int_0^r d^d\, r'\, c(\mathbf{r}')\,, \tag{2.16}$$

where θ is the Heaviside function and $c(\mathbf{r}')$ is the standard correlation function. For small r, this correlation integral behaves as

$$C(r) \propto r^{\nu}\,. \tag{2.17}$$

It has been shown that ν is very close to D [61], [85] This means that we can use it as an approximation of the fractal dimension of the attractor ($D_C \equiv \nu$). With this, we now obtain the Correlation Dimension [84]:

$$\nu \equiv D_C = \lim_{r \to 0} \frac{\log C(r)}{\log r} \,. \tag{2.18}$$

This algorithm, known as the Grassberger-Procaccia algorithm, has proven to be a robust and practical way to obtain estimations of the fractal dimension of a system from a univariate time series when $D < 2$ [85]. Table (2.2) presents the average correlation dimension found with the Grassberger-Procaccia algorithm for an embedding dimension $m = 2$ for all our reference systems, except for the Lorenz system ($m = 3$), and a delay time τ with values going from 1 to 50.

Table 2.2: Correlation dimension of reference systems ($N = 1000$)

	Tent	Logistic	Lorenz
D_C	1.88 ± 0.13	1.59 ± 0.15	1.73 ± 0.09

	Uniform	Persistent	Antipersistent
D_C	1.85 ± 0.02	1.08 ± 0.017	1.92 ± 0.013

Chapter 3

Results

The analysis tools presented in the previous chapter were used to study the data on tropical cyclone activity, as well as samples of the chaotic systems and stochastic processes discussed before. In order to make a useful comparison, the length of the synthetic data sets was matched to that of the observed data (i.e. 72 data points). Building from the vast literature and repertoire of libraries and modules found in Python, some built-in algorithms were used for our analysis tools and some were adapted from them to fit the specific needs of this project.

This chapter briefly introduces the observational data under study and very generally describes it. Then, a statistical analysis using the tools presented in previous sections is carried out followed by a dynamical account of the data. Additionally, each tool is used to analyze the observational data and the samples of our reference systems.

3.1 Tropical cyclone activity data

As mentioned before, the variable under study is the Accumulated Cyclone Energy index, which tracks the wind speed of particular events or seasons. Tropical cyclone data from multiple agencies and historical databases are combined at a central database, which can be found in the International Best Track Archive for Climate Stewardship (IBTrACS) [30]. This database offers global TC best track data of cyclone distribution, frequency and intensity.

Figure (3.1) shows the yearly time series of Accumulated Cyclone Energy in the Northwest Pacific basin. This data set comprises 72 data points and goes from the year 1950 to 2021. It has mean $\mu = 304 \times 10^4\,\mathrm{kt}^2$ and standard deviation $\sigma = 99 \times 10^4\,\mathrm{kt}^2$. A simple visual inspection shows that the data form a non-smooth curve with no apparent trend. There are three clear outliers: two minima with values $110 \times 10^4\,\mathrm{kt}^2$ and $121 \times 10^4\,\mathrm{kt}^2$ observed in 1999 and 2010, respectively, while a clear maximum with value $570 \times 10^4\,\mathrm{kt}^2$ was observed in 2019.

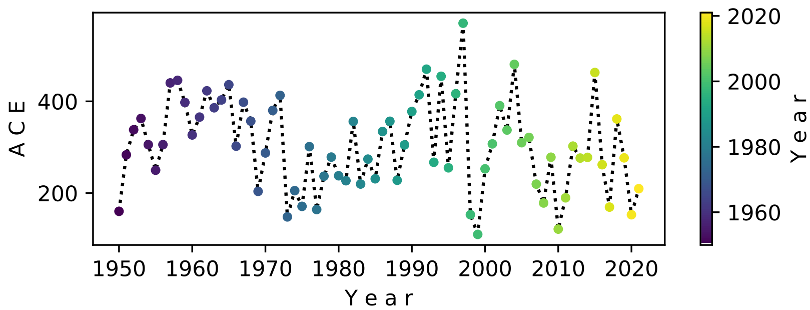


Figure 3.1: Time series of annual Accumulated Cyclone Energy in the Northwest Pacific.

3.2 Statistical analysis

As mentioned earlier, we can find periodicities, as well as identify whether the data have a trend or not with the autocorrelation function. Figure (3.2) shows an autocorrelogram of our observational data with lags ranging from 1 to 71 years with a confidence interval of 95%. This plot shows significant autocorrelations for time lags of 1, 16 and 17 years with values 0.28, -0.38 and -0.36, respectively. Additionally, Fig. (3.3) shows the autocorrelograms of our reference systems.

Building from the autocorrelation found with the autocorrelogram, a moving average and standard deviation was plotted together with the original time series. Figure 3.4 shows this plot with 16- and 30-year moving mean and standard deviation superimposed on the original time series. The 16-year average corresponds to the autocorrelation found for a lag of 16 years. The 30-year smoothing, in turn, conforms to the agreed time average which characterizes climate

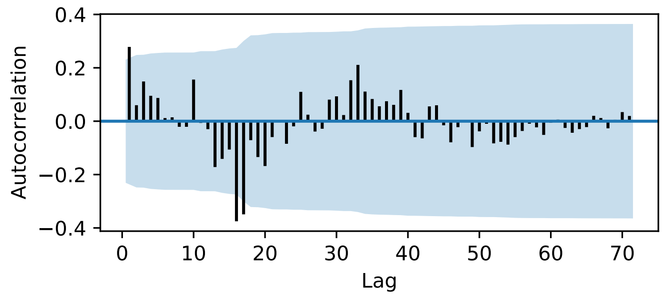


Figure 3.2: Autocorrelation function of the Accumulated Cyclone Energy in the Northwest Pacific. The shaded region indicates 95 % confidence interval.

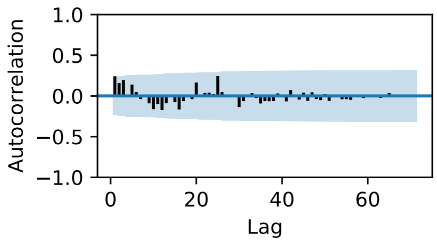
[12]. For the 30-year average a slight downward trend of μ can be identified, while σ appears to increase after around the year 1984.

Plotting the histogram lets us visualize the distribution of values of the data. Figure (3.5) shows the histogram of the data under study and presents the Probability Density Function associated with it. Additionally, a Gaussian using μ and σ from the data is fitted to the plot to assess how the distribution compares to a Gaussian process. As this figure shows, our data are close to a Gaussian distribution that has $\mu = 304 \times 10^4\,\mathrm{kt}^2$ and $\sigma = 99 \times 10^4\,\mathrm{kt}^2$. A closer look at the Probability Density Function reveals a slightly asymmetric bimodal distribution.

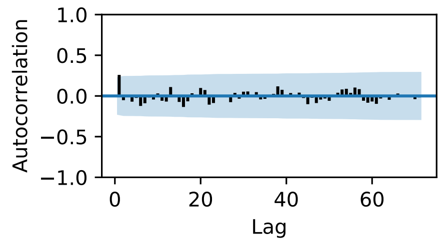
Additionally, the histogram, Probability Density Function and Gaussian fit for our reference systems, namely chaotic systems and stochastic processes, are shown in Fig. (3.6). Comparing these figures with Fig. (2.2), it can be sen that the data preserve their characteristic shape regardless of the shorter time series and number of bins used for the histogram.

As already mentioned, by ranking the values of a given data set we can get the distribution of the data as well as a perspective on the rare, largest elements. Such a representation is given in Fig. (3.7). This figure consists of a linear plot in Fig. (3.7a) and a log-log representation in Fig. (3.7b). The representation in linear scale shows a straight-line profile with a clear outlier for the highest value (first place), and the two lowest values (last two places in the rank). The log-log plot shows.

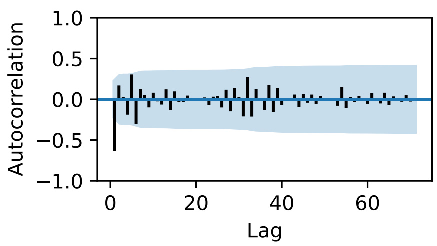
For comparison purposes, Fig. (3.8) shows the ranking representation of our reference systems with $N = 72$ data points. Again, this short number of data points still shows the general



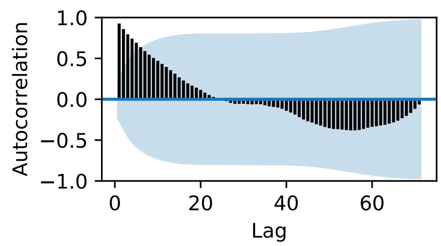
(a) Tent map



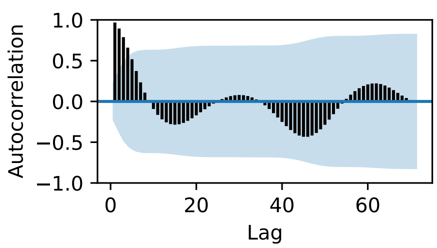
(b) Uniform dist.



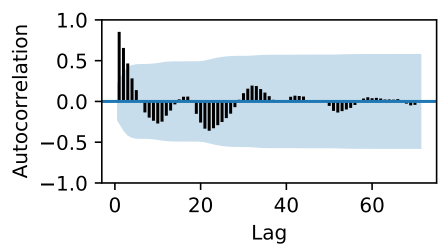
(c) Logistic map



(d) Persistent r.w.



(e) Lorenz system



(f) Antipersistent r.w.

Figure 3.3: Autocorrelation plots of chaotic systems (left) and stochastic processes (right) with $N = 72$ data points. The shaded region marks the 95% confidence interval.

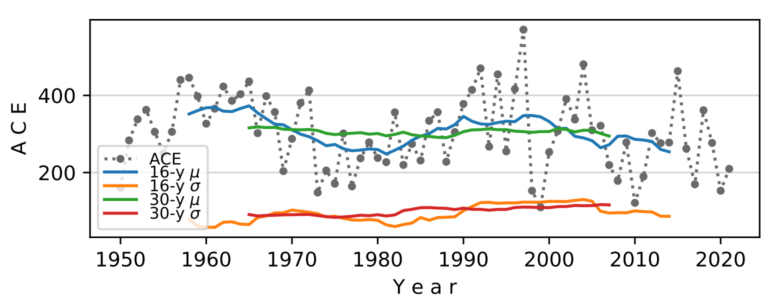


Figure 3.4: 16 and 30 year moving average and standard deviation of annual Accumulated Cyclone Energy in the Northwest Pacific.

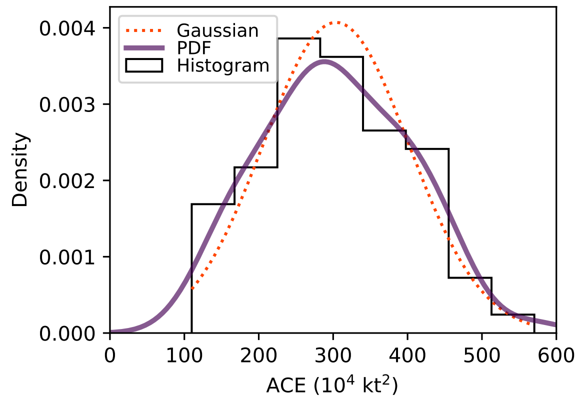


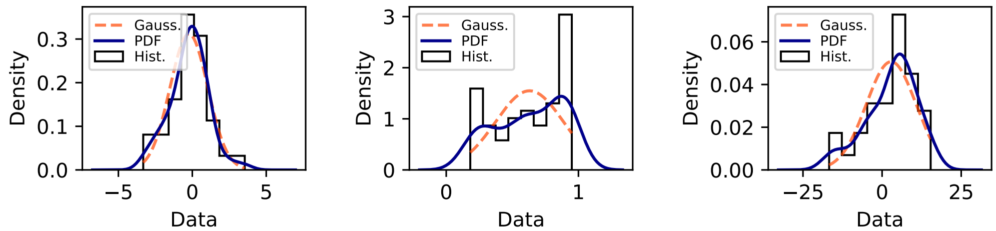
Figure 3.5: Histogram, associated Probability Density Function and fitted Gaussian of the data from the Accumulated Cyclone Energy in the Northwest Pacific

distribution of the data, although the sinusoidal shape of the Gaussian and Gaussian-like distributions is less pronounced and the noise is more evident.

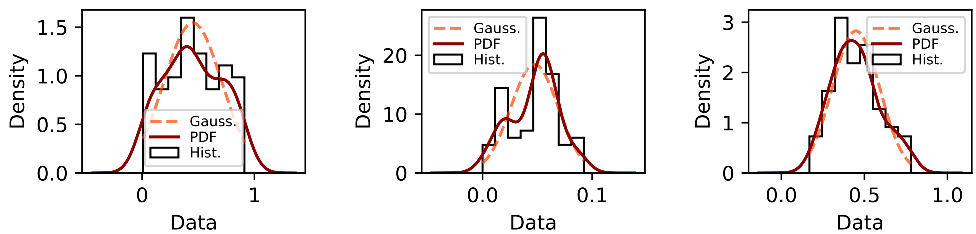
3.3 Dynamical analysis

This section consists of analyses carried out on the observed data and the chaotic and stochastic systems using the tools developed by nonlinear dynamical systems theory, which aim at reconstructing the system that is believed to have generated the data.

Table 3.1 shows the calculated H using rescaled range analysis. Both a linear and nonlinear fit were applied to the data points obtained with the algorithm to compare the performance and similarity of results. The value of H for the Accumulated Cyclone Energy shows that it presents antipersistent behavior. As for the reference systems, the tent map has H close to 0.5, which points at the presence of chaos. Both the Logistic map and Lorenz system show antipersistent behavior, which was not expected as they are in the chaotic regime. The Hurst exponent of all the stochastic processes is close to 0.5. For the time series generated by a uniform distribution this was expected, but for the antipersistent process, H was expected to be lower. The algorithm gave an error when computing H for the persistent random walk. This leads to suspect that it did not work well for such short time series as the ones analyzed in this section.

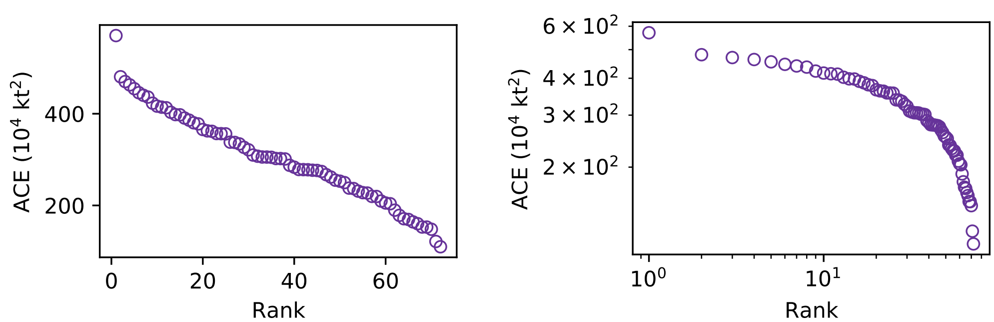


(a) Tent map (b) Logistic map (c) Lorenz system



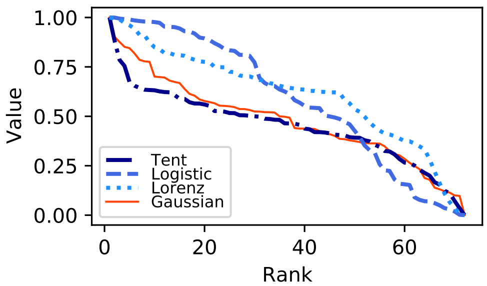
(d) Uniform dist. (e) Persistent r.w. (f) Antipersistent r.w.

Figure 3.6: Histogram, Probability Density Function and Gaussian fit of chaotic systems (upper half) and stochastic processes (lower half).

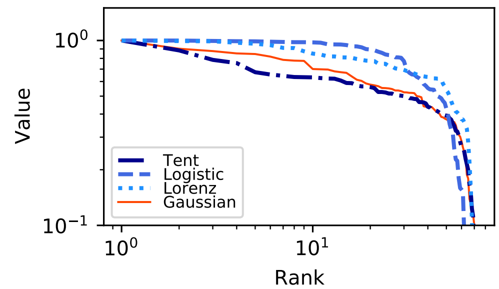


(a) Linear plot (b) Log-log plot

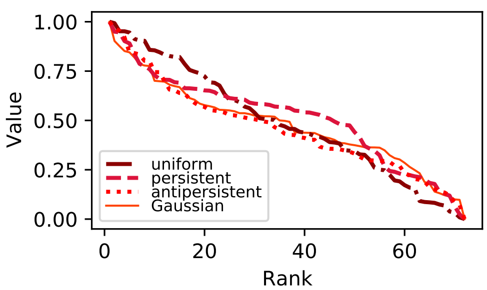
Figure 3.7: Ranking representation of the Accumulated Cyclone Energy in the Northwest Pacific in (a) linear plot and (b) log-log plot of the data.



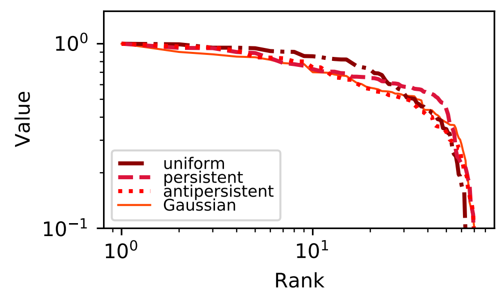
(a) Linear plot



(b) Log-log



(c) Linear plot



(d) Log-log

Figure 3.8: Ranking plot of chaotic systems (upper half) and stochastic processes (lower half).

As the first approximation to determine whether the data have an underlying attractor or a deterministic system generated them, a phase space reconstruction was made using the method of time delay. Figure (3.9) shows a return map of the observed data using a time delay $\tau = 1$ year and embedding dimension $m = 2$. The dispersion and no clear order of the data points suggest the presence of stochastic behavior. Nevertheless, the arrows and color-coded time evolution of the data points suggest they follow quasi-trajectories around two main clusters. This is a signal that there may be some underlying structure which the data points are following. Additionally, the trajectory of data points seems to be following a line of slope one plus some dispersion.

For comparison purposes, Fig. (3.10) shows the phase space reconstruction of the reference systems, i.e. the chaotic and stochastic data sets. The presence of a structure is clear for the chaotic systems. For the stochastic time series, the lack of pattern is seen for the random sequence taken from a uniform distribution (3.10d). On the other hand, the random walk with persistent and antipersistent behavior, shown in Figs. (3.10e and 3.10f), show a clear correlation.

Table 3.1: Hurst exponent of the annual Accumulated Cyclone Energy in the Northwest Pacific basin in contrast with chaotic systems and stochastic processes.

	H (linear fit)	H (nonlinear fit)	Expected
ACE WP	0.399	0.370 $\pm$ 0.071	
Tent	0.576	0.547 $\pm$ 0.037	
Logistic	0.239	0.217$\pm$ 0.056	
Lorenz	0.430	0.375$\pm$ 0.075	≈ 0.5
Uniform	0.538	0.506$\pm$ 0.038	0.5
Persistent	NaN	NaN	0.7
Antipersistent	0.490	0.429 $\pm$ 0.068	0.3

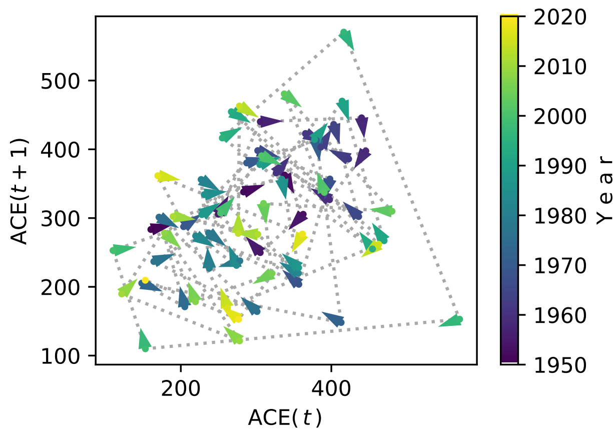
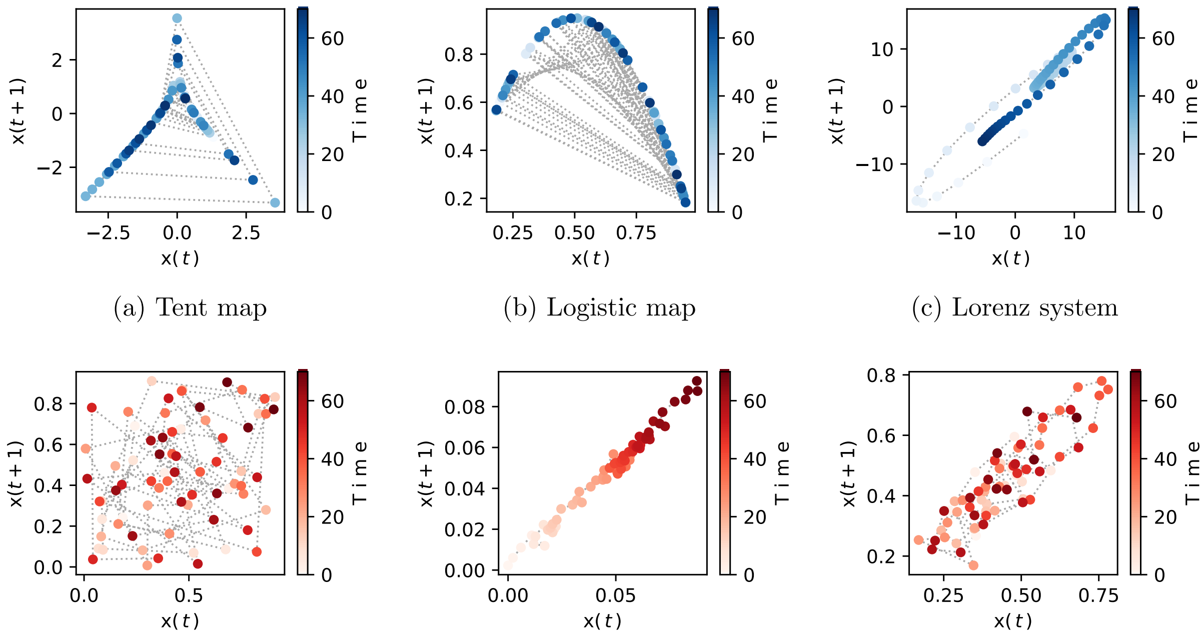


Figure 3.9: Phase space reconstruction of the annual Accumulated Cyclone Energy in the West Pacific basin.

The main difference between these two figures is the dispersion of the points. In Fig. (3.10e), the variance is low and there seems to be a temporal continuity, as the color of the data points shows. Figure (3.10f) shows a greater dispersion of the points (i.e. greater variance) and no temporal relation whatsoever.

In order to find a lower bound for the attractor's fractal dimension, D_f, the correlation dimension was computed for the observed data, as well as the reference systems. This correlation dimension was computed with the Grassberger-Procaccia algorithm described beforehand. The embedding dimension that was used for computing the observed data and the synthetic time series was $m = 2$ and the time delay τ was varied starting from $\tau = 1$ to $\tau = 50$. The



(a) Tent map (b) Logistic map (c) Lorenz system

(d) Uniform dist. (e) Persistent r.w. (f) Antipersistent r.w.

Figure 3.10: Phase space of chaotic systems (upper half) and stochastic processes (lower half).

multiple values were averaged and the resulting D_C's are presented in Table (3.2). It is worth mentioning that these values are close to those presented in Tab. (2.2). Both sets of results show a correlation dimension smaller than 2.

Table 3.2: Correlation dimension of Accumulated Cyclone Energy in the Northwest Pacific and time series generated by chaotic and stochastic systems ($N = 72$).

	ACE
D_C	1.77 ± 0.15

	Tent	Logistic	Lorenz
D_C	1.76 ± 0.18	1.45 ± 0.17	1.71 ± 0.40

	Uniform	Persistent	Antipersistent
D_C	1.63 ± 0.17	1.62 ± 0.15	1.73 ± 0.17

Chapter 4

Discussion

The aim of this project was to characterize the dynamics of tropical cyclone activity in the Northwest Pacific basin through the proxy variable known as Accumulated Cyclone Energy. This analysis was performed on the variable in the form of a time series comprising 72 data points. A parallel aim was to assess the performance of nonlinear time series analysis tools on synthetic (deterministic and stochastic) and observational data. A comparison of the performance of these tools can also be made with the study of data sets of different lengths (in this case 72 and 1000 data points).

Statistical analysis

The first statistics tool used to study the ACE data was the autocorrelation function. The autocorrelogram shows a lack of meaningful correlations, except for significant correlations for a time lag of 16 an 17 years. Comparing the autocorrelogram of the ACE with the synthetic data, we can confirm that the data present no trend and that the next value becomes basically unpredictable for time lags greater than 1 year [65].

A moving average with a window length of 30 years shows a subtle downward trend for the Northwest Pacific, which is in accordance with results obtained by studies focusing on this basin [31], [34]. This downward trend, together with the significant autocorrelation of 16 years may be connected to the El Niño Southern Oscillation (ENSO) phenomenon, which affects global tropical cyclone activity on a multi-decadal scale. The effects of ENSO on tropical cyclone activity in the Northwest Pacific have already been studied previously [31], [35], [37], [86].

The histogram of the data, when using the recommended number of bins for a time series of length N (i.e. $\sqrt{N}$), resembles a Gaussian distribution, albeit with a higher probability for mild and extreme tropical cyclone seasons. When we use 11 bins for the histogram representation, we can see what appears to be a bimodal distribution [87]. This is a sign that building a histogram and extracting the Probability Density Function from it introduces artifacts which may lead to different interpretations.

The ranking representation shows three main outliers in the observed data, as was already observed directly from the time series. Apart from these extreme values, a main cluster that is almost uniformly distributed can be seen. This can be compared to the time series generated by random numbers from a uniform distribution shown in Fig. (3.8c). The rank-ordering representation, as opposed to the histogram, shows the data in their *raw form*, which does not introduce artifacts coming from the use of specific parameters. This makes this representation more appropriate for short data sets [88]. As to the presence of fractal behavior in the data, the ranking representation shows no sign of power-laws in the log-log representation.

Dynamical analysis

The dynamical part of the analysis was more difficult to perform, as larger data sets are normally required to use nonlinear analysis tools. The first analysis, namely computing the Hurst exponent, requires at least ~ 100 data points [89]. The algorithm that was used for computing this quantity did not behave homogeneously with the input data (both for $N = 1000$ and $N = 72$ data points). The conclusions that can be drawn from the Hurst exponent are limited and should be taken cautiously because of the errors that were presented by the algorithm. Both the phase space reconstruction and the ranking representation, however, support the idea that the data are antipersistent. This alleged antipersistence may point at the presence of a negative feedback in tropical cyclone seasons. For example, a specially active season reduces the available thermal energy of the sea and prevents the following season to be as active [90].

As for the methods that work on phase space, the results are ambiguous as to whether or not there is an attractor present. As pointed out in the results section, the phase space reconstruction shows that the trajectories follow what could be interpreted as an attractor with two main clusters. For instance, the phase space reconstruction of the observed data could be vaguely compared to that of the generated data from a random walk that showcases antipersistent behavior.

The dimension of the supposed attractor was found to be small, $D_C = 1.77 \pm 0.15$, which is in accordance to the presence of a low-dimensional attractor. This would mean that at least two variables are needed to describe the dynamics of the system. The low dimensionality of climate and weather variables has been already reported by numerous studies [91]–[95]. Again, estimating the dimension of an attractor reliably requires around $10^{2+0.4D}$ data points [64]. This means that the size of the observed data set cannot yield a reliable estimate of the attractor.

To further probe the idea that the observed data behave like a random walk with antipersistent behavior, surrogate data of the observational records are generated and the analysis tools of this project are applied on it. This analysis is presented in the next section.

Surrogate data

As mentioned in the previous sections, the observed Accumulated Cyclone Energy data are similar to what we would obtain from a stochastic process which exhibits antipersistent behavior. In order to test this conclusion we can perform the same analysis on synthetic data designed to mimic the statistical properties of the observed data, but with an explicit stochastic nature [96].

These surrogate data are obtained using an AutoRegressive Moving Average model, known as an ARMA(p, q) model, given by the expression:

$$y_t = \alpha + \sum_{k=1}^{p} \phi_k y_{t-k} + w_t + \sum_{k=1}^{q} \theta_k w_{t-k}\,, \tag{4.1}$$

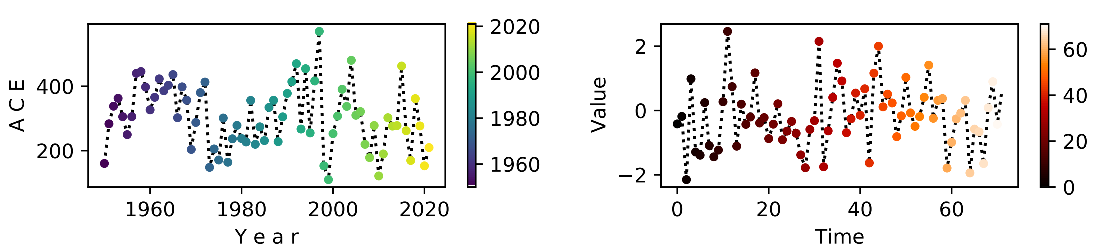
where p and q are the order of the autoregressive and moving average parts respectively. Here, ϕ_i and θ_i are constants and w_t are white noise variables. Finally, α is given by

$$\alpha = \mu \left(1 - \sum_{k=1}^{p} \phi_k \right) . \tag{4.2}$$

In the autoregressive part of the model, the next value of y is a linear combination of its previous values. The moving average part, in turn, represents y as the result of smoothing $q+1$ independent random variables. Here, there is no direct dependence between successive values of y [64].

The first part for generating surrogate data involves estimating the parameters for the ARMA model. Using the Python library `tqdm` and using the Akaike Information Criterion (AIC), we obtained the optimal model: ARMA(0,1), which means only a moving average model of order $q = 1$ is necessary to reproduce the features of the observed data. We then obtained $\theta = 0.3062$. With this information we implemented the model and obtained a time series with $N = 72$ data points to maintain the same length as the observational data.

Figure (4.1) shows the time series of both the Accumulated Cyclone Energy in the Northwest Pacific and the surrogate data generated from the original data. As it can be seen at first glance,

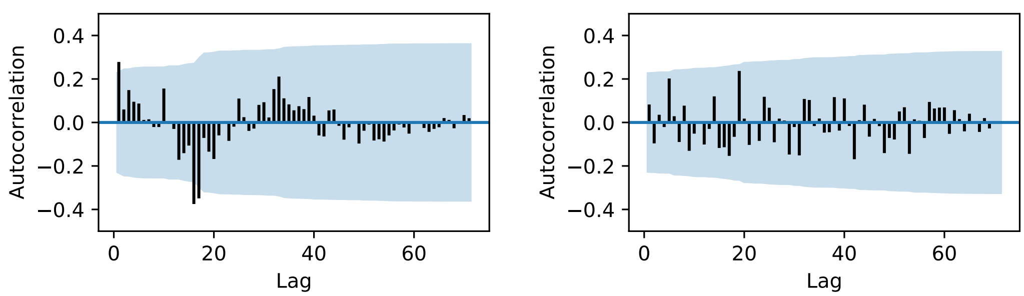


(a) ACE WP　　(b) MA(1)

Figure 4.1: Time series of: a) Accumulated Cyclone Energy in the Northwest Pacific and b) Moving Average model of order 1.

both time series are similar. The only notable difference is that the ACE data shows a smaller variance in the first years compared to the variance seen after the year 1990. This does not necessarily mean there has been a change in TC activity in this basin. This difference in variance may be due to the changes in the way the data were obtained [30], [97].

The autocorrelation function, presented in Fig. (4.2), shows that the surrogate data, in contrast to the observational record, does not have significant autocorrelation values. This means that we can see the presence of randomness in the surrogate data, but not in the observed Accumulated Cyclone Energy.



(a) ACE WP

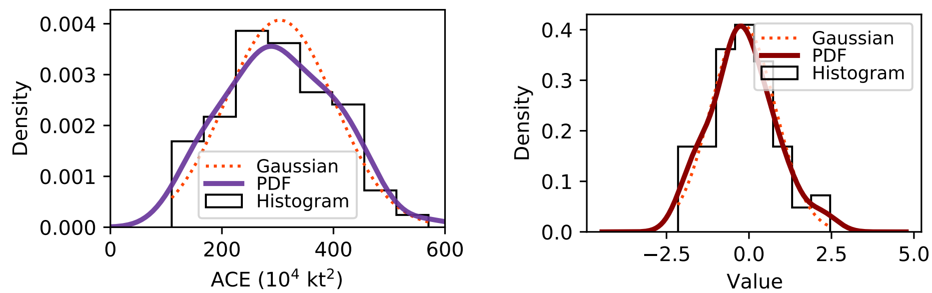
(b) MA(1)

Figure 4.2: Plot of the autocorrelation function of: a) Accumulated Cyclone Energy in the Northwest Pacific and b) Moving Average model of order 1.

The histograms of the ACE data and the synthetic data, shown in Fig. (4.3), clearly show that the observational data have an excess of occurrences for both weak and strong tropical cyclones when compared to a Gaussian. The surrogate data, on the other hand, are closer to the Gaussian benchmark. This reinforces the fact that a stochastic process is at play for the surrogate data.

The ranking representation of both data sets, presented in Fig. (4.4), shows that their distributions are very similar, which means the modeling successfully captured the distribution of values of the experimental data.

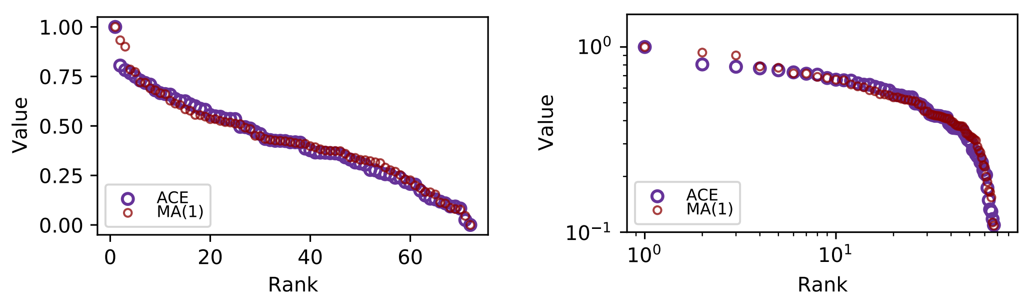
The Hurst exponent obtained for the synthetic data, however, differs from that obtained for the ACE data. We obtained $H = 0.629$ with the linear fit and $H = 0.696 \pm 0.068$ with the nonlinear fit. This means that, according to this computation, the synthetic data show



(a) ACE WP

(b) MA(1)

Figure 4.3: Histogram of: a) Accumulated Cyclone Energy in the Northwest Pacific and b) Moving Average model of order 1.



(a) Linear

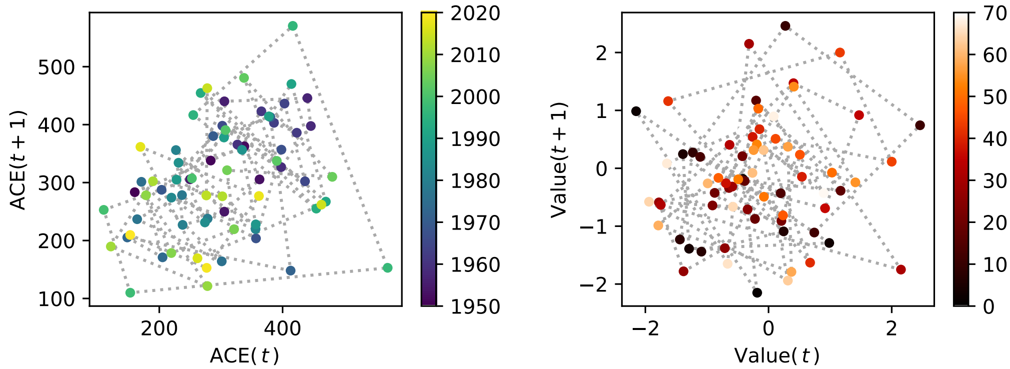
(b) Log-log

Figure 4.4: Rank profile of Accumulated Cyclone Energy in the Northwest Pacific and Moving Average model of order 1 in a) linear plot and b) Log-log representation.

persistent behavior, which is in contrast to the antipersistent behavior found in the ACE data.

Lastly, the phase space reconstructions of both data sets show a very similar distribution of points in the lagged time. What was originally thought to be an attractor with two apparent centroids in the observational data shows no clear structure in the context of this new figure (Fig. 4.5b). As a conclusion of this figure, we can say that we lack data points to point confidently to a defined structure in the data.

From these surrogate data we can draw the conclusion that we can, in principle, model the observational data set as a moving-average stochastic process. Nevertheless, the autocorrelation plot points at the surrogate data as having a clear stochastic nature, whereas the real-world data show some significant autocorrelations that stray from a stochastic description. Additionally, the histogram of both data sets showed a difference in occurrences of both high and low values



(a) ACE

(b) MA(1)

Figure 4.5: Phase space reconstruction of: a) Accumulated Cyclone Energy in the Northwest Pacific and b) Moving Average model of order 1.

of the parameter.

This chapter went over the results and gave an interpretation of them and their relevance to the present study. Additionally, a statistical model called AutoRegressive Moving Average model was used to generate surrogate data and compare it with the observational data set. It was found that the experimental data can be described with an ARMA model, but with some caveats. The following chapter draws from these observations and condenses them into a series of conclusions regarding the validity and usefulness of the results. Furthermore, we outline possible directions for future work within the context of global warming and other pressing issues.

Chapter 5

Conclusion

In this thesis project we studied the annual tropical cyclone activity in the Northwest Pacific. We performed two kinds of data analysis on the time series of observed Accumulated Cyclone Energy. The first kind was based on statistical tools and the second relied on tools developed to study nonlinear dynamical systems.

From the analysis tools we applied to the observational data, as well as the synthetic data sets, the main results are:

- The probability density function suggested the presence of two dynamical states.
- The ranking representation of the data revealed an almost homogeneous distribution similar to that of the antipersistent random walk.
- According to the Hurst exponent and phase space reconstruction, the Accumulated Cyclone Energy shows antipersistent behavior.
- The phase space reconstruction, although not being able to clearly show a determined structure, does confirm the presence of two dynamical states.

Additionally, in accordance with the results reported by previous works, the data showed no signs of increasing activity over the past two decades as compared to when reliable measurements started.

It is important to take this list of conclusions carefully as there are a number of limitations associated to the data and methods used. First, the data set used is inhomogeneous. The introduction of satellite observations in the 1970's improved the quality and precision of observations, but this means that pre-satellite observations are not necessarily consistent with modern data. The second limitation has to do with sample size and the tools used to analyze the data. Statistical and dynamical analysis tools usually require larger number of data points to give reliable results, yet the data set used to carry this work contains only 72 data points. This means that the results that can be obtained with these tools are already limited in the precision and the conclusions that can be reached with them may be biased.

A natural progression of this work is to extend the same analysis to different tropical cyclone basins. Additionally, a variation or extension of these tools could aid in the investigation of the effects of global warming on tropical cyclone activity. Such a connection is being pursued for the North Atlantic basin, where tropical cyclone activity has been shown to be increasing in the last decades [98] (This manuscript has been submitted to the International Journal of Modern Physics C and is currently under revision).

Bibliography

[1] S. Gupta, N. Mastrantonas, C. Masoller, and J. Kurths, “Perspectives on the importance of complex systems in understanding our climate and climate change—The Nobel Prize in Physics 2021,” en, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 32, no. 5, p. 052 102, May 2022.

[2] K. Emanuel, *Divine wind: the history and science of hurricanes*. Oxford University Press, 2005.

[3] V. Lucarini, “Modeling Complexity: The Case of Climate Science,” in *Models, Simulations, and the Reduction of Complexity*, U. Gähde, S. Hartmann, and J. H. Wolf, Eds., De Gruyter, Nov. 2013, pp. 229–254.

[4] K. Morupsisi, “A conceptual model of climate change using dynamical systems theory,” Ph.D. dissertation, 2020.

[5] R. Anthes, *Tropical Cyclones: Their Evolution, Structure, and Effects*. Springer, 1983, vol. 19, Publication Title: Eos, Transactions American Geophysical Union Issue: 4 ISSN: 0096-3941.

[6] S. Eberenz, S. Lüthi, and D. N. Bresch, “Regional tropical cyclone impact functions for globally consistent risk assessments,” *Natural Hazards and Earth System Sciences*, vol. 21, no. 1, pp. 393–415, 2021, Publisher: Copernicus GmbH.

[7] J. B. Elsner and T. H. Jagger, *Hurricanes and climate change*, eng. New York: Springer, 2009.

[8] D. L. Swain, D. Singh, D. Touma, and N. S. Diffenbaugh, “Attributing Extreme Events to Climate Change: A New Frontier in a Warming World,” *One Earth*, vol. 2, no. 6, pp. 522–527, 2020, Publisher: The Authors.

[9] J. M. Done, G. M. Lackmann, and A. F. Prein, “The response of tropical cyclone intensity to changes in environmental temperature,” en, *Weather and Climate Dynamics*, vol. 3, no. 2, pp. 693–711, Jul. 2022.

[10] J. B. Elsner and T. H. Jagger, “United States and Caribbean tropical cyclone activity related to the solar cycle,” en, *Geophysical Research Letters*, vol. 35, no. 18, p. L18705, Sep. 2008.

[11] L. Wu, H. Zhao, C. Wang, J. Cao, and J. Liang, “Understanding of the Effect of Climate Change on Tropical Cyclone Intensity: A Review,” *Advances in Atmospheric Sciences*, vol. 39, no. 2, pp. 205–221, Feb. 2022.

[12] IPCC, *Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 2021, vol. In Press, Type: Book.

[13] Á. Corral, "Tropical cyclones as a critical phenomenon," *Hurricanes and Climate Change-Volume 2*, pp. 81–99, 2010, arXiv: 1103.4654 ISBN: 9789048195091.

[14] M. Pluviose, "A Remarkable Use of Energetics by Nature: The Chaotic System of Tropical Cyclones," Tech. Rep., 2018, Publication Title: International Journal of Applied Environmental Sciences Volume: 13 Issue: 8, pp. 731–754.

[15] Z. J. Lau, T. Pham, S. H. A. Chen, and D. Makowski, "Brain Entropy, Fractal Dimensions and Predictability: A Review of Complexity Measures for EEG in Healthy and Neuropsychiatric Populations," en, *European Journal of Neuroscience*, ejn.15800, Aug. 2022.

[16] D. Rind, "Complexity and climate," *Science*, vol. 284, no. 5411, pp. 105–107, 1999.

[17] C. W. Snyder, M. D. Mastrandrea, and S. H. Schneider, "The Complex Dyanmics of the Climate System," in *Philosophy of Complex Systems*, Elsevier, 2011, pp. 467–505.

[18] R. N. Jones and J. H. Ricketts, "Climate as a complex, self-regulating system," 2021.

[19] K. Emanuel, "A Statistical Analysis of Tropical Cyclone Intensity," en, *Monthly Weather Review*, vol. 128, no. 4, pp. 1139–1152, Apr. 2000.

[20] E. Madrigal-Solís, *Estudio de la Energía Ciclónica Acumulada en el Pacífico Noroeste con herramientas de sistemas complejos*, Zacatecas, Mexico, Oct. 2022.

[21] G. B. Anderson, A. Schumacher, J. M. Done, and J. W. Hurrell, "Projecting the Impacts of a Changing Climate: Tropical Cyclones and Flooding," en, *Current Environmental Health Reports*, vol. 9, no. 2, pp. 244–262, Apr. 2022.

[22] B. Pei, W. Pang, F. Y. Testik, N. Ravichandran, and F. Liu, "Selection of hazard-consistent hurricane scenarios for regional combined hurricane wind and flood loss estimation," en, *Natural Hazards*, vol. 91, no. 2, pp. 671–696, Mar. 2018.

[23] K. Emanuel, "Contribution of tropical cyclones to meridional heat transport by the oceans," en, *Journal of Geophysical Research: Atmospheres*, vol. 106, no. D14, pp. 14 771–14 781, Jul. 2001.

[24] R. E. Hart, "An inverse relationship between aggregate northern hemisphere tropical cyclone activity and subsequent winter climate: TROPICAL CYCLONES AND WINTER CLIMATE," en, *Geophysical Research Letters*, vol. 38, no. 1, n/a–n/a, Jan. 2011.

[25] R. L. Sriver, "Observational evidence supports the role of tropical cyclones in regulating climate," en, *Proceedings of the National Academy of Sciences*, vol. 110, no. 38, pp. 15 173–15 174, Sep. 2013.

[26] K. Emanuel, "Tropical Cyclones," *Annual Review of Earth and Planetary Sciences*, vol. 31, no. 1, pp. 75–104, May 2003.

[27] R. S. Cerveny, J. Lawrimore, R. Edwards, and C. Landsea, "Extreme Weather Records," en, *Bulletin of the American Meteorological Society*, vol. 88, no. 6, pp. 853–860, Jun. 2007.

[28] M. T. Montgomery and B. F. Farrell, “Tropical Cyclone Formation,” en, *Journal of the Atmospheric Sciences*, vol. 50, no. 2, pp. 285–310, Jan. 1993.

[29] P. Taalas, Ed., *Global Guide to Tropical Cyclone Forecasting*. Geneva: World Meteorological Organization, 2019.

[30] K. R. Knapp, M. C. Kruk, D. H. Levinson, H. J. Diamond, and C. J. Neumann, “The International Best Track Archive for Climate Stewardship (IBTrACS),” *Bulletin of the American Meteorological Society*, vol. 91, no. 3, pp. 363–376, Mar. 2010, Publisher: American Meteorological Society.

[31] H. Ramsay, “The Global Climatology of Tropical Cyclones,” in *Oxford Research Encyclopedia of Natural Hazard Science*, Oxford University Press, May 2017.

[32] T. Knutson, S. J. Camargo, J. C. L. Chan, *et al.*, “Tropical Cyclones and Climate Change Assessment: Part I: Detection and Attribution,” *Bulletin of the American Meteorological Society*, vol. 100, no. 10, pp. 1987–2007, Oct. 2019.

[33] K. Emanuel, “Atlantic tropical cyclones downscaled from climate reanalyses show increasing activity over past 150 years,” *Nature Communications*, vol. 12, no. 1, Dec. 2021, Publisher: Nature Research.

[34] P. J. Klotzbach, K. M. Wood, C. J. Schreck, S. G. Bowen, C. M. Patricola, and M. M. Bell, “Trends in Global Tropical Cyclone Activity: 1990–2021,” *Geophysical Research Letters*, vol. 49, no. 6, Mar. 2022.

[35] S. J. Camargo and A. H. Sobel, “Western North Pacific Tropical Cyclone Intensity and ENSO,” *Journal of Climate*, vol. 18, no. 15, pp. 2996–3006, Aug. 2005.

[36] I.-I. Lin, S. J. Camargo, C. M. Patricola, *et al.*, “ENSO and Tropical Cyclones,” en, in *Geophysical Monograph Series*, M. J. McPhaden, A. Santoso, and W. Cai, Eds., 1st ed., Wiley, Nov. 2020, pp. 377–408.

[37] J. Song, P. J. Klotzbach, N. Wei, and Y. Duan, “Modulation of western North Pacific tropical cyclone formation by central Pacific ENSO on decadal and interannual timescales,” *International Journal of Climatology*, Jul. 2022, Publisher: Wiley.

[38] K. A. Emanuel, “An Air-Sea Interaction Theory for Tropical Cyclones. Part I: Steady-State Maintenance,” *Journal of the Atmospheric Sciences*, vol. 43, no. 6, pp. 585–605, Mar. 1986.

[39] K. Emanuel, “Hurricanes, Tempests in a greenhouse,” en, *Physics Today*, vol. 59, no. 8, pp. 74–75, Aug. 2006.

[40] M. Bister and K. A. Emanuel, “Dissipative heating and hurricane intensity,” en, *Meteorology and Atmospheric Physics*, vol. 65, no. 3-4, pp. 233–240, 1998.

[41] H. Riehl, “On the formation of typhoons,” en, *Journal of Meteorology*, vol. 5, no. 6, pp. 247–265, Dec. 1948.

[42] W. M. Gray, “Hurricanes: Their formation, structure and likely role in the tropical circulation,” *Meteorology over the tropical oceans*, vol. 155, p. 218, 1979, Publisher: Royal Meteorological Society Bracknell.

[43] K. Emanuel, “Environmental Factors Affecting Tropical Cyclone Power Dissipation,” *Journal of Climate*, vol. 20, no. 22, pp. 5497–5509, Nov. 2007.

[44] T. R. Knutson and J. Ploshay, “Sea Level Pressure Trends: Model-Based Assessment of Detection, Attribution, and Consistency with CMIP5 Historical Simulations,” *Journal of Climate*, vol. 34, no. 1, pp. 327–346, Jan. 2021.

[45] J.-H. Kim, K.-B. Kim, and H.-Y. Chang, “Solar Influence on Tropical Cyclone in Western North Pacific Ocean,” *Journal of Astronomy and Space Sciences*, vol. 34, no. 4, pp. 257–270, Dec. 2017.

[46] I. Roy, *Climate Variability and Sunspot Activity*, ser. Springer Atmospheric Sciences. Cham: Springer International Publishing, 2018.

[47] M. Singh, R. Krishnan, B. Goswami, *et al.*, “Fingerprint of volcanic forcing on the ENSO–Indian monsoon coupling,” en, *Science Advances*, vol. 6, no. 38, eaba8164, Sep. 2020.

[48] A. P. Cracknell and C. Varotsos, *Understanding global climate change: modelling the climatic system and human impacts*, Second edition. Boca Raton: CRC Press, 2021.

[49] H. Murakami, “Substantial global influence of anthropogenic aerosols on tropical cyclones over the past 40 years,” en, *Science Advances*, vol. 8, no. 19, eabn9493, May 2022.

[50] W. M. Gray, “Forecast of Atlantic seasonal hurricane activity for 1985.,” *Colorado State University*, 1988.

[51] A. Waple, J. H. Lawrimore, M. S. Halpert, *et al.*, “Climate assessment for 2001,” *Bulletin of the American Meteorological Society*, vol. 83, pp. 938–938, 2002.

[52] L. Fridman, “A Comparative Study of Atlantic Seasonal Accumulated Cyclone Energy and Tropical Cyclone Counts,” Ph.D. dissertation, North Carolina State University, 2022.

[53] E. Roca-Flores and G. G. Naumis, “Northeast Pacific annual accumulated cyclonic energy rank-profile,” *International Journal of Modern Physics C*, vol. 32, no. 6, pp. 1–11, 2021, ISBN: 0129183121.

[54] H. M. Fry, “Complex Systems Dynamics,” en, in *Dynamical and Complex Systems*, WORLD SCIENTIFIC (EUROPE), Feb. 2017, pp. 81–120.

[55] M. Mitchell, *Complexity: a guided tour*. Oxford [England] ; New York: Oxford University Press, 2009, OCLC: ocn216938473.

[56] S. Thurner, P. Klimek, and R. Hanel, *Introduction to the Theory of Complex Systems*, en. Oxford University Press, Nov. 2018, vol. 1.

[57] M. Gell-Mann, “Simplicity and Complexity in the Description of Nature,” *Engineering and Science*, vol. 51, no. 3, pp. 2–9, 1988.

[58] H. Sayama, *Introduction to the modeling and analysis of complex systems*, eng, Deluxe color ed. Geneseo, NY: Open SUNY Textbooks, Milne Library, 2015.

[59] M. Fernández-Sanjuan, “Dinámica No Lineal, Teoría del Caos y Sistemas Complejos: Una perspectiva histórica,” *Cienc.Exact.Fís.Nat. (Esp)*, vol. 109, no. 2, pp. 107–126, 2016.

[60] Y. Karaca, "Theory of complexity, origin and complex systems," en, in *Multi-Chaos, Fractal and Multi-Fractional Artificial Intelligence of Different Complex Systems*, Elsevier, 2022, pp. 9–20.

[61] P. Grassberger and I. Procaccia, "Measuring the strangeness of strange attractors," *Physica D: Nonlinear Phenomena*, vol. 9, no. 1, pp. 189–208, 1983, Publisher: Science Direct.

[62] P. Dirac, "The physical interpretation of the quantum dynamics," en, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 113, no. 765, pp. 621–641, Jan. 1927.

[63] E. N. Lorenz, "Deterministic Nonperiodic Flow," en, *Journal of the Atmospheric Sciences*, vol. 20, no. 2, pp. 130–141, Mar. 1963.

[64] C. R. Shalizi, "Methods and Techniques of Complex Systems Science: An Overview," en, in *Complex Systems Science in Biomedicine*, T. S. Deisboeck and J. Y. Kresh, Eds., Series Title: Topics in Biomedical Engineering International Book Series, Boston, MA: Springer US, 2006, pp. 33–114.

[65] J. C. Sprott, *Chaos and time-series analysis*. Oxford ; New York: Oxford University Press, 2003, OCLC: ocm50525812.

[66] R. M. May, "Simple mathematical models with very complicated dynamics," en, *Nature*, vol. 261, no. 5560, pp. 459–467, Jun. 1976.

[67] G. Boeing, "Visual Analysis of Nonlinear Dynamical Systems: Chaos, Fractals, Self-Similarity and the Limits of Prediction," *Systems*, vol. 4, no. 4, p. 37, 2016, arXiv: 1608.04416.

[68] E. N. Lorenz, "Maximum Simplification of the Dynamic Equations," en, *Tellus*, vol. 12, no. 3, pp. 243–254, Aug. 1960.

[69] C. Sparrow, *The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors*, ser. Applied Mathematical Sciences. New York, NY: Springer New York, 1982, vol. 41.

[70] O. Fontanelli Espinoza, R. L. Mansilla Corona, and P. E. Miramontes Vidal, "Distribuciones de probabilidad en las ciencias de la complejidad: Una perspectiva contemporánea," *INTER DISCIPLINA*, vol. 8, no. 22, p. 11, Aug. 2020.

[71] J. Frost, *Introduction to statistics: an intuitive guide for analyzing data and unlocking discoveries*, eng. State College: Statistics by Jim Publishing, 2020, OCLC: 1276800559.

[72] D. Sornette, *Critical phenomena in natural sciences: chaos, fractals, selforganization and disorder: concepts and tools*. Springer Science & Business Media, 2006.

[73] G. K. Zipf, *Human behavior and the principle of least effort: an introduction to human ecology*, eng. Mansfield Centre, CT: Martino Publishing, 1949.

[74] G. Iñiguez, C. Pineda, C. Gershenson, and A.-L. Barabási, "Dynamics of ranking," *Nature Communications*, vol. 13, no. 1, p. 1646, Dec. 2022, Publisher: Nature Research.

[75] K. Fraedrich, "Predictability: Short-and long-term memory of the atmosphere," *Chaos in Geophysical Flows*, pp. 63–104, 2003.

[76] T. Gneiting and M. Schlather, "Stochastic Models That Separate Fractal Dimension and the Hurst Effect," *SIAM Review*, vol. 46, no. 2, pp. 269–282, Jan. 2004, Publisher: SIAM.

[77] H. E. Hurst, “Long-Term Storage Capacity of Reservoirs,” *Transactions of the American Society of Civil Engineers*, vol. 116, no. 1, pp. 770–799, 1951.

[78] A. López-Lambraño, E. Carrillo-Yee, C. Fuentes, A. López-Ramos, and M. López-Lambraño, “Una revisión de los métodos para estimar el exponente de Hurst y la dimensión fractal en series de precipitación y temperatura,” *Revista Mexicana de Fisica*, vol. 63, no. 3, 2017.

[79] M. Tarnopolski, “Correlation between the Hurst exponent and the maximal Lyapunov exponent: Examining some low-dimensional conservative maps,” en, *Physica A: Statistical Mechanics and its Applications*, vol. 490, pp. 834–844, Jan. 2018.

[80] J. de Pedro-Carracedo, D. Fuentes-Jimenez, A. M. Ugena, and A. P. Gonzalez-Marcos, “Phase Space Reconstruction from a Biological Time Series: A Photoplethysmographic Signal Case Study,” *Applied Sciences*, vol. 10, no. 4, p. 1430, Feb. 2020, Publisher: MDPI AG.

[81] N. Boers, J. Kurths, and N. Marwan, “Complex systems approaches for Earth system data analysis,” *Journal of Physics: Complexity*, vol. 2, no. 1, Mar. 2021, Publisher: Web Portal IOP.

[82] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*. Cambridge University Press, Nov. 2003.

[83] F. Takens, “Detecting strange attractors in turbulence,” in *Dynamical systems and turbulence, Warwick 1980*, Springer, 1981, pp. 366–381.

[84] J. Theiler, “Estimating fractal dimension,” *Journal of the Optical Society of America A*, vol. 7, no. 6, p. 1055, Jun. 1990.

[85] P. Grassberger and I. Procaccia, “Characterization of strange atractors,” *Phys. Rev. Letters*, vol. 50, no. 3, pp. 346–349, 1983.

[86] J. C. L. Chan, “Tropical Cyclone Activity in the Northwest Pacific in Relation to the El Niño/Southern Oscillation Phenomenon,” en, *Monthly Weather Review*, vol. 113, no. 4, pp. 599–606, Apr. 1985.

[87] E. Roca-Flores, G. G. Naumis, E. Madrigal-Solís, and K. Fraedrich, “Typhoon complexity: Northwest Pacific tropical cyclone season complex systems analysis,” en, *International Journal of Modern Physics C*, p. 2 350 045, Oct. 2022.

[88] D. Sornette, L. Knopoff, Y. Y. Kagan, and C. Vanneste, “Rank-ordering statistics of extreme events: Application to the distribution of large earthquakes,” en, *Journal of Geophysical Research: Solid Earth*, vol. 101, no. B6, pp. 13 883–13 893, Jun. 1996.

[89] G. Millán, R. Osorio-Comparán, and G. Lefranc, “Preliminaries on the Accurate Estimation of the Hurst Exponent Using Time Series,” in *2021 IEEE International Conference on Automation/XXIV Congress of the Chilean Association of Automatic Control (ICA-ACCA)*, arXiv:2103.02091 [cs, eess], Mar. 2021, pp. 1–8.

[90] K. B. Karnauskas, L. Zhang, and K. A. Emanuel, “The Feedback of Cold Wakes on Tropical Cyclones,” *Geophysical Research Letters*, vol. 48, no. 7, Apr. 2021.

[91] K. Fraedrich, “Estimating the dimensions of weather and climate attractors,” *Journal of Atmospheric Sciences*, vol. 43, no. 5, pp. 419–432, 1986.

[92] P. K. Pal, "Cyclone Track Prediction over the North Indian Ocean," *Monthly Weather Review*, vol. 119, no. 12, pp. 3095–3098, Dec. 1991.

[93] H. N. Shirer, C. J. Fosmire, R. Wells, and L. Suciu, "Estimating the correlation dimension of atmospheric time series," *Journal of the Atmospheric Sciences*, vol. 54, no. 1, 1997.

[94] S. D. Aberson, "Five-Day Tropical Cyclone Track Forecasts in the North Atlantic Basin," *Weather and Forecasting*, vol. 13, no. 4, pp. 1005–1015, Dec. 1998.

[95] B. Rojo-Garibaldi, D. A. Salas-de-León, M. A. Monreal-Gómez, N. L. Sánchez-Santillán, and D. Salas-Monreal, "Nonlinear analysis of the occurrence of hurricanes in the Gulf of Mexico and the Caribbean Sea," *Nonlinear Processes in Geophysics*, vol. 25, no. 2, pp. 291–300, Apr. 2018, Publisher: Copernicus GmbH.

[96] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. Doyne Farmer, "Testing for nonlinearity in time series: The method of surrogate data," en, *Physica D: Nonlinear Phenomena*, vol. 58, no. 1-4, pp. 77–94, Sep. 1992.

[97] S. S. Chand, K. J. E. Walsh, S. J. Camargo, *et al.*, "Declining tropical cyclone frequency under global warming," *Nature Climate Change*, vol. 12, no. 7, pp. 655–661, Jul. 2022.

[98] E. Roca-Flores, G. G. Naumis, E. Madrigal-Solís, K. Fraedrich, and E. Torres, "Hurricane season complexity: The case of North Atlantic tropical cyclones," Unpublished, 2023.

Appendix A

Published article cover page

International Journal of Modern Physics C
(2023) 2350045 (19 pages)

DOI: 10.1142/S0129183123500456

Typhoon complexity: Northwest Pacific tropical cyclone season complex systems analysis

Elio Roca-Flores[*,¶], Gerardo G. Naumis[†], Erick Madrigal-Solís[‡,§] and Klaus Fraedrich[§]

[*]*Doctorado en Ciencias de la Ingeniería*
Instituto Tecnológico de Aguascalientes
Av. López Mateos 1081 Oriente, Fracc. Bona Gens
CP 20256, Aguascalientes, Ags., México

[†]*Departmento de Sistemas Complejos*
Instituto de Física, Apdo. Postal 20-364
UNAM, CP 04510, CDMX, México

[‡]*Licenciatura en Física, Universidad de Guadalajara, Blvd.*
Marcelino García Barragán 1421, Col. Olímpica
CP 44430, Guadalajara, Jal., México

[§]*Department of The Atmosphere in the Earth System*
Max Planck Institute for Meteorology, 20146, Hamburg, Germany
[¶]*elioroca@gmail.com*

Received 7 July 2022
Accepted 26 August 2022
Published

The complex nature of tropical cyclones (TCs) has been recognized in a vast literature yet only few works perform complex systems diagnostics to understand their dynamics. This is especially important in order to study the effects of global warming on TC hazards. Here, such analysis is performed from a data-driven perspective using statistical and nonlinear dynamics diagnostics to the annual Accumulated Cyclonic Energy (ACE) data over the most active basin, the Northwest Pacific, from the years 1950 to 2021. The best quality data period, from 1984 to 2021, is also considered for a separate analysis in order to test the possible differences due to the data acquisition process. The following results are obtained: (i) The use of mobile windows shows a