
Lionel Blankson Amanfu
Embry-Riddle Aeronautical University - Daytona Beach

Follow this and additional works at: https://commons.erau.edu/db-theses

Part of the Applied Behavior Analysis Commons

https://commons.erau.edu/db-theses/5

This thesis is brought to you for free and open access by Embry-Riddle Aeronautical University – Daytona Beach at ERAU Scholarly Commons. It has been accepted for inclusion in the Theses - Daytona Beach collection by an authorized administrator of ERAU Scholarly Commons. For more information, please contact commons@erau.edu.
AN INVESTIGATION OF DETERMINISM AND CHAOTIC BEHAVIOR IN FLIGHT PERFORMANCE DATA: A CHAOS THEORY AND NONLINEAR TIME SERIES ANALYSIS APPROACH

By

LIONEL BLANKSON AMANFU
B.S., Embry-Riddle Aeronautical University, 2005

A Thesis Submitted to the
Department of Human Factors & Systems
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Human Factors & Systems

Embry-Riddle Aeronautical University
Daytona Beach, Florida
Summer 2008
AN INVESTIGATION OF DETERMINISM AND CHAOTIC BEHAVIOR IN FLIGHT PERFORMANCE DATA: A CHAOS THEORY AND NONLINEAR TIME SERIES ANALYSIS APPROACH

By

Lionel B. Amanfu

This thesis was prepared under the direction of the candidate's thesis committee chair, Dahai Liu, Ph.D., Department of Human Factors & Systems, and has been approved by the members of the thesis committee. It was submitted to the Department of Human Factors & Systems and has been accepted in partial fulfillment of the requirements for the degree of Master of Science in Human Factors & Systems.
Abstract

Human flight performance data were investigated using non-linear time series analysis methods to determine deterministic chaotic behavior in the data. Using a sequence of steps of non-linear methods, four flight performance data were used to investigate for the existence of deterministic chaotic behavior. Results revealed that flight performance data may exhibit chaotic behavior. Results also showed a consistent low determinism value in all the data examined which is the defining characteristic of chaotic behavior. It was also found that the data originated from non-stationary process. The Maximal Lyapunov Exponent (MLE) value which indicate chaotic behavior exist in the data revealed that most of the data examined possessed some traces of deterministic chaotic behavior evident by the low Maximal Lyapunov Exponent value.
# Table of Contents

Abstract ......................................................................................................................... III

Table of Contents ........................................................................................................ IV

List of Figures .............................................................................................................. VI

List of Tables ................................................................................................................. VIII

Acknowledgments .......................................................................................................... IX

Introduction .................................................................................................................. 1

Fundamental Concepts ................................................................................................. 3

- **Time Series** ........................................................................................................ 3
- **Dynamic Systems** ............................................................................................... 4
- **Linear vs. Nonlinear System** ................................................................................ 4
- **Random vs. Deterministic Systems** ..................................................................... 5
- **Random vs. Chaos Behaviors** ........................................................................... 5
- **Stationarity vs. Non-stationarity** ........................................................................ 6

- **Chaos Theory** ..................................................................................................... 7

- **Properties of a Chaotic System** .......................................................................... 7
- **Recurrence plot** .................................................................................................. 9
- **Maximal Lyapunov Exponent (MLE)** ................................................................. 10

Review of Literature ................................................................................................... 11

Summary ....................................................................................................................... 22

Objective of Study ........................................................................................................ 23

Method .......................................................................................................................... 24

- **Flight Performance Data** .................................................................................... 24
- **Simple Noise Reduction** ..................................................................................... 26
- **Estimating embedding time delay $\tau$ by Mutual Information Method** ............ 27
- **Phase Space Reconstruction** .............................................................................. 29
- **Recurrence Plot Analysis** ................................................................................... 30
- **Determinism Test** ............................................................................................... 31
- **Stationarity Test** .................................................................................................. 33
- **Maximal Lyapunov Exponent** ............................................................................ 33
- **Software** .............................................................................................................. 37

Results .......................................................................................................................... 38

Summary of Findings .................................................................................................... 56
List of Figures

Figure 1. Graph of the time series of the lateral path deviation from flight performance data ……. 3
Figure 2. Recurrence plots of simple systems with the x and y axis representing time in the plots. ................................................................. 10
Figure 3: Left graph represent random, right graph represent chaos behavior………………… 11
Figure 4. Time series of 5-minutes counts on Interstate 80................................................ 13
Figure 5. Fourier power spectrum of weekday traffic flow data ........................................ 13
Figure 6. Largest Lyapunov Exponent (LLE) of weekday traffic flow data .................... 13
Figure 7. Fourier power spectra for periodic (left) and chaotic (right) logistic functions …… 14
Figure 8. The studied human electrocardiogram data...................................................... 16
Figure 9. Phase space with optimal embedding parameters (a) Before noise reduction, (b) After noise reduction .......................................................................................... 17
Figure 10. Calculation of the Maximal Lyapunov Exponent ............................................ 18
Figure 11. Recording of human gait ................................................................................. 19
Figure 12. Sequence of steps to determine the presence of chaotic behavior in data……… 25
Figure 13. Blue Sky Brown Ground Terrain/NASA Tunnel (Baseline/NASA).............. 35
Figure 14. Elevation Based Generic Terrain/NASA Tunnel (EBG/NASA).................... 36
Figure 15. Blue Sky Brown Ground/Pitch Roll Flight Director (Baseline/PRFD) .......... 36
Figure 16. Elevation Based Generic Terrain / Pitch Roll Flight Director (EBG/PRFD) ....... 37
Figure 17. Airspeed Error Data of 7541 data points....................................................... 38
Figure 18. The embedding time delay at τ=28................................................................. 39
Figure 19. Finding the embedding dimension \( m = 5 \) with time delay \( \tau=28 \) with 7541 data series ................................................................. 40
Figure 20a. Embedding space from original airspeed error data series .......................... 40
Figure 20b. Embedding space after cleaning the original airspeed error data series.............. 40

Figure 21. Graph showing the optimal embedding time delay at $\tau=25$ using the ‘clean’ data..... 41

Figure 22. Graph showing the minimal embedding dimension at $m=5$ using the ‘clean data and $\tau=25$................................................................................................................................................................................. 42

Figure 23. Plots of the embedding phase space at different coordinate systems.................... 42

Figure 24. The recurrence plot of 7493 points........................................................................ 43

Figure 25. This shows the approximated vector field for the embedding space with a
deterministic test value of 0.55.............................................................................................. 44

Figure 26. This shows the average cross prediction error of 11.34, with minimal and maximal
cross prediction error of 0.0054 and 21.12 respectively. The color map displays the average
cross-predictions errors independence on different segment combinations................. 45

Figure 27: Maximal Lyapunov Exponent of airspeed error -0.0633 ................................... 46
List of Tables

Table 1 .................................................................................................................. 47
Table 2 .................................................................................................................. 49
Table 3 .................................................................................................................. 51
Table 4 .................................................................................................................. 53
Acknowledgments

I would like to acknowledge the contributions of the following people and express my appreciation to them for their help in completing this thesis.

My appreciation goes to my thesis committee members for their supervision and support.

Dr. Dahai Liu, of the Human Factors and Systems Department at Embry-Riddle Aeronautical University for providing the data used. Also much appreciation goes to him for his continuous support, supervision and guidance throughout my thesis process.

Dr. Albert Boquet of the Human Factors and Systems Department and Dr. Seenith Siva of the Mathematics Department both of Embry-Riddle Aeronautical University for the encouragement and advice.

I would like to show my appreciation to Matjaz Perc of the Department of Physics at University of Maribor, Slovenia for providing me with the software used to analyze the data.

Finally, my sincere gratitude goes to my parents, Alex and Alberta Amanfu as well as my sister Doreen Amanfu and her family for their continuous encouragement in completing this thesis.
Introduction

In most behavioral science experiments, traditional statistics methods such as mean, standard deviation, linear regression and analysis of variance have been used to analyze experimental data. These methods are largely based on the assumptions that sample data are generated randomly and that general conclusions about the population can be made by analyzing the sample data drawn from the population. However, when the data are not random the conclusion from using such methods doesn’t provide enough relevant meaning of the data.

On the other hand, until recently, most experimental data were thought to be linearly related. Linear data analysis has been used in interpreting random behavior or structures in a data set such as the frequency and linear correlation (Kantz and Schreiber, 1997). When the sample data isn’t random, all linear data methods no longer reveal significant meanings. To understand the true behavior (in this case nonlinear behavior) embedded in the data, nonlinear methods are needed to properly interpret the data and perhaps characterize the system (Bassingthwaighte, Liebovitch, and West, 1994a).

For example, most physiology experimental data look random, and by visual inspection it is not apparent if the data possesses random or chaotic behavior. An example of such data is the fluctuations in the rate of the heart and the blood pressure in the arteries (Bassingthwaighte et al., 1994a). It’s been believed that these fluctuations must be produced by a random process, thus if indeed the fluctuations are random, using linear data analysis will provide adequate interpretation of the data. However, if the fluctuations appear chaotic, then it might be the result of a deterministic physiological mechanism. In other words, knowing the fluctuations possess chaotic behavior, different methods such as nonlinear analysis methods must be used to
understand the mechanism of the system and possibly learn how to control the system (Bassingthwaighte et al., 1994a).

Limited research has been done using nonlinear analysis on psychological processes. This is because in the past, most psychological data were thought to possess random patterns therefore linear analysis were used to analyze the data. In behavioral science this is due to firstly, linearity and randomness have always been the assumption of study and secondly, nonlinear analysis requires large amount of data set which most behavior science study lack. But recently, researchers have started using nonlinear analysis methods which have shown some encouraging results.

This thesis applies nonlinear method analysis on human performance data (in particular flight performance) to investigate if the nature of human performance data contains purely random behavior or chaotic behavior.

Nonlinear analysis tools which has been used in other fields to improve the understanding of the data as well as the ability to make future predictions regarding the behavior of the system will be used to investigate this broad categorization such as random, deterministic chaos, linear, and nonlinear (Shelhamer, 2007a) of the experimental data. A well collected real flight performance data is used as the experimental data set for this analysis.
Fundamental Concepts

Before discussing the literature where chaos theory and nonlinear methods have been applied, some fundamental concepts of nonlinear methods and chaos theory are illustrated first.

*Time Series*

Time series are the set of values \( X \) of a measurement of a function in time \( t \) (Bassingthwaighte et al. 1994a). It is measured usually at successive times, spaced at uniform time intervals. A nonlinear analysis method uses time series to understand the underlying theory of the data points.

The analysis of the time series can also be used to make predictions or future events based on past events. A time series can be defined in the following form (Brockwell and Davis, 1991):

\[
X = \{ X_t : t \in T \}
\]

Figure 1 shows the time series of the lateral path deviation of a flight performance data.

![Flight Performance Time series data](image)

**Figure 1.** Graph of the time series of the lateral path deviation from flight performance data
Dynamic systems are defined as the set of variables whose values change with time (Shelhamer, 2007b). The variables define the course of the systems behavior as a function known as state variables. A state variable is defined as the variable that describes the course of the systems behavior as a function of time. Collectively, the variables describe the state of the system. The next few paragraphs talk about the different types of dynamic systems which are relevant for the purpose of this thesis.

Linear vs. Nonlinear System

A linear system is one that responds in a proportional manner to its input (Shelhamer, 2007a). It is defined by two properties, the scaling property and the superposition property. Scaling means that given an input produces a given output, then doubling the size of the input will double the size of the output. So, for any arbitrary scaling (Shelhamer, 2007a):

\[ u(t) \rightarrow y(t) \Rightarrow au(t) \rightarrow ay(t) \]

Superposition means that given one input produces an output, and a different input produces another output, and the sum of the two inputs to the system will produce an output of the sum of the individual outputs:

\[ u_1(t) \rightarrow y_1(t) \]
\[ u_2(t) \rightarrow y_2(t) \]
\[ \Rightarrow [u_1(t) + u_2(t)] \rightarrow [y_1(t) + y_2(t)] \]

A nonlinear system represents a system whose behavior cannot be expressed as a sum of the behaviors of its input. In particular, the behavior of nonlinear system is not subject to the principle of superposition as linear systems are. A nonlinear system is one whose behavior is not simply the sum of its parts or their multiples.
Random vs. Deterministic Systems

A random system is controlled by some extent by chance. A characteristic of such a system is that given complete information of the dynamics and initial conditions of the system, it is not possible to predict precisely the future course of the system (Shelhamer, 2007b). A deterministic system is the opposite of a random system. That is, given perfect knowledge of the dynamics and initial conditions of the system; the future behavior of the system can be predicted at all times (Shelhamer, 2007b).

A random or deterministic system has different dynamic system behaviors. The next few paragraphs talk about the two system behaviors pertinent to this thesis namely random and chaos behaviors.

Random vs. Chaos Behaviors

A random behavior means the behavior of system is unpredictable. To determine the behavior of the data set (time series), we have to acknowledge that noise (error) is present in all physical measurement, and determine if randomness is inherent in the system or in the measurement process (Shelhamer, 2007b). Chaos behavior arises from deterministic systems whose data set is so complex it might be mistaken for randomness (Bassingthwaigte et al., 1994b). Therefore if it is determined through nonlinear analysis methods that a deterministic system produced a data that contains chaotic behavior, it opens up the possibilities of understanding and possibly controlling that system. There are some data sets that appear random but are in fact chaotic (predictable behavior), and this makes it difficult to distinguish between random or chaotic behavior in the data set.
Stationarity vs. Non-stationarity

Most time series techniques such as nonlinear methods assume data is stationary. Data that are described to be stationary means the statistical properties of the time series of the data such as mean, standard deviation, variance and correlation coefficient do not change over time. Therefore, if the human performance data to be examined is stationary, then the statistical properties of the data in the future can be predicted since the statistical properties do not change over time. But if the statistical properties of the time series of the data change over time then the time series is not stationary which is sometimes called non-stationarity.
Chaos Theory

Chaos theory is a branch of nonlinear analysis, which describes the behavior of complex nonlinear dynamics systems under certain specific conditions that is sensitive to initial conditions. Another definition states chaos theory simply as a highly ordered and often simple system whose output is so complex it mimics random behavior (Bassingthwaighte et al., 1994a). This field was introduced by Lorenz (1963) who was studying the dynamics of turbulent flow in fluids (Levy, 1994). Since its discovery, the interest of chaos theory has grown rapidly and has been used in different disciplines to understand important system dynamics such as electrical signals, mechanical systems and recently in physiological systems. Chaos theory allows us to distinguish between random, probabilistic and deterministic system (Shelhamer, 2007b).

Properties of a Chaotic System

The following list describes the essential properties of a chaotic system (Bassingthwaighte et al., 1994b).

1. **Chaotic system is a deterministic system**: This means that the values of the variables that describe the system in the future are determined by the present values.

2. **Chaotic system is described by either difference or differential equations**: means in a difference equation the values of the variables are calculated at discrete steps or in a differential equation, the values of the variables change continuously in time.

3. **Chaotic system has sensitivity to initial condition**: This means the values of the variables after a given time depend on their initial values. Small changes in the initial values may produce very large changes in the latter values.
4. **The values of the variables are not predicable in the long run**: Based on property 3, the values that are calculated as time elapsed will diverge further from the true values based on their exact initial values.

5. **The values of the variable do not take on all possible values**: As time goes by, the values of the variables seem to fluctuate widely but they do not take on all combinations of values; in other words, the values are not random.

6. **Phase space**: It is the space in which all possible states of a system can be represented, with each possible state of the system corresponding to one unique point in the phase space. The state of the system can be totally described by the values of its variables. For instance, if there are n-variables, then the state of the system can be described at a point in an n-dimensional space whose coordinates are the values of the dynamical variables. This n-dimensional space is called phase space. The phase space developed by Poincaré in 1879, showed that it was easier to analyze the dynamic system by determining the topological properties of trajectory which is the line traced out by the point in the phase space than analyzing the time series of the values directly.

7. **Attractor**: It is the geometric limiting set on which all the trajectories eventually find themselves, that is set of points in the phase space to which all the trajectories are attracted.

8. **Embedding Dimension (m or d)**: Embedding dimension simply means how many state variables are present in the data. The appropriate embedding dimension has to be chosen. The embedding dimension has to be large enough that the attractor is properly embedded in the phase space (Shelhamer, 2007a). The method such as false nearest neighbor can be used to determine the appropriate embedding dimension.
9. **Time Delay (τ):** This is the optimal time selected so as to minimize the interaction between the points of the measured series. When the optimal time delay is used, it opens up the attractor by representing the largest representation of the set points in the phase space. An analogy of what the time delay does to a phase space would be the visualizing the construction of a circle by plotting a sine wave against itself delayed by 90 degrees (Riley & Van Orden, 2005).

*Recurrence plot*

Recurrence plot is a graphical display of the spatial correlation in an attractor in terms of time (Shelhmer, 2007c; Eckmann, Kamphorst, and Ruelle, 1987). The graphical display gives insight into whether the data is periodic, deterministic or random. When the graph displays isolated recurrent points the data is random and when a repeating pattern occurs the data is periodic. When the graph displays diagonal line segments the data is deterministic. The graph also gives insight into whether the data originated from a stationary or non stationary process. Non-stationarity is shown on the graph by the decreasing density away from the main diagonal line segment (Shelhamer, 2007c).

An example is shown in Figure 2. The left is a plot of Gaussian white noise with the number of data points (N=1000), the embedding dimension (M=5) and time delay (τ=5). On the left of Figure 2, the data point at all locations has no obvious dynamical pattern. This behavior is an example of a random system. On the right is the plot of a sine wave (N=200, M=5, and τ=5) which displays diagonal line segments a classic example of a deterministic system. In Figure 2 there are strong diagonal lines parallel to the main diagonal which indicates deterministic
properties. The spacing between the diagonal lines reflects the periodicity of the sine wave (Shelhamer, 2007c).

![Figure 2. Recurrence plots of simple systems with the x and y axis representing time in the plots.](Image)

**Maximal Lyapunov Exponent (MLE)**

The Lyapunov Exponent, also referred as largest/Maximal Lyapunov Exponent is the rate of divergence or convergence of the trajectories of the system in the phase space (Shelhamer, 2007b). Lyapunov Exponent is mathematically calculated and it measures the deterministic chaotic behavior of the system. If the Lyapunov Exponent is zero, the system is neutrally stable. If the Lyapunov Exponent is less than zero, the system attracts to a fixed point or a stable periodic orbit. If the Lyapunov Exponent is positive, the system is chaotic and unstable indicating nearby trajectories of the system diverge exponentially (Li and Li, 2006). This means that for any given nearby points, no matter how close they are, the points will diverge to any random separation as the regions in the phase space are eventually visited. A positive Lyapunov Exponent indicates the system exhibits determinism, which is a characteristic of chaotic systems.
Review of Literature

Applications of Chaos Time Series

The literature where application of chaos nonlinear time series analysis has been applied is presented.

Let's start with an example, in Figure 3, two time series are illustrated. By visually inspecting the graphs they look similar. They probably have approximately the same statistical properties and they both look random. However they are very different upon further investigation. The time series on the left is completely random; generated from a random number generator. The time series on the right computed by the function $x_{n+1} = 3.95x_n(1 - x_n)$, though by visual inspection shows the time series is random but in fact the time series is completely deterministic that is, it exhibits chaotic behavior (Bassingthwaighte et al., 1994a). Many physiological systems exhibit chaotic behavior but when visually viewed appear to be random. If indeed these systems possess chaotic behavior, then appropriate analysis should be applied for physiologist to understand the mechanism of the system and learn to control the system (Bassingthwaighte et al., 1994a).

Figure 3 : Left graph represent random, right graph represent chaos behavior
Physiologically chaotic time series is interesting for a number of reasons according to Conrad (1986). Chaos time series has been used to search processes in microorganisms to identify behavior patterns (Bassingthwaighte et al., 1994b). For example, in biological systems studies; chaos time series has been used in trapping organisms that moves about in an unpredictable way compared to organisms that moves in a predictable pattern (Bassingthwaighte et al., 1994b).

Another area where chaos theory has been used to analyze complex systems is the transportation system. The transportation system is a complex system because its current state and future progress are continuously changing due to a number of countless properties of interacting physical and human elements (Frazier and Kockelman, 2004). So to better understand the data generated by transportation systems such as traffic flow, chaotic data analysis is used to analyze the data to distinguish whether the traffic flow data is random or deterministic. The data obtained from the Freeway Performance Measurement Project (PeMS) run by the University of California, Berkeley was collected over five-week interval from noon April 7, 2003 to noon May 12 2003 from an inductive loop detectors embedded on a section of Interstate 80 near Sacramento, California.

With the recommendation of Nair, Liu, Rilett, and Gupta (2001) who observed that weekend and weekday traffic flows were of different patterns, Frazier and Kockelman (2004) focused their study only on the work week from April 14 through 19. Using the five days count data, various techniques were used to check the existence of chaos in the data. Figures 4, 5, 6 illustrates time series of the 5-minutes count on Interstate 80 between April 14 through 19, the Fourier power spectrum of the data, and the largest Lyapunov Exponent of the data respectively.
In Figure 4, the graph displays the data of work week from April 14 through 19. This data was used together with various techniques to check for the existence of chaos in the data. Frazier and Kockelman (2004) proceeded in determining the presence of chaotic behavior in the sample data known as determinism test. To do this, various methods/techniques were used to determine...
deterministic chaotic behavior (Frazier and Kockelman, 2004). One of such methods used to test the presence of chaos was the Fourier power spectrum. Using Fourier power spectrum the graphical display of Figure 5 was created. The data is periodic if the power spectrum spikes at frequencies that characterize the system lay close to zero for all data points. If the data is chaotic, the spectrum will be broadband and have broad peaks as shown in Figure 7.

![Figure 7. Fourier power spectra for periodic (left) and chaotic (right) logistic functions](image)

Another method used to determine the presence of chaotic behavior was the largest Lyapunov Exponent. This method measured the divergence of nearby trajectories. As the system evolved, the sum of the series of attractor point values converges or diverges (Frazier and Kockelman, 2004). The Lyapunov Exponent measures the rate of convergence/divergence in each dimension and a chaotic system will exhibit trajectory divergence in at least one dimension. If the largest Lyapunov Exponent value exceeds zero, the system is chaotic. The equation used to determine the largest Lyapunov Exponent was:

\[
\lambda_{\text{max}} = \frac{1}{N\Delta t} \sum_{r=0}^{N-1} \ln \left( \frac{|s(t + \Delta t) - s'(t + \Delta t)|}{|s(t) - s'(t)|} \right)
\]
where $N$ was the original length of the data, $s(t)$ the original time series, $\Delta t$ the time delay and $s'(t)$ the derivative of the original data points are distinct but close. As $\Delta t$ increased, so did the Lyapunov Exponent and this theoretically converged to its true value. If largest Lyapunov Exponent exceeded zero, the data exhibited chaotic behavior. Only systems with Lyapunov Exponent value exceeding one are strong candidates for exhibiting properties typical of deterministic chaos. In the study conducted by Frazier and Kockelman (2004), in Figure 6 the graph showed that the Lyapunov Exponent value was between 6.75 and 8.5 indicating that the data studied showed properties of chaotic behavior.

In another popular field, nonlinear time series analysis was used to analyze electrical signal processing. For example, in a study conducted by Kodba, Perc, and Marhl (2005) the presence of chaotic behavior in a simple periodically driven resistor-inductor diode (RLC) was studied. Using nonlinear time series analysis methods, the delay coordinate embedding method was used for reconstructing an attractor (Kodba et al., 2005). In order to use the delay coordinate embedding method, a proper embedding delay and embedding dimension was determined using a mutual information method and a false nearest neighbor method respectively. Mutual information method was used in estimating the proper embedding delay whereas the false nearest neighbor method was used to determine the proper embedding dimension. With the reconstructing attractor, a simple determinism test was performed to distinguish between deterministic chaos behavior and random behavior which sometimes resembles chaos but rather comes from a stochastic system. Kadba et al., (2005) determined the presence of deterministic chaos in the driven RLC circuit by calculating the largest Lyapunov Exponent. In Kadba et al. (2005) study the result of the largest Lyapunov Exponent was 0.33 which indicated the presence of chaotic behavior in the experimental system.
There are plenty of literatures that can be found in this area. For example in a similar research involving electrical signals, researchers from University of Florida Brain Institute and the Malcom Randall Veterans Affairs Medical Center in Gainesville, Florida used chaotic theory concepts with electrical signal stimulation to predict some types of epileptic seizures minutes to hours before they occurred (University of Florida Health Science Center, 1999).

Nonlinear time series analysis has been applied in the field of physiology in recent years. For example a study conducted by Perc (2005a) analyzed the human electrocardiogram for chaotic behavior using simple nonlinear time series analysis methods. In the study, a short densely sampled electrocardiographic recording of the human heart was used. The sample was obtained from the publicly accessible MIT Polysomnographic database. Six hours of recordings was made but only a short insert of the data whereby the subject was normally asleep without any significant movement or apnoea attacks was used for the study. The time series studied consisted of 45,000 data points sampled at every 0.004 seconds. Therefore a total of 180 seconds of electrochemical heart activity was used for this study. The first 10 sec of the data is shown in Figure 8 below (Perc, 2005a).

![Figure 8. The studied human electrocardiogram data](image-url)
In this study, the phase space was reconstructed not until the data which possibly contaminated with noise, was extracted from the data using a simple noise reduction algorithm Perc (2005a). After extracting the ‘clean’ electrocardiographic signal from the data, the mutual information and the false nearest neighbor methods were used to obtain the optimal embedding delay time and the embedding dimension for the phase space reconstruction. Figure 9 below shows the 2D projection of the reconstruction phase space obtained with the embedding parameters (Perc, 2005a).

![Figure 9. Phase space with optimal embedding parameters (a) Before noise reduction, (b) After noise reduction](image)

With the ‘clean’ data time series, the data was tested for determinism and stationarity. Perc (2005a) applied a determinism test which determined whether the data exhibits deterministic properties, and a stationarity test to verify if the system parameters were held constant during the data recordings. With the deterministic and stationarity test determined, the Maximal Lyapunov Exponent was calculated. With a positive Maximal Lyapunov Exponent the studied electrocardiographic recording will indicate properties typical of deterministic chaotic signals (Perc, 2005a). The result from the study showed that the Maximal Lyapunov Exponent
was approximately 0.015 which indicate the data studied exhibited some properties typical of deterministic chaotic signals. Figure 10 illustrates the interpretation of the calculation of the Maximal Lyapunov Exponent. The x-axis $\Delta n$ represents the relative time and the y-axis $S(\Delta n)$ represents the expansion rate dependent of $\Delta n$. The Maximal Lyapunov Exponent is determined by the slope of the linear dashed line which provides a good estimate of the Maximal Lyapunov Exponent of 0.015 (Perc, 2005a).

Figure 10. Calculation of the Maximal Lyapunov Exponent

In another study conducted by Perc (2005b), an analysis of chaotic behavior of the dynamics of the human gait was analyzed using nonlinear time series method. In the study, a short recording of the human gait was obtained from the publicly accessible MIT Gait Database. The original sample data consisted of 90,000 data points that were sampled at intervals of 0.003secs from which only 10,000 was used for the study and calculations. This sample size was chosen according to Perc (2005b) because it was sufficient to get representative results in the shortest possible time. The total number of 30 seconds of human gait was used for the study of which the first 10 seconds is shown in Figure 11 (Perc, 2005b).
Figure 11. Recording of human gait

The time series collected was obtained by using two thin force-sensitive resistors connected parallel, placed inside the subject's right shoe under the heel and toes (Hausdorff Ladin, and Wei, 1995, Perc, 2005b). To complete one stride, the heel stroke the ground, the output voltage of the circuit increased rapidly reaching it maximum as a result of the sudden increase of weight acting upon the force-sensitive resistors. The weight was transferred from the heels to the toes resulting in a light voltage descent which is recorded. Just before the foot loses contact with the ground, the weight which is supported by solely the toes induces a light voltage increase. Finally, when the foot loses contact with the ground there was a sharp voltage descent. The time series from the study exhibited a rather regular activity with a predominant frequency approximately 1.0Hz according to Perc (2005b). Nonlinear time series analysis was applied to the data to understand the dynamics which yielded the observed behavior as shown above in Figure 11. The embedding time delay and the embedding dimension were carefully chosen to reconstruct a phase space. The mutual information method was applied to determine the embedding time delay $\tau$. Perc (2005b) explained that $\tau$ should not be larger than the typical time in which the system loses it memory of its initial state. That is, if $\tau$ was chosen bigger, the embedding space would look more or less random since it will consist of uncorrelated points.
Next, the embedding dimension \( m \) was determined by applying the false nearest neighbor method. The method proposed by Kennel, Brown, and Abarbanel (1992) is the best method of determining a minimal \( m \). This method is based on the assumption that the phase space is a deterministic system and that it folds and unfolds smoothly with no sudden irregularities in its structure (Perc, 2005b).

With the embedding delay time and dimension determined, the embedding phase space was reconstructed. A deterministic test was applied as proposed by Kaplan and Glass (1995) that measures the average directional vectors in a coarse-grained embedding space. This means that, neighboring trajectories in a small portion of the embedding space should all point in the same direction which will indicate the determinism of the system. For the data being studied it showed that the data was deterministic. With this result, it can now be determined if the data exhibit chaotic behavior. To determine chaotic behavior, Perc (2005b) applied the Maximal Lyapunov Exponent to test the divergence of the nearby trajectories in the phase space. If the Maximal Lyapunov Exponent was greater than zero, it indicated a strong chance that the data exhibited properties characteristic of chaotic behavior. In the human gait study, the Maximal Lyapunov Exponent was 0.21±0.02 which shows that the data of the human gait possessed properties typical of deterministic chaotic system.

Furthermore, chaos analysis has been applied widely in signal processing fields, but recent in physiological systems. An example of chaos analysis applied to physiological data is cardiovascular chaos studies. Glass and Winfree (1984) used chaos analysis on embryonic chick heart cells which were either beating spontaneously or were periodically simulated to determine existence of chaotic behavior (Bassingthwaighte et al., 1994a). In another example, Markus and Hess (1985) and Markus, Kuschmitz, and Hess (1984) used chaos analysis to study the metabolic
pathway that converts the free energy in glucose into adenosine triphosphate (ATP) which is used as an energy source for many biochemical reactions (Bassingthwaigte et al., 1994b). Finally, chaos analysis has been used in electroencephalogram (EEG) data which measures the activity of the brain by Mayer-Kress and Lynes (1987) whereby they analyzed whether the EEG data from a person resting quietly was purely from a random process or deterministic process (Bassingthwaigte et al., 1994b).

There are also applications in human behavior studies, for example in the field of speech perception, where chaos analysis concepts have been used to understand a phenomenon in speech perception known as categorical perception. According to Tuller (2005) this phenomenon occurs when within an acoustic parameter range it becomes difficult to separate between different stimuli that are labeled as the same speech segment (Tuller, 2005). In other words listeners have great difficulty in distinguishing two stimuli (e.g. the word ‘say’ and ‘stay’) which belong to same phonetic category (Tuller, 2005). This phenomenon is naturally nonlinear and therefore using nonlinear methods against traditional empirical methods will allow a deeper understanding of categorical perception. More examples of nonlinear methods used in the behavioral and cognitive science have been compiled in the book by Riley, and Van Orden (2005) entitled *Tutorials in contemporary nonlinear methods for the behavioral sciences.*
Over the past two decades, nonlinear methods have become popular in analyzing systems of different disciplines. Recently these methods have been applied to physiological systems to enhance the understanding of physiological behavior, such as human heart rate variations, and cardiovascular pressures. Most of the data from physiological systems appear random which if true allows us to use linear statistics method to analyze the data for relevant interpretation. However, if data from physiological systems are not random, then appropriate nonlinear methods must be used to examine and interpret the data and possibly characterize the system.

From the literature, examples were given where data visually appearing to be random have shown characteristics of chaotic behavior when nonlinear methods have been applied to the data. In analyzing such data, nonlinear methods was used to determine if the data was chaotic. The underlying feature of data exhibiting chaotic behavior is that the data must be deterministic that is, the data must exhibit periodic behavior. Finally, the Lyapunov Exponent was calculated which indicated a strong evidence of chaotic behavior if the Lyapunov Exponent is positive.

Limited literature was found on human behavior data where chaos theory and nonlinear analysis methods have been performed. Human behavior is driven by physiological systems which sometimes produce chaotic or random behavior. Knowing the data is random or chaotic the appropriate analysis tools will be applied to examine the data. In most psychological experimental data, traditional statistics methods have been often used to understand the data. But until recently, it was known that even though physiological data look random, it may just be a front for chaotic behavior. It is natural to believe that the data is random but further investigation sometimes will reveal that the data possesses chaotic behavior.
Objective of Study

Nonlinear time series methods have been applied in few disciplines to better understand the behavior of data but in the field of aviation and human performance data, the research is practically non-existing. It is important to investigate the nature of data prior to conducting any statistical analysis. This thesis will use a flight performance data collected in a previous study to determine if the data possesses deterministic chaotic behavior. Understanding the data through nonlinear time series will provide authentic results into how the data is represented, that is chaotic or random behavior.

Data for this thesis came from a recent research which studied the effectiveness of synthetic vision system (SVS)-Hits and velocity-vector based command augmentation system (V-CAS) (Liu, Goodrich and Peak, 2005). The study used a modified 1978 F33C Bonanza and eight non-IFR pilots and four IFR pilots to investigate the benefits and interaction terrain portrayals, guidance symbology and control-system response type on single pilot performance (SPP).

Looking at one of the data set for this thesis as shown in Figure 1, one cannot but suspect that the data set was as a result of either a completely random system or chaotic system. The purpose of this thesis is to apply nonlinear time series analysis methods on pilot's flight performance data to investigate deterministic chaotic behavior present in the time series data.
Method

In this section, a comprehensive framework of methods including various equations and algorithm used in nonlinear time series analysis methods are presented. The sequence of steps to take to analyze the data to determine if the data possess chaotic behavior are presented as well as a brief background on the flight performance data to be investigated.

Flight Performance Data

The data to be analyzed was obtained from NASA from a study which investigated the effect of synthetic vision system (SVS) concepts with or without velocity-vector based command augmentation system (V-CAS). The study evaluated the benefits and interactions of two level of terrain portrayal, guidance symbology and control-system response on SPP in the context of lower-landing minimums approaches (Liu, et al., 2005). Flight performance measurement consisted of flight technical error (FTE), pilot perceived workload and situational awareness (SA) and subjective preference. For this thesis only measurement of the flight technical error data will be analyzed. A test airplane modified 1978 Model F33C Bonanza; S/N CJ-144 was used by the subjects to record the flight performance data. Pilot flight performance was measured based on their flight plan using both fly by wire mechanism and the conventional flight control. In the cockpit were 2 high brightness 8’’ x 10’’ LCD display with resolutions of 1024 x 768 pixels. The left display was used to display the PFD concepts while the right display was used to present complimentary navigation providing real time platform view of the approach procedure and terrain referenced to an exocentric view of the aircraft. Two PC class computers run the display which received position and state information from air-data, attitude and heading reference, and also the computer running the PFD was used to record the flight time error and airspeed data. A final computer was interfaced with the flight control to record the control

Though there was data available for each 12 pilots, to keep our study balanced, two instruments rated pilots and two non-instruments rated pilots data were randomly selected to analyze the behavior of the flight performance data which include FTE (vertical deviation and horizontal deviation) and airspeed error. Data from the four pilots were randomly selected instead of all twelve pilots because the pilots had similar training experience that is, private pilot training and instrument rating respectively, hence in randomly selecting four pilots data the expectation would be that the result from the four would be similar to remaining eight pilot’s data.

In order to determine the chaotic behavior of the data, a sequence of steps has to be completed to come to such a conclusion. The following tests (algorithms) are used to establish the presence of chaos behavior in the flight performance data:

Figure 12. Sequence of steps to determine the presence of chaotic behavior in data
**Simple Noise Reduction**

Similar to physiological experimental measurements, flight performance data are bound to be contaminated by noise either from the measurement instruments used for collecting the data or the system itself (Bassingthwaighte et al., 1994b); therefore a ‘clean’ flight performance data is extracted from the noised data by implementing a simple noise reduction algorithm (Perc, 2005a). The noise reduction algorithm is implemented to try and decompose the time series into two components, with one containing the signal (clean flight performance data) and the other containing random fluctuations (Kantz and Schreiber, 1997). To implement the noise reduction algorithm, \( m, \tau, \varepsilon \) need to be calculated beforehand. The embedding dimension \( m \) has to be larger so that the chaotic behavior in the data can be greatly determined as well as improve the chances of selection an appropriate neighboring points with increasing \( m \). The delay time \( \tau \) is calculated such that \( m\Delta \tau \) ranges from 1/3 to 2/3 of the average time interval of the data. According to Perc (2005a), \( \varepsilon \), a small constant usually not larger than the standard deviation of the data; should be larger than the noise level in the data but still small enough not to average out the typical curvature radius of the time series. Once the \( m, \tau, \varepsilon \) have been determined, the noise reduction algorithm is applied to the data. The purpose of the algorithm is to substantially reduce the noise level while mostly preserving the curvature radius of the time series. The ‘clean’ flight performance data obtained after noise reduction is used to reconstruct the phase space. For further and precise algorithm description refer to page 51 of *Nonlinear Time Analysis* (1997) by Kuntz and Schreiber.
Estimating embedding time delay \( \tau \) by Mutual Information Method

When estimating \( \tau \), one has to consider that \( \tau \) has to be large enough so that the information measuring the value of \( x \) at time \( i + \tau \) is very different from the information already know at time \( i \). If this criterion is met then it can be guarantee that enough information will be gathered about all the other system variables that influence the value of \( x \) to reconstruct the phase space (Kadba et al., 2004). Based on the above requirement, Fraser and Swinger (1986) proposed the first minimal mutual information method between \( x_i \) and \( x_{i+r} \) as the optimal embedding delay \( \tau \). This quantifies the amount of information about the state of \( x_{i+r} \) if presumed to know the state of \( x_i \). The algorithm of calculating the mutual information begins with a given time series for instance equation (1). The minimum (\( x_{\text{min}} \)) and the maximum (\( x_{\text{max}} \)) of the time series is calculated and the absolute value of their difference

\[
|x_{\text{max}} - x_{\text{min}}|
\]

is partitioned into \( j \) equal sized intervals (boxes). Finally, the calculation of the mutual information is given by the expression:

\[
I(\tau) = \sum_{h=1}^{j} \sum_{k=1}^{j} P_{h,k}(\tau) \ln \frac{P_{h,k}(\tau)}{P_h P_k}.
\]

where:

\( P_h \) and \( P_k \) = the probabilities that the variable assumes a values inside the \( h^{th} \) and \( k^{th} \) box respectively

\( P_{h,k}(\tau) \) = joint probability that \( x_i \) is in box \( h \) and \( x_{i+r} \) is in box \( k \)

If \( j \) is large enough, the partitioning of the whole the data is fine and the value of the mutual information doesn’t explicitly depend on box size. The first minimum \( I(\tau) \) is the optimal choice for the embedding delay.
Estimating Embedding Dimension $m$ by False Nearest Neighbor

In establishing the proper embedding dimension $m$, the false nearest neighbor method by Kennel et al. (1992) is applied in order to determine the minimal $m$ that is required to resolve the deterministic structure of the system in the reconstructed phase space (Perc, 2005b). For this method to work, one has to assume the phase space of the deterministic systems folds and unfolds smoothly with no sudden irregularities appearing in the structure. This will ensure that those points that are close to the reconstructed embedding phase space stay sufficiently close during forward iteration. If this criterion is met, according to Kodba et al. (2005) the distance between two points $p(i)$ and $p(j)$ of the reconstructed phase space, which are initially only a small $\varepsilon$ apart cannot grow further as $R_\varepsilon$ and $\varepsilon$. $R_\varepsilon$ known as the threshold (a constant), according to Kennel et al. (1992) is equal to 10 has proven to be a good value for most data set.

In calculating $m$ the following algorithm is used. Given $p(i)$ in the phase space, the neighbor $p(j)$ is found so that

\[ || p(i) - p(j) || < \varepsilon \]

[6]

Where:

$|| ... ||$ = is the square of the norm

$\varepsilon$ = small constant not larger than standard deviation of the data

The normalized distance $R_i$ is calculated between the $(m+1)^{th}$ embedding coordinate of points $p(i)$ and $p(j)$ according to equation 7 (Kodba et al. 1992).
Applying the time series with various $m=1, 2, 3, \ldots, \infty$, $R_i$ is determined and if $R_i$ is larger than $R_{tr}$, the point $p(i)$ is known to have a false neighbor and recorded. What this means is for example, given the flight performance data, if $m=2$ then we can say that the flight performance time series has two active degree of freedom that is the experiment can be perform using no more than two first order differential equations.

Phase Space Reconstruction

With the embedding dimension ($m$) and time delay ($\tau$) properly determined and the data examined for any random fluctuations; the phase space is reconstructed from the ‘clean’ data. According to Takens (1981) reconstructing of the phase space has to be based off of the original time series

$$\{x_0, x_1, x_2, \ldots, x_i, \ldots, x_n\}$$

and using the proper embedding dimension, embedding time delay and the ‘clean’ time series, the reconstruction of the phase space is successfully created given by the equation

$$\{S_n = s(x_n)\}$$

where $S_n$ is the time series of the ‘clean’ data used to create an embedding phase space given by:

$$\{S_n = s_{n-1}, s_{n-2}, \ldots, s_{n}\}$$
It must be emphasized that the reconstructing of the phase space is crucial to the investigating of the flight performance data to be analyzed. This is because the success of reconstructing the phase space will provide the results that would prove meaningful as well as insightful into understanding the data. For example, using the embedding phase space, analysis of important properties of time series such as determinism and stationarity can be performed. The following sections discuss the algorithm used to test the embedding phase space for determinism and stationarity of the ‘clean’ flight performance data.

*Recurrence Plot Analysis*

To determine the stationarity and determinism of the ‘clean’ flight performance data, a graphical programming tool called a recurrence plot is applied according to Perc (2005b) to observe the stationarity and deterministic behavior of the data. According to Perc (2005b), since recurrent behavior is a characteristic of oscillation systems; if the data exhibits regular oscillations then the embedding phase space \( i, j \) can be arbitrarily close, that is

\[
\left\| p(i) - p(j) \right\| = 0
\]

so that \( i \) and \( j \) differ exactly by some integer of the oscillation period. If the flight performance data exhibit chaotic behavior, \( \left\| p(i) - p(j) \right\| \) is always finite. The recurrence plot algorithm plots a 2D square grid using \( \varepsilon, \tau, m \) and the embedding phase space points \( (i,j) \) that satisfies \( \left\| p(i) - p(j) \right\| < \varepsilon \). Interpretation of the recurrence plot is one of the most important features of the recurrence plot. The two features of the recurrence plot are the large and small scale structures of the recurrence plot also known as the typology and texture respectively (Perc, 2005b).
inspecting the typology and texture of the recurrence plot visually the stationarity and the
determinism of the system can be determined. According to Perc (2005b), if the recurrence plot
shows a homogenous typology the ‘clean’ flight performance data originated from a stationary
process. Otherwise a non-homogenous and disruptive typology will indicate the data originated
from a non-stationarity process characterized by white areas of the recurrence plot. The texture
provides information about the determinism verses stochastic origin of the signal characterized
by diagonal lines. Even though the recurrence plot algorithm answers the question of
determinism and stationarity of the data, the determinism and stationarity tests has to be
conducted to verify the results from the recurrence plot.

Determinism Test

Using the ‘clean’ flight performance time series data, the determinism test is applied as
proposed by Kaplan and Glass (1992). This test measures the average directional vectors in the
embedding phase space. According to Perc (2005a), he explained that determinism occurs when
neighboring trajectories in small portions of the embedding phase space are all pointing in the
same direction ensuring a unique solution in the space phase. For the algorithm to work, the
embedding space has to be coarse-grained into equally sized boxes. The average direction vector
$V_k$ is calculated at each pass $p$ of the trajectory through the $k^{th}$ box. This generates a unit vector
$e_p$ whose direction is determined by the phase space point where the trajectory enters the box and
the phase space point where the trajectory leaves the box. The average vector is given by:

$$ V_k = \frac{1}{N} \sum_{p=1}^{N} e_p $$  [11]
where \( n \) = number of passes through the \( k \)th box

When all passes have been completed for all occupied boxes in the embedding phase space; the resulting sum of each average directional vector yield a directional approximation for the vector field of the system. If the time series in our case the ‘clean’ flight performance data originate from a deterministic system and the box partitioning is done well, then the directional vector field should consist of only vectors that have a unit length of one. Therefore, if the system is deterministic the average of the lengths of directional vectors must equal one while for a random system the average of all the lengths of the directional vectors must approximate close to zero.

In another article by Perc (2006), a definite measure for determinism (\( \kappa \)) was proposed which weighted the entire average directional vector \( \mathbf{V}_k \) with respect to the average displacement per step \( R_i^* \) of a random walk. The determinism value \( \kappa \) is calculated by the equation (Perc, 2006)

\[
\kappa = \frac{1}{A} \sum_{i=1}^{A} \frac{(v_i)^2 - (R_i^*)^2}{1-(R_i^*)^2}
\]  

where

\( A \) = total number of occupied boxes

\( R_i^* \) is obtained by the given equation

\[
R_i^* = c_m P_i^{1/2}
\]

Where

\( c_m \) = constant depending on the embedding dimension.
As mentioned earlier if $\kappa=1$ it means the data exhibited deterministic behavior while $\kappa=0$ or less exhibited random behavior.

**Stationarity Test**

To confirm the results from the recurrence plot that the time series originated from a stationary process, a stationarity test by Schrieber (1993) is applied to the 'clean' flight performance data. The purpose of the stationarity test is to determine whether the mean values and variances of a series vary with time. Schrieber (1993) explained that since regular statistics such as the mean and standard deviation doesn't help in analyzing irregular signals, a nonlinear statistic known as cross-prediction error must be applied to the data for analysis. The stationarity test algorithm uses the cross-prediction error which splits the time series into several short non-overlapping segments and uses a particular data segment to make predictions in another data segment (Perc, 2005a). This process is repeated for all possible combinations meaning that for a total of four segments of the data there are $4^2$ possible combinations. The average prediction error ($\delta_{gh}$) is calculated when points in segment $g$ make a prediction in segment $h$; so given that a $\delta_{gh}$ against the total average $\delta_{gh}$ is significantly higher, the data can be thought as originating from a non-stationary process (Perc, 2005a).

**Maximal Lyapunov Exponent**

Maximal Lyapunov Exponent determines the rate of divergence or convergence of initially nearby trajectories in a phase space over a time period (Tabah, 1992). Generally for an $m$-dimensional phase space, there will be $m$ different Lyapunov Exponents which we denote by $\lambda_i$ where $i=1, 2, 3, \ldots, m$. These different Lyapunov Exponent when arranged in the order from
the largest to the smallest value form the Lyapunov spectrum \( \lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_m \) (Perc, 2005b). The system is chaotic when that one or more of the Lyapunov Exponent is larger than zero. The Maximal Lyapunov Exponent is calculated using the algorithm proposed by Wolf, Swift, Swinney, and Vastano (1985).

When calculating \( \lambda_i \), given an embedding phase space, the points \( p(t) \) and its neighboring point \( p(i) \) are located which satisfies:

\[
\| p(i) - p(t) \| \leq \epsilon
\]

The points are iterated and both points forward in time for a fixed evolution time \( \nu \) which should be between \( \tau \) and \( m\tau \). If the system is chaotic, over time \( \nu \) the distance given by the equation

\[
\| p(i + \nu) - p(t + \nu) \| = \epsilon \nu
\]

should be larger than the initial \( \epsilon \); for chaotic behavior and for regular behavior \( \epsilon \approx \epsilon \nu \) (Perc, 2006). After each time evolution \( \nu \) a replacement step is attempted in which a new point \( p(j) \) is found in the embedding space, whose distance to be evolved (point \( p(t + \nu) \)) should be small (\( \epsilon \)), under the constraint that angular separation between the vectors constituted by the points \( p(t + \nu) \), \( p(i + \nu) \), \( p(t + \nu) \) and \( p(j) \) are small. This process is repeated until the initial point of the trajectory reaches the last one. The Maximal Lyapunov Exponent is calculated by the equation

\[
\lambda_{\text{MAX}} = \frac{1}{M\nu} \sum_{i=1}^{M} \ln \frac{\epsilon \nu}{\epsilon}
\]
where
\[ M = \text{total number of replacement steps.} \]

If the Lyapunov Exponent converges which means \( \lambda_{\text{MAX}} > 0 \), the data indeed possesses chaotic behavior. It must be noted that for the algorithm to converge, the embedding space has to be constructed properly and the points have to be densely populated. But if the points are sparse the accuracy of the Maximal Lyapunov Exponent is arguable.

**Experimental Data**

Two randomly chosen data from non-instruments rated pilots and instrument rated pilots respectively (Liu, Goodrich and Peak, 2006) were selected and individually analyzed. Each pilot flew with the aircraft using conventional controls and then with the V-CAS. The pilots performed two flights each consisting of four different terrain portrayals and guidance and position awareness symbology concepts. Figures 13-16 shows the different terrains and guidance control used in the study.

![Figure 13. Blue Sky Brown Ground Terrain/NASA Tunnel (Baseline/NASA)](image-url)
Figure 14. Elevation Based Generic Terrain/NASA Tunnel (EBG/NASA)

Figure 15. Blue Sky Brown Ground/Pitch Roll Flight Director (Baseline/PRFD)
For each activity, three distinct data sets each consisting of at least 5,000 data points were recorded. These were the airspeed error, lateral path deviation and vertical path deviation. Each pilot flew two flights and hence eight individual data sets were recorded for each pilot, making a total of thirty-two data sets to analyze. In the next few pages the results of each of the thirty-two data sets are presented.

**Software**

The software used to analyze the data was provided by Matjaz Perc (2007). It is a collection of algorithms for the analysis of signals from nonlinear sources by using the algorithms above. The software is accessible to the user through a Graphic User Interface prompting the user to enter valid parameters. With the appropriate parameters, the appropriate executable file is run on the chosen algorithm.
Results

The next few pages presents the result found from running the sequence of steps discussed above to determine the presence of chaotic behavior in the flight performance data.

Pilot 1: Non-Instrument rating pilot using conventional controls

Data: Airspeed Error (from 90 knot/hr)   Terrain/Guidance Control: EBG/NASA

Figure 17. Airspeed Error Data of 7541 data points

Figure 17 shows the graph of the airspeed error data of a non-instrument rated pilot. By visibly inspection the graph, it cannot be determined if the data posseses chaotic or random behavior. The noise reduction algorithm was first applied to the original data to remove most of the noise in the data. Before performing the noise reduction, the embedding dimension and the time delay are found.
Finding the embedding time delay by Mutual Information Method (M.I.)

The optimal embedding delay is most likely 28. The Shannon entropy of data is 4.212.

Figure 18. The embedding time delay at $\tau=28$

In Figure 18, the optimal embedding time delay $\tau=28$ is determined by observing that the slope of the curve occurs at its most optimal at $\tau=28$. This value is the minimum embedding time delay of the airspeed error of 7541 data points. With the determination of the embedding time delay, together with the data series, the embedding dimension can be determined using the false nearest neighbor method.
Finding the embedding dimension by the False Nearest Neighbor (FNN) method

![Graph showing the False Nearest Neighbor method](image)

Figure 19. Finding the embedding dimension $m = 5$ with time delay $\tau = 28$ with 7541 data series

In Figure 19, with the time delay of $\tau = 28$ and 7541 data series, the embedding dimension is determined by observing the graph above that at $m = 5$ the slope of the graph is at its minimum. With the embedding time delay and dimension calculated, the noise reduction method is performed on the data to reduce the noise in the data.

**Noise Reduction of airspeed error data**

Figure 20a. Embedding space from original airspeed error data series.

Figure 20b. Embedding space after cleaning the original airspeed error data series.
Figure 20a shows the graph of the original airspeed error data. As noticed, the curvature of the graph isn’t smooth indicating that there is some noise in the data. After applying the noise reduction method to the data, Figure 20b shows a much refined curvature of the graph which shows a reduction of noise level in the data. This is known as the ‘clean’ data used for the analysis of determining the chaotic or random behavior of the data. Using the ‘clean’ data now, the embedding time delay and embedding dimension has to be determined again for use in the reconstructing of the phase space of the data.

Finding the new embedding time delay by Mutual Information Method (M.I.)

Figure 21. Graph showing the optimal embedding time delay at \( \tau = 25 \) using the ‘clean’ data
Finding the new embedding dimension (DIM) using False Nearest Neighbor Method (FNN)

Figure 22. Graph showing the minimal embedding dimension at $m=5$ using the 'clean data and $\tau=25$.

Plotting the embedded phase space using the $m=5 \tau=25$ and 'clean' data.

Figure 23. Plots of the embedding phase space at different coordinate systems.
In Figure 21, and 22 the new time delay and embedding dimension are determined. Together with the ‘clean’ data, the phase space of the airspeed error data points is reconstructed as shown in Figure 23. Figure 23 also shows the plots of different time delays found in Figure 21 and the different coordinate systems based on the embedding dimension \( m=5 \). Since the embedding dimension 5 corresponds to five different coordinate systems, a combination of the different time delays and coordinate systems are plotted to have a good representation of the phase space. The figure above shows the different phase space that allows us to clearly see the pattern of the data at different time delays.

In determining the presence of chaotic behavior in the airspeed error data, the recurrence plot analysis is performed which provides insight about the data originating from a non-stationary object as well as the data showing deterministic patterns.

*Plotting the recurrence plot*

![Recurrence plot](image)

Figure 24. The recurrence plot of 7493 points
Figure 24 shows the constructing of a 2D graph of dots from the data series. The coordinates represents the positions in time of the points within the observed time series that are neighbors in the reconstructed embedding space. The recurrence plot provides a powerful visualization of the time series, revealing the extent of non-stationarity and patterns at a glance. The graph shows a small square pattern created by the attractor in which the trajectory remains for a time in a region before moving another neighbor. As you move along the diagonal, there is no remarkable pattern which suggests that the data lacks deterministic characteristics and it can be observed that there are white spaces as you move along the diagonal suggesting that the data originated from a non-stationarity process.

*Deterministic Test*

The result of the deterministic test of the embedding phase space presented in Figure 23 is shown in Figure 25. The result showed a deterministic value of $\kappa=0.55$. This value reveals a weak determinism in the data which supports the analysis of the recurrence plot. Since the deterministic value of $\kappa=1$ is ideal for determining deterministic behavior we can confidently make the argument that the reconstructed phase space doesn't strongly exhibit periodic or deterministic behavior.

Figure 25. This shows the approximated vector field for the embedding space with a deterministic test value of 0.55
Stationarity Test

Figure 26. This shows the average cross prediction error of 11.34, with minimal and maximal cross prediction errors of 0.0054 and 21.12 respectively. The color map displays the average cross-predictions errors independence of different segment combinations.

In Figure 26, as described in the method’s section, if the average cross prediction is significantly above average for any given combination of $i$ and $j$ then the system could be thought as originating from an non-stationarity process. In our case, since the maximal cross prediction error is almost two times larger than the average cross prediction error, we can clearly state and contest that the data originated from a non-stationarity process.

Maximal Lyapunov Exponent

With the information we now know about the stationarity and determinism of the data, we can predict that the data might not possess chaotic behavior since the deterministic test result revealed the data loosely exhibit deterministic behavior. But to confirm this conclusion, the Maximal Lyapunov Exponent must be found to be certain that the data indeed doesn’t possess
chaotic behavior. As seen in Figure 27 below, the slope of the linear part of the graph indicates a good estimate of the Maximal Lyapunov Exponent of the airspeed error data which in this case is a negative value of $\lambda = -0.0633$. The interpretation of this negative value tells that the data doesn’t possess deterministic behavior which confirms what was already known about outcome of the nature of the airspeed error data.

Figure 27: Maximal Lyapunov Exponent of airspeed error -0.0633
In the next few paragraphs, the results of the thirty-two data series from the four pilots are summarized.

Table 1
Summary of determining chaotic behavior of non-instrument pilot data using conventional controls

<table>
<thead>
<tr>
<th>CONVENTIONAL CONTROL</th>
<th>BASELINE EBG/NASA</th>
<th>BASELINE EBG/PRFD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non Instrument Pilot</td>
<td>Pilot 1</td>
<td>Pilot 2</td>
</tr>
<tr>
<td><strong>Airspeed Error</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Original Data Series</td>
<td>7541</td>
<td>7836</td>
</tr>
<tr>
<td>‘Clean’ Data Series</td>
<td>7493</td>
<td>7812</td>
</tr>
<tr>
<td>Deterministic Value</td>
<td>0.55</td>
<td>0.588</td>
</tr>
<tr>
<td>Stationarity Value</td>
<td>11.34</td>
<td>1.85</td>
</tr>
<tr>
<td>Maximal Lyapunov</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exponent</td>
<td>-0.06</td>
<td>-0.215</td>
</tr>
<tr>
<td><strong>Lateral Deviation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Original Data Series</td>
<td>7541</td>
<td>7836</td>
</tr>
<tr>
<td>‘Clean’ Data Series</td>
<td>7173</td>
<td>7698</td>
</tr>
<tr>
<td>Deterministic Value</td>
<td>0.88</td>
<td>0.84</td>
</tr>
</tbody>
</table>
### CONVENTIONAL CONTROL

<table>
<thead>
<tr>
<th></th>
<th>EBG/NASA</th>
<th>EBG/PRFD</th>
<th>BASELINE /NASA</th>
<th>BASELINE /PRFD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non Instrument Pilot</td>
<td>Pilot 1</td>
<td>Pilot 2</td>
<td>Pilot 1 Pilot 2</td>
<td>Pilot 1 Pilot 2</td>
</tr>
<tr>
<td>Stationarity Value</td>
<td>18.79</td>
<td>22.88</td>
<td>54.35</td>
<td>26.12</td>
</tr>
<tr>
<td></td>
<td>30.61</td>
<td>18.45</td>
<td>39.38</td>
<td>48.41</td>
</tr>
<tr>
<td>Maximal Lyapunov Exponent</td>
<td>-0.02</td>
<td>-0.035</td>
<td>-0.08</td>
<td>-0.45</td>
</tr>
<tr>
<td>Vertical Deviation</td>
<td>54.35</td>
<td>26.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deterministic Value</td>
<td>0.549</td>
<td>0.78</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>Stationarity Value</td>
<td>24.81</td>
<td>9.2</td>
<td>37.62</td>
<td>11.52</td>
</tr>
<tr>
<td></td>
<td>26.53</td>
<td>24.67</td>
<td>211.59</td>
<td></td>
</tr>
<tr>
<td>Maximal Lyapunov Exponent</td>
<td>0.016</td>
<td>-0.04</td>
<td>-0.02</td>
<td>-0.03</td>
</tr>
</tbody>
</table>

In Table 1, the results of determining the existence of chaotic behavior in the airspeed error, lateral deviation and vertical deviation data of non-instrument rated pilots using conventional controls are presented. For each terrain portrayal for instance EBG/NASA, Pilot 1 and Pilot 2 results are compared. Using the airspeed error original data for example, the data was cleaned and the embedding phase space reconstructed using the ‘clean’ data. The recurrence plot analysis was performed and confirmed by the stationarity and deterministic test. In the case of EBG/NASA portrayal using the airspeed error data, it was found that Pilot 1 had a deterministic and stationarity value of 0.55 and 11.34 respectively while Pilot 2 had a deterministic and
stationarity value of 0.6 and 1.85 respectively. This means that firstly, the airspeed data from Pilot 1 and Pilot 2 originated from a non-stationarity process. And secondly, the airspeed error data of Pilot 1 showed the data was not deterministic while Pilot 2 was deterministic because the Maximal Lyapunov Exponent value was negative which indicate that Pilot 1 airspeed error data didn’t exhibit chaotic behavior ($\lambda = -0.06$) while Pilot 2’s data showed a loosely weak deterministic behavior ($\lambda = 0.12$).

Table 2
Summary of determining chaotic behavior of non-instrument pilot data using VCAS controls

<table>
<thead>
<tr>
<th>VCAS CONTROL</th>
<th>BASELINE</th>
<th>BASELINE</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EBG/NASA</td>
<td>EBG/PRFD</td>
<td>/NASA</td>
<td>/PRFD</td>
<td>Pilot 1</td>
<td>Pilot 2</td>
<td>Pilot 1</td>
<td>Pilot 2</td>
</tr>
<tr>
<td>Non Instrument Pilot</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pilot 1</td>
<td>Pilot 2</td>
<td>Pilot 1</td>
<td>Pilot 2</td>
</tr>
<tr>
<td>Airspeed Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pilot 1</td>
<td>Pilot 2</td>
<td>Pilot 1</td>
<td>Pilot 2</td>
</tr>
<tr>
<td>Original Data Series</td>
<td>7541</td>
<td>7023</td>
<td>7311</td>
<td>7866</td>
<td>6365</td>
<td>5300</td>
<td>6306</td>
<td>6065</td>
</tr>
<tr>
<td>‘Clean’ Data Series</td>
<td>7517</td>
<td>6951</td>
<td>7224</td>
<td>7818</td>
<td>6344</td>
<td>5279</td>
<td>6285</td>
<td>6032</td>
</tr>
<tr>
<td>Deterministic Value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stationarity Value</td>
<td>3.1</td>
<td>0.77</td>
<td>3.16</td>
<td>1.48</td>
<td>0.77</td>
<td>2.9</td>
<td>0.51</td>
<td>1.35</td>
</tr>
<tr>
<td>Maximal Lyapunov Exponent</td>
<td>-0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Lateral Deviation

| Original Data Series  | 7541     | 7023     | 7311 | 7866 | 6365    | 5300    | 6306    | 6065    |
In Table 2, the results of determining the existence of chaotic behavior in the airspeed error, lateral deviation and vertical deviation data of non-instrument rated pilots using VCAS controls are presented. For each terrain portrayal for instance EBG/PRFD, Pilot 1 and Pilot 2 results are compared. Using the lateral deviation original data for example, the data was cleaned and the embedding phase space reconstructed using the clean data. The recurrence plot analysis was performed and confirmed by the stationarity and deterministic test. In the case of EBG/PRFD portrayal using the lateral deviation data, it found that Pilot 1 had a deterministic and
stationarity value of 0.85 and 18.04 respectively while Pilot 2 had a deterministic and stationarity value of 0.86 and 33.49 respectively. This information tells us that the data of Pilot 1 and Pilot 2 originated from a non-stationarity process as well as it being deterministic because their Maximal Lyapunov Exponent value of Pilot 1 exhibited some deterministic chaotic behavior \( (\lambda = 0.07) \) as well as Pilot 2 \( (\lambda = 0.05) \). Similar steps described above were performed on each of the terrain portrayals and their corresponding data to arrive at their Maximal Lyapunov Exponent which determined the existence of chaotic behavior in the data examined.

Table 3
Summary of determining chaotic behavior of instrument-rated pilot data using conventional controls

<table>
<thead>
<tr>
<th>CONVENTIONAL CONTROL</th>
<th>BASELINE EBG/NASA</th>
<th>BASELINE EBG/PRFD</th>
<th>BASELINE NASA</th>
<th>BASELINE PRFD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instrument Pilot</td>
<td>Pilot 3</td>
<td>Pilot 4</td>
<td>Pilot 3</td>
<td>Pilot 4</td>
</tr>
<tr>
<td>Airspeed Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Original Data Series</td>
<td>7788</td>
<td>7321</td>
<td>7804</td>
<td>8505</td>
</tr>
<tr>
<td></td>
<td>6834</td>
<td>6278</td>
<td>6575</td>
<td>6143</td>
</tr>
<tr>
<td>'Clean' Data Series</td>
<td>7710</td>
<td>7289</td>
<td>7777</td>
<td>8478</td>
</tr>
<tr>
<td></td>
<td>6764</td>
<td>6260</td>
<td>6533</td>
<td>6115</td>
</tr>
<tr>
<td>Deterministic Value</td>
<td>1.42</td>
<td>2.86</td>
<td>2.07</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>0.76</td>
<td>0.12</td>
<td>0.84</td>
<td>1.67</td>
</tr>
<tr>
<td>Stationarity Value</td>
<td>-0.03</td>
<td>-0.12</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>Maximal Lyapunov Exponent</td>
<td>-0.03</td>
<td>-0.12</td>
<td>0.03</td>
<td>0.03</td>
</tr>
</tbody>
</table>

51
<table>
<thead>
<tr>
<th>Instrument Pilot</th>
<th>Pilot 3</th>
<th>Pilot 4</th>
<th>Pilot 3</th>
<th>Pilot 4</th>
<th>Pilot 3</th>
<th>Pilot 4</th>
<th>Pilot 3</th>
<th>Pilot 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lateral Deviation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Original Data Series</td>
<td>7788</td>
<td>7321</td>
<td>7804</td>
<td>8505</td>
<td>6834</td>
<td>6278</td>
<td>6575</td>
<td>6143</td>
</tr>
<tr>
<td>‘Clean’ Data Series</td>
<td>7668</td>
<td>7204</td>
<td>7687</td>
<td>8327</td>
<td>6654</td>
<td>6161</td>
<td>6507</td>
<td>6071</td>
</tr>
<tr>
<td>Deterministic Value</td>
<td>0.89</td>
<td>0.84</td>
<td>0.81</td>
<td>0.85</td>
<td>0.88</td>
<td>0.09</td>
<td>0.11</td>
<td>0.1</td>
</tr>
<tr>
<td>Stationarity Value</td>
<td>7.89</td>
<td>14.11</td>
<td>26.26</td>
<td>33.19</td>
<td>3.7</td>
<td>17.46</td>
<td>23.37</td>
<td>14.18</td>
</tr>
<tr>
<td>Maximal Lyapunov Exponent</td>
<td>-0.09</td>
<td>-0.11</td>
<td>-0.1</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Vertical Deviation</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Data Series</td>
<td>7788</td>
<td>7321</td>
<td>7804</td>
<td>8505</td>
<td>6834</td>
<td>6278</td>
<td>6575</td>
<td>6143</td>
</tr>
<tr>
<td>‘Clean’ Data Series</td>
<td>7698</td>
<td>7249</td>
<td>7690</td>
<td>8421</td>
<td>6735</td>
<td>6185</td>
<td>6467</td>
<td>6026</td>
</tr>
<tr>
<td>Deterministic Value</td>
<td>0.78</td>
<td>0.75</td>
<td>0.76</td>
<td>0.82</td>
<td>0.75</td>
<td>0.78</td>
<td>0.78</td>
<td>0.87</td>
</tr>
<tr>
<td>Stationarity Value</td>
<td>6.41</td>
<td>10.68</td>
<td>17.73</td>
<td>9.21</td>
<td>13.95</td>
<td>12.35</td>
<td>14.96</td>
<td>48.24</td>
</tr>
<tr>
<td>Maximal Lyapunov Exponent</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.026</td>
<td>-0.1</td>
</tr>
</tbody>
</table>

In Table 3, the results of determining existence of chaotic behavior in airspeed error, lateral deviation and vertical deviation data of instrument rated pilots using conventional controls is presented. For each terrain portrayal for instance BASELINE/NASA, Pilot 3 and Pilot 4 are compared. Using the vertical deviation original data for illustration, the data was cleaned and the embedding phase space reconstructed using the clean data. The recurrence plot analysis was
performed and confirmed by the stationarity and deterministic test. In the case of BASELINE/NASA portrayal using airspeed error data, it found that Pilot 3 had a deterministic and stationarity value of 0.75 and 13.95 respectively while Pilot 4 had a deterministic and stationarity value of 0.77 and 12.35 respectively. This means that the data Pilot 3 and Pilot 4 originated from a non-stationarity process as well as the data being deterministic because the Maximal Lyapunov Exponent value of Pilot 3 indicates there is some chaotic behavior (\( \lambda = 0.02 \)) as well as Pilot 4 (\( \lambda = 0.09 \)) in the data examined. Similar steps described above were performed on each of the terrain portrayals and their corresponding data to arrive at their Maximal Lyapunov Exponent to determine whether the data exhibited chaotic behavior.

Table 4
Summary of determining chaotic behavior of instrument-rated pilot data using VCAS control

<table>
<thead>
<tr>
<th>VCAS CONTROL</th>
<th>BASELINE/</th>
<th>BASELINE/</th>
<th>BASELINE/</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EBG/NASA</td>
<td>EBG/PRFD</td>
<td>NASA</td>
</tr>
<tr>
<td>Instrument Pilot</td>
<td>Pilot 3</td>
<td>Pilot 4</td>
<td>Pilot 3</td>
</tr>
<tr>
<td>Airspeed Error</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Original Data Series</td>
<td>10470</td>
<td>7439</td>
<td>7385</td>
</tr>
<tr>
<td>'Clean' Data Series</td>
<td>10332</td>
<td>7395</td>
<td>7358</td>
</tr>
<tr>
<td>Deterministic Value</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stationarity Value</td>
<td>1.6</td>
<td>2.48</td>
<td>1.42</td>
</tr>
<tr>
<td>Maximal Lyapunov Exponent</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In Table 4 above, the results of determining existence of chaotic behavior in airspeed error, lateral deviation and vertical deviation data of instrument rated pilots using VCAS controls are presented. For each terrain portrayal for instance BASELINE/PRFD, Pilot 3 and Pilot 4 are compared. Using the airspeed error original data for illustration, the data is cleaned and the embedding phase space reconstructed using the clean data. The recurrence plot analysis is
performed which is confirmed by performing the stationarity and deterministic test. In the case of the BASELINE/PRFD portrayal using airspeed error data, it was found that Pilot 3 had a deterministic and stationarity value of 0.57 and 0.05 respectively while Pilot 4 had a deterministic and stationarity value of 0.47 and 2.08 respectively. This means that the data of Pilot 3 and Pilot 4 originated from a non-stationarity process as well as it being deterministic. If the data is deterministic, then it implies the data should possess chaotic behavior. But the Maximal Lyapunov Exponent value $\lambda = -0.12$ determined of Pilot 3 indicate a negative value which signifies the data doesn’t possess chaotic behavior while the Lyapunov Exponent value of $\lambda = 0.12$ Pilot 4 indicate the data possesses chaotic behavior. Similar steps as described above were performed on each of the terrain portrayals and their corresponding data to arrive at their Maximal Lyapunov Exponent to determine whether the data exhibited chaotic behavior.
Summary of Findings

In summarizing the results found, it is evident in Tables 1, 2, 3 and 4 that not all the data examined exhibited chaotic behaviors. Only a few of the data showed chaotic behavior as highlighted in Tables 1, 2, 3, and 4. Also the results revealed that all the data originated from a non-stationary process. Determinism which is a characteristic of chaotic behavior was not strongly present in all the data examined. The deterministic test conducted showed that at least all the data did possess some traces of determinism as evident the data’s low deterministic test value. The Maximal Lyapunov Exponent value which indicates the nature of the behavior in a data showed a mixture of low negative and positive Lyapunov Exponent values. In the next section, explanation to the presence of chaotic behavior in some of the data will be further discussed to understand the type of data being analyzed and also attempt to rationalize the discrepancies of the deterministic test and the Lyapunov Exponent values which are important into establishing chaotic behavior in the data examined.
Discussion

Presence of Chaotic Behavior in data

The objective of this thesis was to investigate the presence of chaotic behavior in human performance data. Nonlinear time series methods until recently have been applied in few disciplines to better understand the behavior of data but in the field of aviation and human performance data, the research is practically non-existing. This thesis used flight performance data collected in a previous study to determine deterministic chaotic behavior in the data.

Thirty-two data time series was analyzed to obtain insight into the behavior of human flight performance data. The noise reduction method was applied to the original data to reduce the noise in the data. Using non-linear analysis methods and the ‘clean’ data, the optimal embedding time delay and dimension was determined to reconstruct the embedding phase space. The ‘clean’ data was used to perform a stationarity and deterministic tests as well as calculate the Maximal Lyapunov’s Exponent. The outcome of the thirty-two data series showed some of the data possessed deterministic chaotic properties. The results also revealed that not all human flight performance data are necessarily chaotic. This ambiguity in the results needs to be looked into to ascertain what could have contributed to some of the data not showing deterministic chaotic behavior. The next few paragraphs will attempt to explain and understand the reasons for the inconsistency in results.

Overall, comparing the results of the non-instrument rated pilots (pilot 1 and pilot 2) using conventional controls in Table 1, it can be found that some of the data (airspeed error, vertical and horizontal deviation data) possessed some level of deterministic chaotic behavior. Chaotic behavior between pilot 1 and pilot 2 of the airspeed error data was evident only in Baseline/NASA terrain portrayal using conventional control with Maximal Lyapunov Exponent.
\[ \lambda = 0.06 \] and \[ \lambda = 0.03 \] of pilot 1 and pilot 2, respectively. This Maximal Lyapunov Exponent values observed in Table 1 is however low suggesting there is some deterministic chaotic behavior in the data indicative of the loosely weak deterministic seen in Table 1. The positive maximum Lyapunov Exponent value which is defining feature of chaotic behavior is quite low for pilot 1 and 2. In this scenario, the Maximal Lyapunov Exponent value is approximately zero but since the Lyapunov Exponent value obtained is slightly greater than zero, it can be speculated that there is some level of deterministic chaotic behavior in the data. Data not exhibiting chaotic behavior was evident in both pilot's vertical deviation data using EBG/NASA and BASELINE/NASA terrain portrayals. The Maximal Lyapunov Exponent for pilot 1 was \[ \lambda = -0.016 \] and \[ \lambda = -0.05 \] while pilot 2 was \[ \lambda = -0.04 \] and \[ \lambda = -0.03 \] clearly indicating the non-existing chaotic behavior in the data.

Furthermore, results of the non-instrument rated pilots (pilot 1 and pilot 2) in Table 2 using VCAS controls; again showed mixed results of some data possessing deterministic chaotic behavior. However, result of pilots using the VCAS showed more chaotic behavior between the airspeed error, lateral and vertical deviation data and their corresponding terrain portrayals than the pilot 1 and pilot 2 using conventional controls. For example in Table 2 the presence of chaotic behavior is strongly evident in the lateral deviation data where all four terrain portrayals indicated a positive Lyapunov Exponent between pilot 1 and pilot 2. More presence of chaotic behavior is seen in the airspeed error data using EBG/PRFD (\[ \lambda = 0.03, \lambda = 0.03 \]) and BASELINE/PRFD (\[ \lambda = 0.11, \lambda = 0.02 \]) between pilot 1 and 2 respectively. Also there was chaotic behavior exhibited in the vertical deviation data where the Maximal Lyapunov Exponent value were \[ \lambda = 0.06, \lambda = 0.02 \] for pilot 1 and pilot 2 respectively using the BASELINE/NASA terrain portrayal. Again, as observed earlier, the Maximal Lyapunov Exponent value is quite low.
leading to suggest that there is some level of deterministic chaotic behavior in the some of the data. Data not exhibiting chaotic behavior was evident only in both pilot’s vertical deviation data using EBG/PRFD whereby the Maximal Lyapunov Exponent was $\lambda = -0.01$, $\lambda = -0.06$ for pilot 1 and pilot 2 respectively.

The result of the instrument rated pilots (pilot 3 and pilot 4) in Table 3 using conventional controls showed few chaotic behaviors in the data analyzed contrary to the expectation that the presence of deterministic chaotic behavior would be more prevalent compared to the non-instrument pilots. Chaotic behavior was apparent in the airspeed error data using EBG/NASA ($\lambda = 0.02$, $\lambda = 0.04$) for pilot 3 and 4 respectively, lateral deviation data using BASELINE/NASA ($\lambda = 0.05$, $\lambda = 0.04$) and BASELINE/PRFD ($\lambda = 0.06$, $\lambda = 0.01$) for pilot 3 and 4 in that order and finally, vertical deviation data using EBG/PRFD ($\lambda = 0.04$, $\lambda = 0.04$) and BASELINE/NASA ($\lambda = 0.02$, $\lambda = 0.09$) for pilot 3 and 4 correspondingly. Data not exhibiting chaotic behavior was present in the lateral deviation data using EBG/PRFD ($\lambda = -0.11$, $\lambda = -0.1$), vertical deviation data using EBG/NASA ($\lambda = -0.03$, $\lambda = -0.03$) and BASELINE/PRFD ($\lambda = -0.026$, $\lambda = -0.01$) of pilot 3 and 4 respectively. Once again it must be noted that the Maximal Lyapunov Exponent value is quite low but enough to suggest that there is some deterministic chaotic behavior in some of the data.

Finally, in examining the results of the instrument rated pilots using VCAS control (pilot 3 and 4) in Table 4, deterministic chaotic behavior was present in the airspeed error data ($\lambda = 0.05$, $\lambda = 0.21$) and vertical deviation data ($\lambda = 0.02$, $\lambda = 0.01$) using EBG/NASA terrain of pilot 3 and 4 respectively. The positive Maximal Lyapunov Exponent indicating chaotic behavior was found in the vertical deviation data using EBG/PRFD ($\lambda = 0.03$, $\lambda = 0.04$) and BASELINE/PRFD ($\lambda = 0.03$, $\lambda = 0.01$) of pilot 3 and 4 accordingly. Since the Lyapunov
Exponent value observed here is slightly greater than zero, it can be considered that there is some level of deterministic chaotic behavior in the data. Data not exhibiting chaotic behavior was only present in the airspeed error data using BASELINE/NASA terrain with Maximal Lyapunov Exponent of $\lambda = 0.03$ and $\lambda = 0.01$ of pilot 3 and pilot 4 correspondingly.

Inconsistency in the Results

In Tables 1, 2, 3, and 4, it was found that some of the data showed weak and strong determinism, a characteristic of chaotic behavior. It also found that all data exhibited properties of non-stationarity. The data examined originated from a non-stationary process with some of the data possessing some level of determinism. But as we have become aware of some of the data examined exhibited deterministic chaotic behavior which were expected however, this was not consistent between the non-instrument and instrument rated pilots' data and the pilot's data that have similar training level. As a result of the enormous disparities of chaotic and random behavior in the data among the four terrain portrayals, this investigation cannot conclusively provide the solid proof to suggest that flight performance data possesses chaotic behavior but rather, this investigation can make known that some flight performance data do exhibit some deterministic chaotic behavior. The questions that remains unanswered are what might have accounted for these disparities in the result hence the inability to confidently suggest human performance data exhibits deterministic chaotic behavior for the data investigated and also what could have factored into to the low Maximal Lyapunov Exponent value and as well as the weak deterministic value obtained for the data that showed signs of deterministic chaotic behavior.

In answering the question presented above, the disparities in the results may be due to prior pilot experience and training. Comparing the occurrence of chaotic behavior present
between pilot 1 and pilot 2 (instrument rated) and pilot 3 and pilot 4 (non instrument rated) using conventional controls, it revealed that deterministic chaotic behavior was only present in both pilot 1 and pilot 2’s airspeed error data using the BASELINE/NASA terrain portrayal (see Table 1). But chaotic behavior was however present in most of both pilot 3 and pilot 4’s data (i.e. airspeed error, vertical and lateral data) as indicated by their positive Maximal Lyapunov Exponent values highlighted in Table 3. This revelation from the results appears to support the notion that instrument rated pilots, by virtue of their level of pilot experience or training should have more determinism in their ability to control the airplane. However, in comparing the occurrence of chaotic behavior present in data of the pilots (instrument and non instrument rated) using VCAS controls, it was found there was a higher count of chaotic behavior exhibited by non-instrument rated pilot than their counterparts as seen in Table 2 and 4. This rather surprising result therefore renders the notion of instrument rated pilots having more determinism in their ability to control the airplane erroneous. By this analysis therefore, it can be argued that the inconsistency of the existence of deterministic chaotic behavior in the data are not reflective of the pilot’s experience or training because as observed, regardless of level of pilot experience or training, some of the data from both instrument and non-instrument pilots showed deterministic chaotic behavior signifying that pilot experience couldn’t have been a contributing factor to the disparities of the presence of chaotic behavior in the data examined.

Another influencing factor that may have resulted in the low determinism and low Maximal Lyapunov Exponent values is the mechanics of the airplane. The airplane was controlled by pilots to collect the data and the collected data consist of two components-human and machine characteristics. That is, for example for a pilot to roll out of a turn, the human characteristic is the ability for the pilot to turn the yoke to initiate the turn and the machine
characteristic is the mechanism for the plane to accurately perform the turn. The data composed of this mixture of human and machine characteristics makes the data analyzed in this study not purely a human performance data. The machine component may have introduced some degree of randomness in the data and this may have resulted in some of the data recording low determinism and Maximal Lyapunov Exponent values. Methods such as Detrended Fluctuation Analysis (DFA) have been used in isolating different component of signals. This method could be applied to the data to remove the machine component of the data leaving a purely human performance data which if used would probably result in a higher determinism and a higher Maximal Lyapunov Exponent.
Conclusions

The goal of this study was to understand the nature of flight human performance data. In order to accomplish this goal, non-linear methods were used to understand if chaotic behavior existed in the flight performance data.

A sequence of methods was applied to determine the presence of chaotic behavior in the data. The embedding time delay and embedding dimension is determined using the false nearest neighbor and mutual information method respectively. This information is used to reconstruct the embedding phase space. The phase space was used for further analysis to determine if the data was from a deterministic system by performing a deterministic test. A stationarity test was performed to determine if the data collected was from a stationary or non-stationary system. The Maximal Lyapunov Exponent was determined for the existence of chaotic behavior in the data.

The investigation conducted found that some flight performance data possessed chaotic behavior. Chaotic behavior was seen in some instrument rated and non-instrument rated pilots’ flight performance data; however there was no conclusive pattern to confidently suggest flight performance data exhibits chaotic behavior. Influencing this kind of result includes the pilot’s prior experience and training, and the operations of the aircraft control and mechanics of the airplane.

In most nonlinear methods, analyzing time series data requires the sample size to contain well over 10,000 data points and more. In this study, most of the data points examined was between 8,000-11,000 data points. Therefore, in future studies the sample size must contain more than 10,000 points to meaningfully use nonlinear methods to accurately make analysis of the data. Another limitation encountered in this study was the inability to know the ages of the pilots as well as the number of years of flight experience and amount of training received; which as
noted earlier may be a factor that might have influenced the inconsistencies in the results. Knowing this information would be valuable in understanding the data better in determining whether instrument rated pilots or non-instrument rated pilot data exhibit chaotic behavior and by of magnitude. If this limitation is resolved in future studies, it will allow better understanding whether flight performance data truly processes chaotic behavior.

Another limitation that needs to be pointed out is the nature of the data (machine and human characteristics). This has to be distinguished in order to analyze the human characteristics for the presence of deterministic chaotic behavior in flight performance data. A method such as Detrended Fluctuation Analysis (DFA) could be used to remove the pure human characteristic component of the data for analysis. Such data will provide the confidence needed to conclusively answer the question of the presence of deterministic chaotic behavior in the data. Also, knowing the data has deterministic chaotic behavior; a method such as fractal analysis which assumes that the data is already deterministic could be used in facilitating understanding the nature of data and enable the prediction of future nature of similar human performance data.

In conclusion, this thesis revealed flight performance data may exhibit chaotic behavior but we cannot conclusively confirm chaotic behavior exist in flight performance data due so the limitations discussed above. Future research must emphasis on analyzing the data in order to extract the human characteristic component of the data for analysis of the presence of deterministic chaotic behavior in human flight performance data. Humans are a complex system, and in order to understand this system better, applying the appropriate analysis is vital for understanding human performance from the data collected. Models such as non-linear analysis will be very useful for behavior scientists to obtain insights into the performance data which traditional technique cannot provide.
References


Management, 28, 61-73.


