Review Article

Application of Chaos Theory in the Prediction of Motorised Traffic Flows on Urban Networks

Aderemi Adewumi,[1] Jimmy Kagamba,[2] and Alex Alochukwu[1]

[1]*School of Mathematics, Statistics & Computer Science, University of KwaZulu-Natal, Westville Campus, Durban 4000, South Africa*
[2]*African Institute for Mathematical Sciences, Km 2 Route de Joal, Centre for IRD, BP 1418, M'Bour, Senegal*

Correspondence should be addressed to Aderemi Adewumi; laremtj@gmail.com

In recent times, urban road networks are faced with severe congestion problems as a result of the accelerating demand for mobility. One of the ways to mitigate the congestion problems on urban traffic road network is by predicting the traffic flow pattern. Accurate prediction of the dynamics of a highly complex system such as traffic flow requires a robust methodology. An approach for predicting Motorised Traffic Flow on Urban Road Networks based on Chaos Theory is presented in this paper. Nonlinear time series modeling techniques were used for the analysis of the traffic flow prediction with emphasis on the technique of computation of the Largest Lyapunov Exponent to aid in the prediction of traffic flow. The study concludes that algorithms based on the computation of the Lyapunov time seem promising as regards facilitating the control of congestion because of the technique's effectiveness in predicting the dynamics of complex systems especially traffic flow.

1. Introduction

In recent times, urban traffic road networks are faced with severe congestion problems as a result of the accelerating demand for mobility. The excessive congestion in the form of immense traffic jams on urban roads has hindered mobility along these roads. This is one of the major challenges encountered in most mega cities around the world with urban road networks and in turn has a serious effect on road users which includes economic, health, and environmental problem such as vehicle emission and air pollution, arising out of increased fuel consumption during the long periods of congestion. U.S. Bureau of Transport Statistics in 2007 recorded that, due to traffic congestion, Americans residing in urban areas were coerced to travel more 4.2 billion hours and spent about \$87.2 billion in purchasing extra 2.8 billion gallons of fuel [1, 2].

Urban planning and complex traffic network studies have been explored explicitly to potentially mitigate congestion and its associated problems on urban roads. Several efforts and studies have been made in time past by researchers on two major areas that affect urban traffic, namely, traffic flow modeling and prediction and information communications technology which is meant to give guidance to drivers through updated information about their desired routes [3]. However, without fundamental knowledge of the dynamics of vehicles on road networks, these studies were mainly based on costly and obsolete classical travel surveys on traffic flow and travel times and to some extent failed to provide the necessary information needed by road users in order to cope with the increasing urban demand for mobility [4].

One of the major concerns of traffic managers in traffic management system is traffic volume estimation, a major component of Intelligent Transport System (ITS), as it helps in the decision making and efficient traffic management planning when monitoring the current traffic flows in the road networks. Thus, to reduce the effect of congestion on urban road networks, accurate prediction of the Motorised Traffic Flow as well as traffic estimation is of paramount importance as it provides information on road accidents and level of congestion along the roads [3, 5]. Real time traffic flow data are useful for traffic volume estimation and help in forecasting traffic trends by determining the traffic flow patterns. Traffic data collection, predicting traffic patterns,

and forecasting traffic trends are usually performed for pavement design, fuel-tax revenue projection, and highway planning. However, the monitoring activities necessary for accurate Annual Average Daily Traffic (AADT) estimates are expensive in terms of costs and personnel. Thus, aside providing information for road accidents frequency and congestion, traffic estimation is also an issue for tactical purposes of transportation [6].

Based on the reports on experimental data found in literature, traffic flow patterns are highly predictable and often exhibit irregular and complex behaviours which changes abruptly when entering or leaving a congestion zone [3, 7]. Shang et al. in [8] reported on the irregularity and complexity of traffic flow as one approaches congestion zones in a traffic stream. They stated that the current nature and future dynamics of traffic flows highly depend on continuously interacting properties such as human behaviour and traffic characteristics. [9] noted that some of the main characteristics responsible for the complex behaviour in a traffic flow stream are variations in headways and spacing.

Several methods have been used in time past for short-term traffic flow prediction flows, including ARIMA-type models, Artificial Neural Networks, SARIMA models, Generalised Linear models, Nonparametric Statistical methods, Dynamic Neural Networks, Support Vector Regression models, and STARIMA models just to mention but a few. A brief review of some related work on Traffic Flow Prediction is presented below.

Catriona and Casper in [10] presented a Linear Multiregression Dynamic Model (LMDM) which uses concept of graph for the traffic flow forecasting where the time series of flows at different sites are represented by the nodes and the structure between the flows at different sites as well as the independence is represented by the edges connecting the nodes. The idea of using graphical dynamic model approach in their work for traffic flow forecasting follows from that of [11, 12], respectively, with focus on forecasting traffic flows in two separate motorway networks using UK as a case study. Based on the distinctive features of their the LMDM, their model can be used for testing real-time instances. They illustrated how the LMDM can be used for forecasting and validated their model on some networks. The performance of the proposed approach was compared with other models in literature.

Dauwels et al. in [13] proposed a unified model by developing different forecasting models that is matrix and tensor based by applying partial least squares (PLS), higher order partial least squares (HO-PLS), and N-way partial least squares (N-PLS) for the time series prediction. Their focus was on collective prediction for multiple road segments and prediction-horizons against the known prediction for individual road segments and prediction horizons. One interesting feature of the developed models was the ability to carry out feature selection efficiently and simultaneously carry out traffic condition forecasting for multiple road segments and prediction-horizon. The computational performance of the proposed models which was validated on generic road networks consisting of expressway and arterial roads, in particular, an urban subnetwork in Singapore by

performing a multihorizon speed prediction, showed that the proposed models performed better than the Support Vector Regression (SVR) traditional based model for longer prediction horizons. For the short prediction horizons, lower prediction errors were seen in SVR compared to the PLS based methods with N-PLS achieving higher accuracy when compared with PLS and HO-PLS. In fact their proposed unified models achieved same prediction accuracy as compared to the individual models but can be faster than the traditional based model for moderately sized networks.

ARIMA is one of the most precise methods for traffic flow prediction when compared to other known methods. In particular, Seasonal ARIMA (SARIMA) models have been shown to perform better than the other traditional based models but often times it is faced with some restriction in applicability as a result of using huge historical database for model development. Kumar and Vanajakshi, having this background knowledge in their work in [14], tried to overcome such drawbacks by proposing a prediction scheme approach using the SARIMA model for short-term traffic flow prediction which needs on limited input data for model development. They validated their proposed approach with using both historic and real-time data considering cases where peak period occurred both in morning and evening. The data used for the analysis and model development was from a 3-lane arterial roadway in Chennai, India, with limited flow data from 3 consecutive days. The results of the values of the predicted flows were compared with that of the actual flow values. Thus, the proposed approach will work in most cases where database is a major challenge when using ARIMA for traffic flow prediction model development.

Previous studies have shown that ANN has stable and consistent performance even if there is an increase in the travel time interval for the traffic flow prediction. This was so evident in [15] by Kumara et al. where ANN based model for a neural network was used for short-term prediction of traffic flow with heterogeneous condition for nonurban highway. Their model incorporates speed, density, traffic volume, and time as input variable but considers the speed separately in contrast to most work in literature where average speed of combined traffic flow was considered. For other works in literature that applied Artificial Neural Network or ARIMA-type models for traffic flow prediction, see [16–19]. Moreover, previous and recent research findings have shown that policy makers using existing traffic flow models to predict traffic flows have not been able to mitigate the congestion problem to fairly acceptable levels as expected and hence the need to come up with robust methodology for predicting traffic flows.

Chaos Theory is a novel science paradigm with numerous applications that have not been deeply explored and seems very promising with respect to the analysis and prediction of complex systems like traffic flows, although at the moment little empirical evidence exists to confirm this notion. It can be used to analyze the traffic flow patterns in urban road network by utilizing the intrinsic deterministic nature of the traffic flow in order to reduce congestion on urban road networks. In this paper, we report a systematic review of Chaos Theory and propose an approach to predicting Motorised Traffic Flow on Urban Road Networks based on

Chaos Theory with emphasis on the Largest Lyapunov Exponent method for prediction, the most common, effective, and direct technique of analyzing the presence of Chaos in a given dynamical system. This work contributes to this research field in the sense that the proposed approach is different from other conventional models found in literature and serves as an alternative method for predicting Motorised Traffic Flow on Urban Networks. Also, the effectiveness of the Largest Lyapunov Exponent prediction method seems very promising in terms of prediction accuracy as well as reducing the congestion problems on urban network, although this is yet to be fully validated using computer based algorithm on empirical traffic flow data.

The layout of this paper is as follows. The congestion problem on urban road networks is introduced in Section 1 with brief review of related works on Motorised Traffic Flow Prediction Models. Section 2 gave an insight on Motorised Traffic Flow and Traffic Flow Variability by highlighting some of the main characteristics for the complex behaviour of traffic flow stream. A systematic review of Chaos Theory is presented in Section 3 with emphasis on its application to the analysis and prediction of Motorised Traffic Flow in Urban Road Networks based on the Largest Lyapunov Exponent Prediction Method. The conclusion and directions for future work are drawn in Section 4.

2. Motorised Traffic Flow

2.1. Headway and Spacing. One of the applications of ITS as earlier mentioned is predicting road traffic volumes in order to make efficient traffic management and planning over a network as well as implementing road safety measures. [20] in reporting Shang et al.'s study in their paper noted that the differences in the distribution of various vehicle types, human driving habits (high driver perception-reaction times), space, and time headway, are among the principle causes of chaotic behaviour in traffic flows with the time and space headway been the main factors causing variations in observed traffic distributions and its transformation [9]. A proper knowledge of the above mentioned will be helpful in understanding traffic flow and to some extent provide theoretical foundation for short-term traffic flow forecasting. For the purpose of this study, our focus is on the linking space and time headways and variation of traffic flows on a given road network.

Based on the study carried in [9], suppose we have a traffic stream composed of two consecutive vehicles in a single lane road such that we have a follower-vehicle, i a leader-vehicle, and $i + 1$ as shown in Figure 1.

It can be observed that vehicle, i, is some distance, h_{s_i}, from its pacesetter, $i + 1$, termed as the space headway (usually expressed in metres, m). h_{s_i} comprises of the distance to the leader-vehicle, g_{s_i} (the space gap), and the self-length of the follower-vehicle, l_i. Hence, h_{s_i} is given by

$$h_{s_i} = g_{s_i} + l_i. \tag{1}$$

g_{s_i} is measured from the follower-vehicle's anterior bumper to the leader-vehicle's hind bumper. The hind bumper of
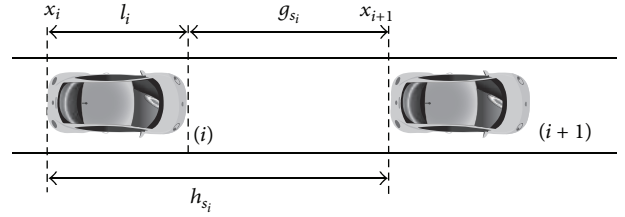

FIGURE 1: Two consecutive vehicles (a follower-vehicle, i at position, x_i and a leader-vehicle, and $i + 1$ at position, x_{i+1}) in a single lane road (after [9]).
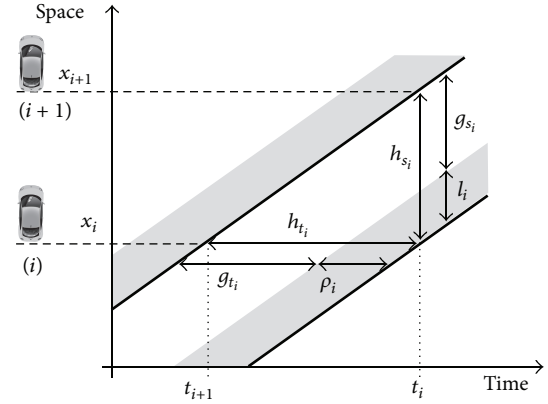

FIGURE 2: Trajectories of a two-car traffic stream (after [9]).

the vehicle represents the vehicle's position. Thus, the space headway, h_{s_i}, can be expressed as

$$h_{s_i} = x_{i+1} - x_i. \tag{2}$$

From (1), each of the two vehicles has also a time headway associated with it. Thus, h_{t_i} (measured in seconds, s) comprises a time difference, g_{t_i}, and a time of occupancy, ρ_i, given by

$$h_{t_i} = g_{t_i} + \rho_i. \tag{3}$$

Both space and time headway can be envisaged in a space time diagram as shown in Figure 2. Thus, the positions x_i and x_{i+1} of the two vehicles, i and $i + 1$, can be plotted with respect to time, tracing out two vehicle trajectories, as the vehicles are in motion.

Figure 2 is called a time-space diagram. The respective speeds of the two vehicles can be derived from the diagram by drawing the tangent line. For simplicity, we assume that both vehicles travel at a constant speed resulting into parallel trajectories.

In single-lane traffic (microscopic traffic model), vehicles always keep their relative order. However, for multilane traffic (macroscopic traffic model), this principle can no longer be obeyed due to overtaking manoeuvres, resulting into irregular vehicle trajectories. If the same time-space diagram were to be drawn for several lanes (in multilane traffic), then some vehicles' trajectories would suddenly appear or fade away at the point where there exists a change of lane.
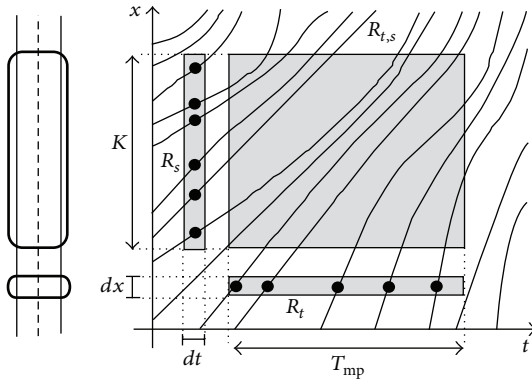
FIGURE 3: A time-space diagram showing nonlinear trajectories of several vehicles where movements are bounded by three regions of measurement, that is, R_t, R_s, and $R_{t,s}$ (after [9]).

FIGURE 4: Vehicle volume distributions by classification of vehicles in California in 1996 (source: [22]).

Figure 3 shows the relative trajectories of vehicular traffic stream in a multilane facility.

In Figure 3, the three regions of measurement are always bounded in both time and space (that is to say, a period of measurement, T_{mp}, and a length of road section, K). Black dots were used to represent single measurements made in the diagram. The following describes what the three bounded regions represent:

(i) R_t represents measurements taken at certain locations that are fixed in space, (dx), done in T_{mp} time period. An example of such a measurement is one obtained by an underground automatic inductive loop.

(ii) R_s represents measurements taken at particular instances in time, (dt), along length, K, of a road section. An example is results taken from aerial photographs.

(iii) $R_{t,s}$ represents a region where general measurements are made. This region normally takes other forms of shape other than a rectangular one (as illustrated in Figure 3). An example of such a measurement is results of video cameras.

It becomes more complicated to represent the vehicle trajectories on the space diagram as a result of the disorderliness in the dynamics of the vehicle movements along the traffic stream. This causes variability in the traffic flow.

2.2. Traffic Flow Variability. Traffic flows are subject to variations over numerous time scales, namely, yearly, monthly, weekly, and daily. It also varies directionally as well as from place to place. Aside the fact that roads carry different volumes of traffic, the characteristics of the vehicles using these roads also change depending on the road facility [21]. For example, one road with about 10,000 vehicles per day may have very little truck traffic, while another road with the same volume of vehicles may have 2,000 trucks per day mixed with 8,000 ordinary cars. Similarly, one road section may be traversed by 1,000 heavily loaded trucks per day while a nearby road is used by 1,000 partially loaded trucks
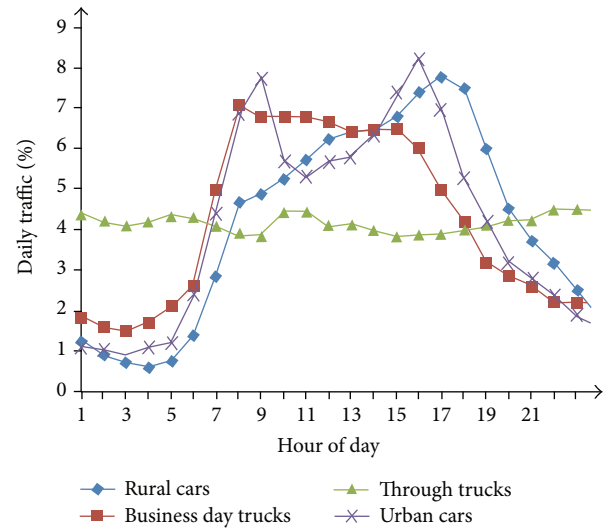
(Traffic Monitoring Guide, 2013). We illustrate with the aid of graph the two major types of traffic volume variations, namely, Time-of-Day and Day-of-Week Variation based on the findings of the Federal Highway Authorities [3].

2.2.1. Time-of-Day Variation. The Federal Highway Administration (FHA) in 1996 reported that most truck travel falls into one of two basic time-of-day patterns, namely, a pattern that is centered on travel during the business hours of a day (working hours) and a pattern that shows almost constant travel all day through (twenty-four-hour day). Figure 4 summarizes the research findings of FHA in 1996.

As can be seen in Figure 4, cars tend to follow either the traditional two-humped urban commute pattern or the single-hump pattern commonly seen in rural areas, where traffic volumes continue to grow throughout the day until they begin to taper off in the evening. However, the truck pattern differs from the rural car pattern; in that it peaks in the early morning (many trucks make deliveries early in the morning to help prepare businesses for the coming workday) and tapers off gradually, until early afternoon, when it declines quickly. The other truck pattern (travel constantly occurring throughout the day) is common with long haul trucking movements. In addition, at any specific location, time-of-day patterns may differ significantly as a result of local trip generation patterns that differ from the norm. For example, Las Vegas, Nevada, generates an abnormal amount of traffic during the night because that city is very active late at night. In heavily congested urban areas, the commute period traffic volume peaks flatten out and can last three or more hours.

A close observation at Figure 4 reveals that cars tend to follow either the traditional two-humped urban commute pattern (double peaked pattern) or the single-hump pattern commonly seen in rural areas, where traffic volumes continue to grow throughout the day until they begin to taper off in
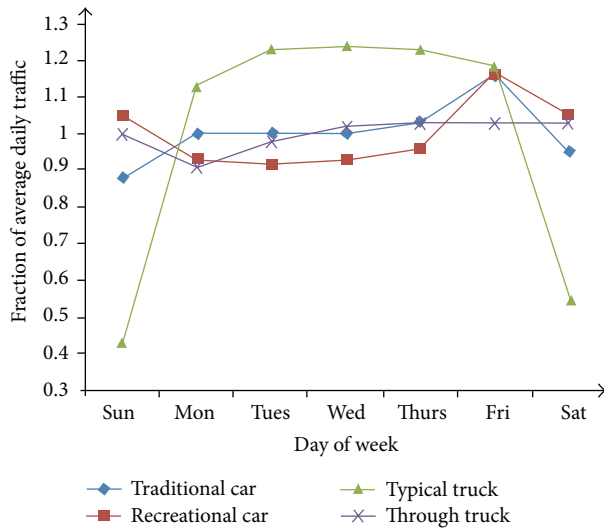
FIGURE 5: Day of the week traffic variations in California in 1996 (source: [22]).

the evening. However, the truck pattern differs from the rural car pattern; in that it peaks early in the morning due to the fact that many trucks make deliveries early in the morning to help prepare businesses for the coming workday and tapers off gradually, until early afternoon, when it declines quickly. The other truck pattern maintains a constant pattern throughout the day which is mostly common with long haul trucking movements.

Moreover, time-of-day patterns usually differ significantly with respect to places at any specific location as a result of local trip generation patterns that differ from the norm. For example, a city with night clubs or recreational facilities will generate an abnormal amount of traffic during the night hours or other hours of operation because that city is very active late at night. Also, in heavily congested urban areas, the commute period traffic volume peaks flatten out and can last three or more hours.

2.2.2. Day-of-Week Variation. The same study also revealed that there exists a large difference in daily patterns of the ordinary vehicle categories and typical trucks since truck travels are mainly business motivated as opposed to ordinary vehicles whose drivers have several travel objectives. Figure 5 illustrates the day-of-the-week variations.

It is evident from the graph that the day-of-week traffic variations are highly responsible for the traffic congestion that comes in form of jams on urban roads. A good example is the stampede observed along Kwame Nkrumah circle in Accra, Ghana, whose immense traffic jams are estimated to have caused annual losses of about $125 million to travelers along this road in 2014 (monetary value of lost time during traffic jam), as pointed out by traffic experts of the Ghana Institute of Engineers [23].

To mitigate this problem of congestion on urban roads, it is very necessary to carry out a substantial traffic estimation which requires a method of high precision to forecast a

complex entity such as traffic flow. This is the main reason for proposing an alternative way of addressing complex systems like traffic flows, using effective techniques based on Chaos Theory (which studies dynamic systems) to analyse and predict traffic flow patterns.

3. Chaos Theory Review

3.1. Introduction. Several systems exist in everyday life that evolve with time. Such systems are difficult to predict accurately on long-term scale even with robust statistical prediction models. Examples of such system include weather, turbulent fluids (flowing across planes), population infected by epidemic, and stock market indices and they are generally referred to as dynamical systems [24].

These systems are said to exhibit "Chaos." Chaos in a simple term refers to any state of confusion or disorder that is showing the absence of some kind of particular order. Many work exists in literature that addresses dynamical systems as well as the chaotic behaviour. Kiel and Elliott in [25] described how many disorganised systems can spontaneously acquire organisation. For example, a shapeless liquid mass upon cooling can be transformed into an exquisite shape. Zhang and Jarrett in [26] studied the dynamic behaviour of road traffic flows in an origin-destination network. Their proposed dynamic model is a modification of the static conventional model by Dendrinos which also describes the traffic flow variability of the O-D network flows. They showed that the O-D flow patterns varies depending on whether the dimension is lower or higher. The characterization of the chaotic attractors by positive Lyapunov Exponents and fractal dimensions agrees with the fact that Largest Lyapunov Exponents provide the best measure of Chaos in any dynamical system. See [27] for details on the search for chaos in traffic-flow dynamics.

A Chaotic system can be described as one that is complex, aperiodic (it never exactly repeats), and sensitive to its initial conditions. Chaos Theory is novel Science paradigm in the field of nonlinear analysis which is used to describe the realms of nonrepeating and highly complex dynamic systems. This discipline is accredited to a meteorologist from the Massachusetts Institute of Technology (MIT), Muhmoudabadi [6], who described Chaotic systems to sensitively depend on initial conditions. He termed this behaviour as "*The Butterfly Effect*" (where the flap of butterfly's wings in Brazil sets off a tornado in Mexico). For clarity purposes, Chaotic processes should not be confused with random processes because Chaos does not imply randomness in any sense. Chaotic processes do not have any kind of distribution like random processes such as Brownian motion that exhibit a Gaussian distribution [28]. Furthermore, Chaotic processes are perfectly deterministic while random process are attached to some prior probabilities. Some properties of chaotic systems outlined below will help in understanding the behaviour of Chaotic systems.

3.2. Properties of Chaotic Systems. Chaotic systems have a number of distinctive characteristics which are used to

describe the dynamic evolution of such systems. These characteristics include the following.

Sensitivity to Initial Conditions. As already introduced in Section 3.1, Chaotic systems are highly dependent on initial conditions, a property sometimes regarded as "*The Butterfly Effect.*" Two trajectories emerging from two different close-by initial conditions diverge exponentially from one another as the system evolves in phase space (a phase space is a representation of all possible states (configurations) of a dynamic system, and each possible state mapped by unique points [29]) [30]. In order to make accurate prediction of long-term behavior of Chaotic systems, the initial conditions must be known in their entirety and to high levels of precision.

Determinism. Chaotic systems are strictly deterministic. A deterministic system is one where for a given time interval there is only one future state that follows from the current state [31]. These systems can be described by Ordinary Differential Equations (ODE's). At least three variables are needed for Chaos in continuous-time systems as opposed to Chaos in discrete systems that requires only a single variable [29]. The reason is that the space time trajectories have to be aperiodic and finitely bounded in some region. However, it is unlikely to have a single trajectory intersecting itself due to the fact that every point has a unique mapping in space [29].

Nonlinearity. Intuitively, a nonlinear system is a system whose outputs and inputs are not proportional to each other. In other words, a nonlinear system is a system which cannot be decomposed into parts and reassembled into the same thing. This is a situation where the relationship between variables describing a system is not simply static or directly proportional to the output, but instead it is dynamic and varies [32]. Nonlinear dynamic systems exhibit nonlinear time series (discussed later in Section 3.4). In the case of nonlinearity, there is no periodicity (nonrepetitive system) as compared to linearity where the system repeats itself over a time period.

Instability. Chaotic systems have a sustainable irregular manner caused by sensitive dependence on initial conditions and thus predictions for a given system can only be made on short-term scales to high precision [29].

Attractors. These are d-dimensional sets of states, $\mathcal{X} \in \mathbb{R}^d$ (points in phase space) invariant under the system's dynamics where all states in close proximity asymptotically approach each other [33]. Many dynamic systems in nature have attractors and it has been discovered by researchers that all Chaotic systems' dynamics of evolution emerge into a certain type of attractors called strange attractors which are sensitively dependent on their initial conditions [24]. The four known types of attractors are briefly described as follows:

(i) Point attractor: a system is said to have a point attractor if the system evolves to a fixed point, for example, a single singing pendulum bob (see Figure 6(a)).
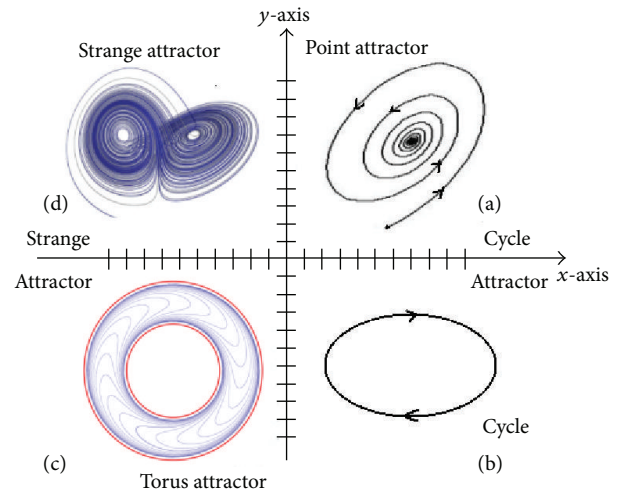

FIGURE 6: Different types of attractors constructed in 2-dimensional phase space; (a) point attractor, (b) limit cycle, (c) limit torus, and (d) strange attractor (after [29]).

(ii) Limit cycle: if the system is cyclic and its position in the cycle can be predicted, then the system is said to have a limit cycle, for example, planetary motions (see Figure 6(b)).

(iii) Limit torus: a system that has a limit torus is similar to that of a limit except that the system's trajectories are bounded within a region of a ring torus; for example, the "halo" ring of planet Jupiter is a torus composed of mainly dust particles in motion (see Figure 6(c)).

(iv) Strange attractor: if a system takes an aperiodic irregular shape and never repeats itself in time, the system is said to have a strange attractor. Such an attractor can also be described as a limit region (object with fractional (fractal) dimension) within phase space which is ultimately occupied by all trajectories of a dynamical system. Examples of such strange attractors include the famous Lorenz attractor illustrated in Figure 6(d) [30], Hénon attractor, and logistic map attractor.

Fractal Dimensionality. It is an already established fact that that the geometrical dimension of a line, plane, and box is 1, 2, and 3, respectively. However, many examples seen in our everyday life as well as many objects are not geometrically smooth like the ones mentioned above. Complex, noninteger dimensions are called fractal dimensions [35]. This is usually used to measure the complex nature of a given Chaotic system. When a Chaotic system's evolution is represented in phase space, the topological dimension, d, of the space state of the system's trajectories is a noninteger. A famous example of a plot with fractal dimension is Mandelbrot's plot ($z = z^2 + c$), which lies in the category of fractals, which are shapes that infinitely repeat themselves in smaller magnifications (scales) [34]. Figure 7 is an illustration of Mandelbrot's plot in a 2-dimensional complex plane.
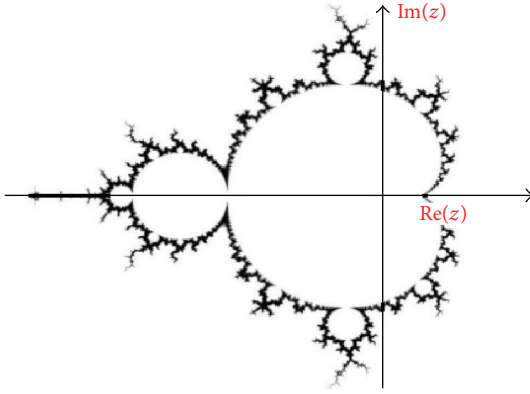
FIGURE 7: Mandelbrot's plot that is self-replicating according to some predetermined rule such that the boundary of the set has fractal dimensions (drawn in a 2-dimensional complex plane) (after [34]).

Other examples of shapes in nature with fractal dimensions include coastlines and slow flakes [34].

To summarize the properties of Chaotic systems, we note that there are two important characteristics that make chaotic systems very complex and our focus is on these characteristics:

(i) The strange attractor, which contains a large number of unstable system trajectories.

(ii) The ergodicity (ergodicity is a system behaviour that is averaged over time and space for all the system's states) in the dynamics of the system trajectories. In other words, as the system evolves temporarily, a small neighbourhood of every point in one of the unstable orbits within the attractor is visited [29, 36].

We note that in Chaos Theory, there is no need for prior knowledge of probabilities unlike in statistical physics. Under appropriate circumstances, it has been reported that algorithms based on Chaos Theory have shown the capability of attaining high level of performance, far better than those obtained using classical stochastic methods or techniques based on signal processing, and these can be applied in the following areas among others [37].

In meteorology, Chaos Theory is used to predict slight changes in weather, air, and aerosol movements in the atmosphere and so forth as studied by Lorenz in the late 1960s [30]; it is used in most biological processes such as heart beat detection, circadian rhythms, in particular, and electrocardiographic recording of a pregnant woman [29]. In economics and finance, Chaos Theory is used in foreign exchange rates and stock market indices for market crash forecasting. This is based on the Mandelbrot fractal hypothesis which predicts a market crash every two decades starting from 1987 up to date [34]. Moreover, Chaos Theory is also applied in traffic flow predictions, which is still an open and new area for research opportunities. This is the main motivation for this review [38].

3.3. Limitations and Control of Chaotic Systems. Although Chaotic systems have good characteristics that are suitable for analysis of complex behaviours, there are significant factors that hinders one from accurately predicting the behaviour of complex system. They include sensitive dependence on initial conditions which are in most cases unknown as most assumptions made often lead to error, the current stage of this "new" discipline of science (just half a century old) as one is not yet very sure of how much data is required to precisely reconstruct phase space and determine the fractal dimension of a given system (discussed in Section 3.4.1), the nature of the calculations involved in Chaos Theory which are repetitive, high extensive, and tedious which can only be done with the help of computers with high accuracy and precision [29, 30, 39].

Chaos systems can be controlled in order to reduce computational errors due to the adverse effects of the above limitations. Shewalo et al. in [29] stated that Chaos in systems can be controlled in exactly three ways which we have summarised below without full details (see [29] for full details). First, the systems parameters can be changed heuristically so that the range of fluctuations is limited. Secondly, one can apply perturbation to the Chaotic system which causes the system to organise itself using Ott-Grebogi-Yorke method, and finally the relationship between the system and the environment is changed using Pyragas method.

Having established the fact that Chaotic systems exhibit nonlinearity property with time evolution in previous sections, we now briefly describe the Chaotic Time Series and how it can be applied in the prediction of dynamical systems such as traffic flow based on the nonlinearity concept.

3.4. Chaotic Time Series Prediction. Phase space dynamics can be used to analyse and make predictions of dynamical systems. Nonlinear processes resulting in higher dimensional objects (called attractors when drawn in phase space) are characterised by nonlinear time series that intrinsically describe the behaviour of the system under study [40].

One can make prediction for a given time series using phase space techniques which is often referred to as the determinism test of a system. Such techniques are based on the fundamental fact that trajectories in close proximity asymptotically approach each other within the phase space [7]. A dynamical system can be represented in an m-dimensional finite dimensional vector space, $\mathbb{R}^m$, by the following equation:

$$\{X_{n+1}\} = F\left(X_n\right)_{\{n=0,1,2,\dots\}}, \tag{4}$$

where n is the dimension of the vector space, $\{X_n\}$ represents n-phase space time evolution points, and $F(X_n)$ is an arbitrary function representing the system's behaviour (but is usually unknown). This is as a result of the fact that, in most cases, elements of X_n are very difficult to observe empirically; that is, one may only be able to measure a single variable for a given time series and still have no explicit knowledge on the system's nonlinear dynamics [20].

Taking a look at traffic systems in particular, they highly depend on human and physical factors in a given road facility

and this even becomes more complicated due to the presence of immeasurable quantities such as traffic laws and social codes. Nevertheless traffic flow patterns are deterministic and their time series have been found to be nonlinear [4, 20].

Since Chaotic systems exhibit a nonlinearity property, developing a Chaos prediction model for a given dynamical system is based on nonlinear time series analyses which mainly involves two steps, that is,

(i) reconstruction of the phase space from a given data set,

(ii) developing of a methodology for predicting the phase space dynamics.

These steps can be explored following Takens' Fundamental Embedding Theorem (from 1981) [33]. We note that the reconstruction of the original time series data is done using this theorem known as the foundation of all Chaos based predictions [36].

Theorem 1 (Takens' Theorem). *Let $\mathcal{X} \in \mathbb{R}^d$ be a bounded set. In the Cartesian product space of C^1 mappings on $\mathcal{X}$ and the C^1 function $\mathcal{X} \to \mathbb{R}$, there exists an open and dense subset, U such that if $(T, f) \in U$, then the reconstruction map, $R_{map}^{(m)}$, is embedding whenever $m > 2\dim(\mathcal{X})$. Moreover the embedding is continuously differentiable and also has a differentiable inverse (C^1 diffeomorphism). We have a deterministic system, $T : \mathcal{X} \to \mathcal{X}$, and we also have a read-out function, $f : \mathcal{X} \to \mathbb{R}$. If $m > 2\dim(\mathcal{X})$, there exists a precise deterministic rule, g, for predicting the next state of a time series.*

Interpretation of Takens' Theorem. The proof of Takens' Theorem is omitted in this work but the following definitions and interpretation of the theorem will give an understanding of the theorem.

Definition 2 (a diffeomorphism). A diffeomorphism is a map between manifolds (smooth space system states), which is differentiable and has a differentiable inverse.

$\mathcal{X}$ is called the attractor set corresponding to the following time series:

$$\{x_t\} = \{x_0, x_1, x_2, \ldots\} = \{f(x_0), f(x_1), f(x_2), \ldots\}, \quad (5)$$

and we can rebuild the system's dynamics by the rule, g, which states that

$$\dim(\mathcal{X}) \le m \le 2\dim(\mathcal{X}) + 1, \quad (6)$$

where all the d-dimensional manifolds (space-states) of the system's attractor, $\mathcal{X}$, can be embedded in an $m = (2d + 1)$-dimensional reconstructed space while preserving the geometrical invariants, and $d = \dim(\mathcal{X})$.

This simply means that all the information about the system's complex d-dimensional attractor can still be captured in the discretized reconstructed m-dimensional phase space. Based on the knowledge of the outcomes of Theorem 1, we can now determine the topological parameters of the system's attractor.

3.4.1. Reconstruction of Phase Space. During reconstruction, new space states are created that are (in the sense of diffeomorphisms) equivalent to the original space states so that the relevant geometrical properties of the system are always preserved. The set of reconstructed trajectories, X, corresponds to a matrix in which each row is a vector in phase space; that is,

$$X = \left[\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_i, \ldots, \vec{X}_M\right]^T, \quad (7)$$

where $\vec{X}_i$ is the system state at discrete time, i, and for a real time series with N-points, $\{x_1, x_2, \ldots, x_N\}$, each $\vec{X}_i$ is denoted by

$$\vec{X}_i = \left[x_i, x_{i+\tau}, x_{i+2\tau}, \ldots, x_{i+(m-1)\tau}\right], \quad (8)$$

where τ is the reconstruction delay time (lag) and m is the embedding dimension.

Therefore the matrix X is $M \times m$ matrix given by

$$X = \begin{pmatrix} x_1 & x_2 & \cdots & x_M \\ x_{1+\tau} & x_{2+\tau} & \cdots & x_{M+\tau} \\ x_{1+2\tau} & x_{2+2\tau} & \cdots & x_{M+2\tau} \\ \vdots & \vdots & \vdots & \vdots \\ x_{1+(m-2)\tau} & x_{2+(m-2)\tau} & \cdots & x_{M+(m-2)\tau} \\ x_{1+(m-1)\tau} & x_{2+(m-1)\tau} & \cdots & x_{M+(m-1)\tau} \end{pmatrix}^T, \quad (9)$$

where the constants M, N, m, and τ are related by the equation $M = N - (m-1)\tau$ and by Theorem 1, $m > 2d$ where d is the dimension of the system's attractor.

Now, suppose we have a scalar observed nonlinear times series, say from empirical traffic data,

$$\{x_t\}_{\{t=1,2,3,\ldots,N\}}. \quad (10)$$

The vector for each reconstructed single point time series is given by

$$X_t = \{x_t, x_{t+\tau}, x_{t+2\tau}, \ldots, x_{t+(m-1)\tau}\}, \quad (11)$$

and it follows that

$$X_{t+1} = \{x_{t+1}, x_{t+1+\tau}, x_{t+1+2\tau}, \ldots, x_{t+1+(m-1)\tau}\}, \quad (12)$$

where τ is the time delay and m is the embedding dimension (as before), and that

$$[t + 1 + (m-1)\tau] \le N. \quad (13)$$

Consider Figure 8 illustrating the time series, $\{X_t\}$ and $\{X_{t+1}\}$ in a time-space diagram, and a phase space diagram, respectively.

Figure 8(b) gives a probable representation of the strange attractor of the data set, whose result is set of points of the above two time series plotted in 2-dimensional phase space. However the trajectories of the attractor (as in the diagram) may appear to intersect each other but they actually never cross even in higher dimensions.
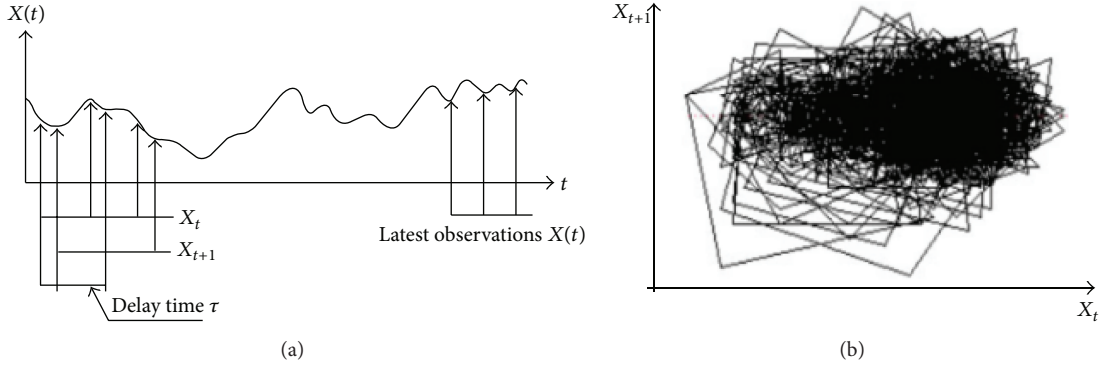
FIGURE 8: Illustration time series, $\{X_t\}$ and $\{X_{t+1}\}$, plotted in a space time (a) and a 2-dimensional phase space plot (b) (after [39]).

Parameters (topological parameters) such as the dimension of d, the attractor, τ, the delay time, and m, the embedding dimension, are necessary for reconstruction of the systems' dynamics in phase space before any predictions can be made. These parameters can be determined by the procedures in the order described below. First, we compute the delay time, τ, as follows.

(i) Determination of Delay Time, τ. There are several approaches for determining the delay time. The first approach as pointed out in [33] is by computing the Auto Correlation Function (ACF) of the data given by the following equation:

$$\mathscr{C}(r) = \frac{1}{N}\sum_{i=0}^{N-1}(x_i - \langle x\rangle)(x_{i+1} - \langle x\rangle), \qquad (14)$$

where $\langle x\rangle$ is the arithmetic mean of the observations, given by

$$\langle x\rangle = \frac{1}{N}\sum_{i=0}^{N-1}x_i. \qquad (15)$$

The choice of τ is determined by the duration after which x_i (or x_t) and x_{i+1} (or $x_{t+\tau}$) become uncorrelated, although [20] claims that it is difficult to obtain this.

Another method of determining τ is to calculate the nonlinear Auto Correlation Function called the Average Mutual Information (AMI), $I(\tau)$. AMI is a standard technique that tells us how much information we can obtain about a measurement taken from one time series, say $\{x_t\}$, that is affected by another measurement taken from another time series, $\{x_{t+\tau}\}$, sampled after a time interval, τ [41]. In other words $I(\tau)$ is a measure of the mutual dependence between two time series, and it is given by

$$I(\tau) = \sum_{ij}p_{ij}\ln p_{ij}(\tau) - 2\sum_{i}p_i\ln p_i(\tau), \qquad (16)$$

where p_i is the probability that x_t takes the ith bin of a histogram, p_{ij} is the probability that x_t is in the ith bin, and $x_{t+\tau}$ is in the ith bin.

The concept of bin in a histogram will help in understanding how the information is obtained. We define a bin of
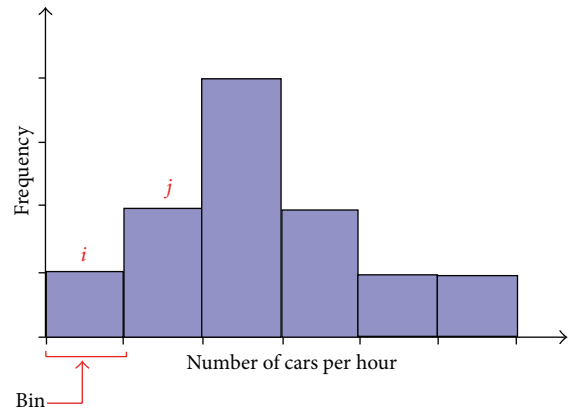

FIGURE 9: An example of a histogram plot illustrating arbitrary bins, i and j (after [42]).
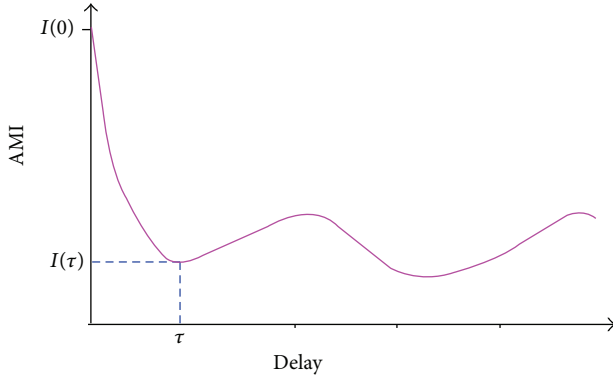
a histogram intuitively with the following example. The bar graph of a histogram simply shows how many data points fit within a certain range. That range is called the bin (sometimes called the bin width). See Figure 9.

For instance, suppose we want to plot a histogram graph after counting the number of cars passing through a certain area per hour. Using histogram chart in Figure 9, we might decide to plot it using the intervals 1–10, 11–20, 21–30, and so on. In this case, our bin would be 10 and every bar on your histogram represents a range of ten cars. The same data could be plotted on a range of 5 as 1–5, 6–10, 11–15, and so on. Here, our bin would be five. Obviously, the smaller the bin is, the more information we obtain about our data set, and vice versa. The narrower the bin is, the more you miss out on the point of a histogram.

Thus, we can compute the above probabilities (p_i and p_{ij}) and hence $I(\tau)$, by the Fraser and Swinney (1986) algorithm that is fully described by [42]. This algorithm can be directly applied to a given time series.

$I(\tau)$ is plotted against increasing values of τ and this plot is known as the AMI graph. This takes a shape such as the one illustrated in Figure 10.

To obtain the most appropriate value of τ, the first minimum in the AMI graph is chosen. This is because the first

FIGURE 10: AMI against τ plot (after [41]).

FIGURE 11: A 2-dimensional plot of the Hénon attractor showing $x^{(b)}$ called the FNN of $x^{(a)}$ and $x^{(c)}$ called the TN of $x^{(a)}$ (after [41]).

minimum preserves both the independence and correlation of the values of the two time series of x_t and $x_{t+\tau}$ and with this we can have a good approximation of the coordinates for the reconstructed vectors [7].

Claim. In [43], the criterion suggests that $I(\tau)/I(0) \approx 1/5$ if τ time series works well for down sampled data. $\tau \leq 5T_s$, where T_s is the sampling time of the data set.
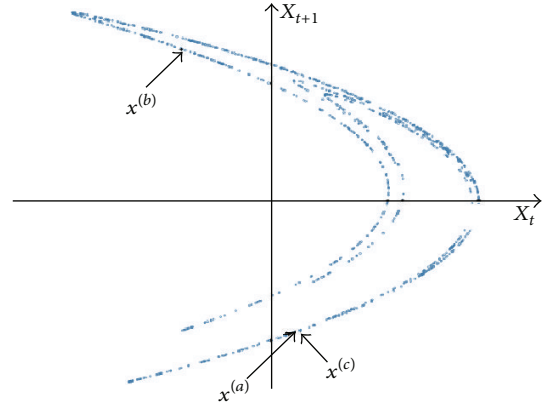
Next, we compute the embedding dimension, m, as follows.

(ii) Determination of Embedding Dimension, m. This is done by computing the False Nearest Neighbours (FNN) Method [43]. This method is based on the assumption that two points that are in close proximity in the appropriate embedding dimension, m, must remain close as we move to higher dimensions [44]. However, if the embedding dimension is too small, then the points that are truly farther apart could seem to be neighbours, and such points are known as FNN.

Now, suppose 2 points, $x^{(a)}$ and $x^{(b)}$, are in close proximity in phase space. We compute the Euclidean distance of the 2 points given by $|x^{(a)} - x^{(b)}|$ in 2 consecutive embedding dimensions, m_0 and $m_0 + 1$ for $(m_0 \geq 2)$. Then, we determine whether a certain ratio (which is a function of the Euclidean distances in dimensions, m_0 and $m_0 + 1$) is greater than some predetermined value. One detects FNN within a given vector when the points close in dimension, m, move a significant distance apart in the following state while doing the computation. In dimension, m_0, the Euclidean distance is obtained as follows:

$$R_{m_0}^2 = \sum_{m=1}^{m_0-1} \left\{ x^{(a)}\left(t + m\tau\right) - x^{(b)}\left(t + m\tau\right) \right\}^2. \quad (17)$$

Moving from dimension, m_0, to dimension, m_0+1, means that position of points in phase space changes by an amount equal to $x(t + d\tau)$ and this has a contribution to each delay vector. It follows that the Euclidean distance in the dimension, $m_0 + 1$, is given by

$$R_{m_0+1}^2 = R_{m_0}^2 + \left| x^{(a)}\left(t + d\tau\right) - x^{(b)}\left(t + d\tau\right) \right|^2. \quad (18)$$

The relative distance between the 2 dimensions gives the following relationship (a ratio):

$$\frac{\sqrt{R_{m_0+1}^2 - R_{m_0}^2}}{R_{m_0}} = \frac{\left| x^{(a)}\left(t + d\tau\right) - x^{(b)}\left(t + d\tau\right) \right|}{R_{m_0}}. \quad (19)$$

Based on this criterion, [43] states that if the ratio in (19) above is found to be greater than some predetermined value, R_{tol}, called the tolerance threshold, then the points $x^{(a)}$ and $x^{(b)}$ are characterised as "False Nearest Neighbour" (FNN).

In the same way, $R_{m_0+1} > \sigma/R_{\text{tol}}$, where σ is the statistical standard deviation of the attractor's time series data set around the mean, $\langle x \rangle$.

Claim. Reference [43] showed that, for several dynamical systems, $R_{\text{tol}} \approx 15$.

The authors of [20] stated that the claim presented in [43] was later empirically confirmed by a study on the eruption of Vatnajökull volcano of Iceland that $9 \leq R_{\text{tol}} \leq 17$, and a value of $R_{\text{tol}} = 10$ has proved to give good results. Thus, FNN is calculated for a given observed time series to determine the sufficient delay time necessary for phase space reconstruction.

Consider Figure 11 showing the Hénon attractor to help us intuitively understand the difference between FNN and "True Neighbours" (TN).

The above procedure is repeated for all possible pairs of points in dimensions of ascending order until the fraction of FNN drops to zero (or gets close to zero), a process usually termed as "unfolding" of the attractor. The percentage of FNN should drop to zero when the appropriate embedding dimension, m, is achieved.

For a given dynamical system such as traffic flow, a suitable value of R_{tol} has to be chosen although 10 is usually the best value as stated above. Based on this criterion, we note that a graph of the percentage of FNN against increasing values of embedding dimension, m, is plotted, which takes a shape similar to the one illustrated in Figure 12.

Normally, the value of m corresponding to the first minimum value of FFN% (for curve (a) in Figure 12) above
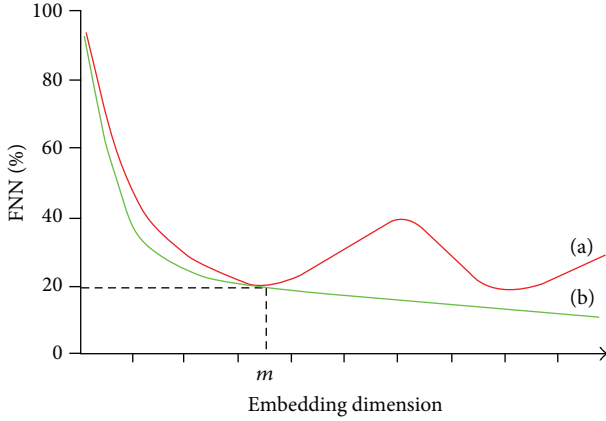
FIGURE 12: Plot of the percentage of FNN against embedding dimension, m (after [41]).

zero is taken as the most appropriate embedding dimension of the reconstructed time series. This is because by then the percentage of FNN has substantially reduced and the attractor is unfolded.

Noise Reduction. In the case of clean Chaotic data (having no random noise), it is expected that the percentage of FNN is reduced to zero when the proper embedding dimension is found. If the time series data is too noisy, however, it is likely that the method fails due to futile attempt of trying to unfold the noise in the data. Apart from determining the optimal embedding dimension, m, the FNN method is a good indicator of a noisy data set. From Figure 12, if FNN% converge in the range of increasing values of m (i.e., $\lim_{m \to \infty}$ FNN% $\to 0$) as shown in curve (b) of Figure 12, then there is high possibility of random noise, which may be responsible for spreading the data, and therefore, it needs to be filtered [45]. As a stochastic process, noisy data must not unfold at any given dimension in phase space (in this case, we have no clear-cut minimum). Moving average and low-pass filter are commonly used methods for noise reduction in data sets although it is not discussed in this work [6].

We now discuss the different methodologies for prediction of Chaotic system's behaviour having discussed the topological parameters of the attractor.

3.4.2. Methodology for Prediction. Literature suggests that it is very necessary to check for Chaos in a given data set before predictions are made. The reason for the check is that there might be presence of random data, which are often assumed to be chaotic, in the data set.

There are several methods used to test for Chaos in a time series data set of a dynamical system. The following methods covered in this work were briefly discussed. They include computation of the (i) Correlation Dimension, d_c; (ii) Hurst Exponent, $\mathscr{H}$; (iii) Kolmogorov Entropy, K; and (iv) Largest Lyapunov Exponent (LLE), $\lambda_{\max}$.

(i) Correlation Dimension, d_c. This method has been widely used by physicists to test for Chaos in dynamical systems [33].

It provides a measure of which points in a given data set of an attractor affect each other. This parameter provides one of the best measures used in differentiating between stochastic and Chaotic systems.

The Correlation Function, $\mathscr{C}(r)$, is given by

$$\mathscr{C}(r) = \lim_{N \to 0} \frac{2}{N(N-1)} \sum_{i,j=1}^{N} H(u), \qquad (20)$$

where $H(u)$ is the Heaviside step function given by

$$H(u) = \begin{cases} 1; & u > 0 \\ 0; & u \le 0, \end{cases} \quad u = r - \left| X_i - X_j \right|, \qquad (21)$$

where r is the radius of the sphere whose center is at X_i or X_j and N is the number of points in the reconstructed attractor's data set.

If the time series is characterized by an attractor, then

$$\mathscr{C}(r) \propto \mu r^{d_c}, \qquad (22)$$

where μ is a constant of proportionality and d_c is the Correlation Dimension or the gradient of the $\log \mathscr{C}(r)$ against $\log r$ plot denoted by

$$d_c = \lim_{r \to 0} \frac{\log \mathscr{C}}{\log r}, \qquad (23)$$

where d_c can also be estimated by the method of least squares or a smooth line over a certain range of m values referred to as the *scaling region*. This region can be estimated by determining the local slope given by

$$d_c = \frac{d[\log \mathscr{C}]}{d[\log r]}. \qquad (24)$$

Reference [33] states that d_c provides the lower bound of the dimension, d, of the attractor and satisfies the inequality

$$d_c \le d. \qquad (25)$$

To observe the existence of Chaos in the data, a plot of the Correlation Dimension against increasing embedding dimension values is obtained. The plot takes such a shape as illustrated in Figure 13.

If $d_c \in \mathbb{R} \setminus \{\mathbb{Z}\} < \infty$, then Chaos exists in the data set. The closest integer above the scaling region of the curve gives the least value of phase space variables used in the modeling of the actual dimension, d, of the attractor.

Note. If d_c is unbounded and is observed to increase with increasing embedding dimension, m, that is, $\lim_{m \to \infty} d_c \to \infty$, then the system is considered to be stochastic.

Now we define an upper bound for the d-dimension of the attractor called the *Limit capacity*, d_l, which satisfies the following inequality:
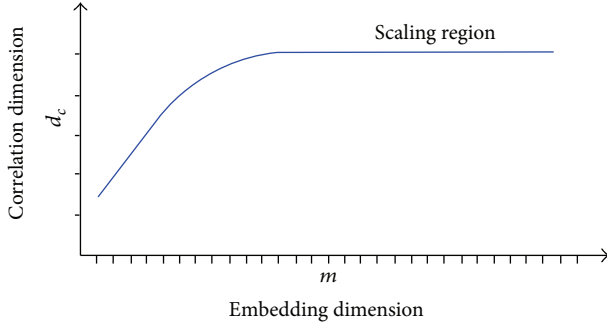
$$d_c \le d \le d_l. \qquad (26)$$

FIGURE 13: Plot of the Correlation Dimension, d_c, against embedding dimension, m (after [20]).

To determine d_l, we let $N(\epsilon)$ be the number of spheres of radius, r, for $0 < r < 1$, such that all the points of the attractor are covered by the spheres. Then it follows that

$$d_l = \lim_{\epsilon \to 0} \frac{\ln N(\epsilon)}{\ln(1/\epsilon)}. \tag{27}$$

In practice, we do not know the prior dimension, d, of the attractor and the most appropriate value of m of the newly reconstructed dynamics. Therefore, the dimensional estimate of d is found by increasing values of m (starting with $m = 2$) until a stable value of d is achieved (as described in Section 3.4.1(ii)).

(ii) Hurst Exponent, $\mathcal{H}$. Similar to the Lyapunov Exponents, a well-established parameter that is commonly used for testing for the Chaos in systems is the Hurst Exponent [38].

The Hurst Exponent, $\mathcal{H}$, is a measure of the degree to which a given time series can be statistically expressed as a random walk (i.e., Brownian motion).

If a time series vector, x_t, on average moves away from its original position by an amount that is directly proportional to $\sqrt{\Delta t}$ (where Δt represents a time interval), it is said that its Hurst Exponent is 1/2 as stressed by [39] in reporting Kantz and Schreiber's work of 1997.

Therefore, one can determine whether the time series data is randomly distributed or not. This is obtained through the square root relation between increments after a certain time interval as follows:

$$\Delta x^2 \propto \Delta t^{2\mathcal{H}}, \tag{28}$$

where $\mathcal{H}$ is the Hurst Exponent and $0 \le \mathcal{H} \le 1$ and Δt is the time interval.

Reference [45] claimed that the relationship between the Hurst Exponent, $\mathcal{H}$, and Correlation Dimension, d_c is

$$d_c = 2 - \mathcal{H}. \tag{29}$$

In a data set where $\mathcal{H} = 1/2$, we conclude that the data is randomly distributed and is not correlated, while for $\mathcal{H} > 1/2$, we say that the data set has a positive correlation, and finally when $\mathcal{H} < 1/2$, the time data set has negative correlation.

(iii) Kolmogorov Entropy, K. A change in volume gives information about the sum of the corresponding Lyapunov Exponents which is equal to the Kolmogorov Entropy, K, given by

$$K = \sum_{\lambda_i > 0} \lambda_i, \tag{30}$$

where λ_i is the spectrum of Lyapunov Exponents (seen later in Section 3.4.2(iv)) [46].

For $N(\epsilon)$-number of spheres (as defined before in part (ii)) and embedding dimension, m, if a time series is completely deterministic (Chaotic), then

$$\lim_{N(\epsilon) \to \infty} \lim_{m \to \infty} K \longrightarrow 0. \tag{31}$$

On the other hand, for a completely random time series, the value K will not converge to single value, that is, $(\lim_{N(\epsilon) \to \infty} \lim_{m \to \infty} K \to \infty)$. Therefore, lower values of K imply higher predictability of the system and vice versa.

(iv) Largest Lyapunov Exponent, $\lambda_{\max}$. As far as we know, computation of Lyapunov exponents provides the best measure of Chaos in any dynamical system [46]. For this reason, we are going to explicitly explain and focus on this method since it is the most direct and most effective technique used for analysing the Chaotic behavior in a given dynamical system which is helpful in making predictions. Lyapunov exponents can clearly explain all the information contained in a time series. Thus, can be used to determine the length of the predicting period for any dynamical system, as argued out by [20].

Having established that the exponential divergence of nearby trajectories is the hallmark of Chaotic behaviour as explained by [30], the Lyapunov spectrum of exponents is given by

$$\lambda_{i\{i=1,2,\dots,n\}}, \tag{32}$$

where n is the number of points in the reconstructed data set.

If the exponents are arranged in descending order such that

$$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n, \tag{33}$$

then the following relationships are true:

(i) The length of the principle axis of spectrum is proportional to $e^{\lambda_1 t}$.

(ii) The area determined by 2 principle axes is proportional to $e^{(\lambda_1 + \lambda_2)t}$.

(iii) The volume of the first k-principle axes is proportional to $e^{(\lambda_1 + \lambda_2 + \cdots + \lambda_k)t}$,

where t is time interval for the system to evolve from one state to another in phase space.

To understand the above relationships, we compute the he Euclidean distance between 2 points in phase space. Suppose
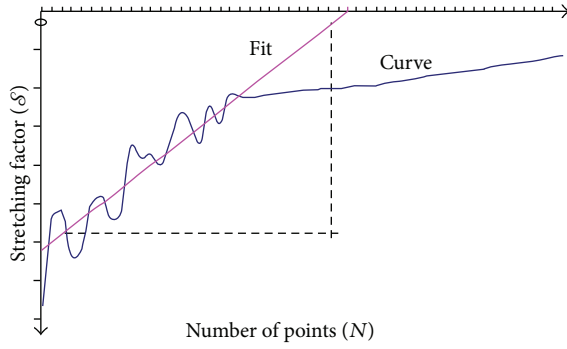
FIGURE 14: A plot of the Stretching Factor, $\mathcal{S}$, against number of points, N, in the data set (after [20]).

that originally we have 2 points in phase space that is $x^{(n_0)}$ and $x^{(n_1)}$ whose Euclidean distance is given by

$$\left| x^{(n_0)} - x^{(n_1)} \right| = \delta_0. \qquad (34)$$

After a time interval, t, the system evolves and the new distance is given by $\delta = \delta_0 e^{\lambda_1 t}$, where $\lambda_1 > 0$, called the Lyapunov exponent. Thus, computing the Euclidean distances between points in consecutive higher dimensions will give the area and the volume, respectively.

Our focus is mainly on the Largest Lyapunov Exponent (LLE), $\lambda_{\max} = \lambda_1$, which gives evidence for determinism of a given system. In reporting Rosemstein et al.'s study, Shang et al. in [20] suggest that, after determining the most suitable topological parameters τ and m of the attractor, a point $x^{(n_0)}$ is chosen and all the neighbouring points $x^{(n_i)} = [x^{(n_1)}, x^{(n_2)}, \ldots, x^{(n_j)}]$, called True Neighbours (TN), closer than the distance, r (for chosen arbitrarily between 0 and 1), are found.

A number of N-trajectories are utilized in finding the closest points on the predicted trajectory, $x(t_1 + m\tau)$, which is used as the starting vector during the computation of the LLE. This procedure is repeated for N-number of points along the orbits and an average quantity, $\mathcal{S}$, known as the Stretching Factor given by (35) is calculated. One has

$$\mathcal{S} = \frac{1}{N} \sum_{n_i=1}^{N} \left(\ln \frac{1}{u_{x^{(n_0)}}} \sum \left| x^{(n_0)} - x^{(n_i)} \right| \right), \qquad (35)$$

where $u_{x^{(n_0)}}$ is the number of neighbours around $x^{(n_0)}$.

Claim. Xue and Shi in [36] stated that if $20 \leq |u_{x^{(n_0)}}| \leq 30$, then a good approximation of the LLE can be obtained. A plot of $\mathcal{S}$ against the number of points N (or $t = N\Delta t$) yields a curve that has a linear inverse in one region which is followed by a plateau in another region. This plot takes the shape as illustrated in Figure 14.

The least squares approach gives a smooth line (fit) on Figure 14 and its slope gives an estimation of LLE, $\lambda_{\max}$.

Prediction. If $\lambda_{\max} \in \,]0, 1[$, then the system under analysis is not a Chaotic system but rather a stochastic one, and so we cannot make any predictions based on Chaos Theory.

If $0 < \lambda_{\max} < 1$, then it implies that there is Chaos in the system. For practical purposes, we compute the approximate period limit, $\Delta t_{\max}$ (often called *Lyapunov time*) for accurate prediction since it is a function of the LLE, $\lambda_{\max}$.

The Lyapunov time, $\Delta t_{\max}$, is given by

$$\Delta t_{\max} = \frac{1}{\lambda_{\max}}. \qquad (36)$$

If $\lambda_{\max} \to 0$ implies $\Delta t_{\max} \to \infty$, then long-term accurate predictions are possible. Initially, one starts with a vector, $X(t_1)$, followed by selecting k-closest trajectories (not points) on the system's attractor which is then followed by choice of k-closest points to $X(t_1)$ (one on each trajectory). It follows that we precisely know the dynamic evolution of the system after time, $\Delta t_{\max}$.

In the same way, if $\lambda_{\max} \to \infty$ implies $\Delta t_{\max} \to 0$, long-term accurate predictions are not possible, but rather short-term ones can be made. With the Lyapunov time, $\Delta t_{\max}$, we can precisely predict any observed quantity (say traffic flow) for this time [20].

Practically in traffic flow analysis, the one-dimensional traffic flow time series data is replaced with m-dimensional reconstructed data. The reconstructed time series data is then plotted, and this is followed by analysis of the previous observations which are neighbours to the preceding ones, and short-term predictions are finally made.

4. Conclusions

This study have shown how Chaos Theory can be used in the analysis of dynamical systems via a systematic review of the characteristic features of Chaotic system. In particular, it showed how Chaos Theory can be used for Motorised Traffic Flow Time Series Prediction in Urban Transport Network based on the the method of computation of the Largest Lyapunov Exponent, $\lambda_{\max}$, which is the best method so far for analysis and prediction of chaotic behaviours of a given complex system like traffic flow as reported by most researchers in literature. Using the Largest Lyapunov Exponent prediction method, it was shown how the Lyapunov time, $\Delta t_{\max}$, can be obtained which is the time interval for making accurate predictions of traffic flows.

In order to make a complete and robust prediction model for traffic flow, there is need to develop a computer based algorithm that will compute the time delay, embedding dimension, and Lyapunov time of a real time series from empirical traffic flow data. Thus, the validation aspect of the proposed approach and comparison with other known conventional models for traffic flow prediction especially in the area of prediction accuracy is still in progress and left for our future work so as to enable us have access to available traffic flow data sets. Moreover, there is need to come up with a concrete relationship (most preferably a mathematical equation) that links the Lyapunov time with traffic flow so as to aid in proper traffic predictions. The effect of noise on traffic flow data as well as determining the type of noise and magnitude is also an important area to look into in our future work. Thus, by effectively incorporating all these into

the present work, it is believed that the proposed approach will be useful in reducing the congestion problem on urban traffic networks.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.
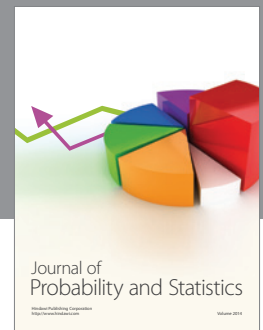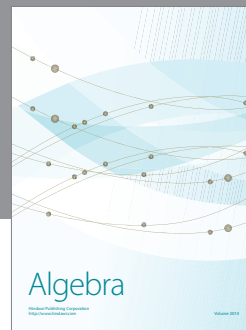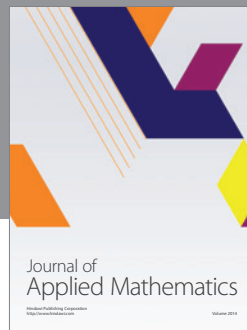
Acknowledgments

References

[1] M. Garavello and B. Piccolli, *Traffic Flow on Networks: Conservation Laws Model*, American Institute of Mathematical Sciences, Springfield, Mo, USA, 2006.

[2] U.S. Bureau of Transportation Statistics, National Transportation Statistics Electronic, January 2015, http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_statistics/index.html.

[3] N. H. Gartner and G. Improta, *Urban Traffic Networks: Dynamic Flow Modeling and Control*, vol. 2 of *Transportation Analysis*, Springer, Rome, Italy, 2nd edition, 2011.

[4] A. Y. Abul-Magd, "Modeling highway-traffic head-way distributions using superstatistics," *Physical Review E*, vol. 76, no. 5, Article ID 057101, 2007.

[5] G. Mwesige, H. Farah, U. Bagampadde, and H. N. Koutsopoulos, "A stochastic model for passing rate at passing zones on two-lane rural highways in Uganda," in *Proceedings of the 93rd Annual Meeting of the Transportation Research Board*, Washington, DC, USA, January 2014.

[6] A. Muhmoudabadi, "The assessment of applying chaos theory for daily traffic estimation," in *Proceedings of the International Conference on Industrial Engineering and Operations Management*, vol. 1, Tehran, Iran, January 2014.

[7] A. S. Nair, L. R. Jyh-Cham Liu, and G. Saurabh, "Non-linear analysis of traffic flow," in *Transportation Systems: Theory and Application of Advanced Technology, IFAC Symposium, Tianjin, PRC*, vol. 1, pp. 773–784, 1999.

[8] P. Shang, M. Wan, and S. Kama, "Fractal nature of highway traffic data," *Computers & Mathematics with Applications*, vol. 54, no. 1, pp. 107–116, 2007.

[9] B. De Moor and S. Maerivoet, *Traffic Flow Theory*, vol. 8999.40.-a, 02.50.-r,45.70, 2008.

[10] M. M. Q. Catriona and J. A. Casper, *Forecasting Traffic Ows in Road Networks: A Graphical Dynamic Model Approach*, Department of Mathematics and Statistics, The Open University, Milton Keynes, UK, 2008.

[11] M. E. Whitlock and C. M. Queen, "Modelling a traffic network with missing data," *Journal of Forecasting*, vol. 19, no. 7, pp. 561–574, 2000.

[12] C. M. Queen, B. J. Wright, and C. J. Albers, "Eliciting a directed acyclic graph for a multivariate time series of vehicle counts in a traffic network," *Australian & New Zealand Journal of Statistics*, vol. 49, no. 3, pp. 221–239, 2007.

[13] J. Dauwels, A. Aslam, M. T. Asif et al., "Predicting traffic speed in urban transportation subnetworks for multiple horizons," in *Proceedings of the 13th International Conference on Control Automation Robotics & Vision (ICARCV '14)*, pp. 547–552, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, December 2014.

[14] S. V. Kumar and L. Vanajakshi, "Short-term traffic flow prediction using seasonal," *European Transport Research Review*, vol. 7, no. 3, article 21, 2015.

[15] K. Kumara, M. Paridab, and V. K. Katiyar, "Short term traffic flow prediction for a non urban highway using artificial neural network," *Procedia—Social and Behavioral Sciences*, vol. 104, pp. 755–764, 2013, Proceedings of the 2nd Conference of Transportation Research Group of India (2nd CTRG).

[16] H. Dong, L. Jia, X. Sun, C. Li, and Y. Qin, "Road traffic flow prediction with a time-oriented ARIMA model," in *Proceedings of the 5th International Joint Conference on INC, IMS and IDC*, pp. 1649–1652, Seoul, Republic of Korea, August 2009.

[17] D. K. Parbat and P. B. Nagarnaik, "Artificial neural network modeling of road traffic noise descriptors," in *Proceedings of the 1st International Conference on Emerging Trends in Engineering and Technology (ICETET '08)*, pp. 1017–1021, Civil Engineering Department, G. H. Raisoni College of Engineering, Nagpur, India, July 2008.

[18] G. Fusco and C. Colombaroni, *An Integrated Method for Short-Term Prediction of Road Traffic Conditions for Intelligent Transportation Systems Applications*, Department of Civil, Constructional and Environmental Engineering, Sapienza University of Rome Via Eudossiana, Roma, Italy, 2013.

[19] G. A. Ali and C. F. Bakheit, "Comparative analysis and prediction of traffic accidents in Sudan using artificial neural networks and statistical method," in *Proceedings of the 30th Annual Southern African Transport Conference (SATC '11)*, Sudan University of Science and Technology, Pretoria, South Africa, July 2011.

[20] P. Shang, X. Li, and S. Kamae, "Chaotic analysis of traffic time series," *Chaos, Solitons & Fractals*, vol. 25, no. 1, pp. 121–128, 2005.

[21] Q. C. Diliman, "Traffic impact assessment for sustainable traffic management and transportation planning in Urban areas," in *Proceedings of the Eastern Asia Society for Transportation Studies*, vol. 5, pp. 2342–2351, 2005.

[22] F. L. Hall, *Traffic Stream Characteristics, Traffic Flow Theory*, US Federal Highway Administration, Washington, DC, USA, 1996.

[23] E. Appiah, http://www.myjoyonline.com/news/2015/may-19th/we-lose-over-125m-just-staying-in-traffic-experts-say.php.

[24] P. Cvitanovic, *Chaos, and What to Do About It?* vol. 1 of *Ebook*, 2015.

[25] L. D. Kiel and E. W. Elliott, *Chaos Theory in the Social Sciences: Foundations and Applications*, University of Michigan Press, Ann Arbor, Mich, USA, 1997.

[26] X. Zhang and D. F. Jarrett, "Chaos in a dynamic model of traffic flows in an origin-destination network," *Chaos*, vol. 8, no. 2, pp. 503–513, 1998.

[27] D. S. Dendrinos, "Traffic-flow dynamics: a search for chaos," *Chaos, Solitons & Fractals*, vol. 4, no. 4, pp. 605–617, 1994.

[28] E. Pepke, "Chaos Theory: What is the Difference Between Chaotic Behaviour and Random Behaviour?, Differentiating Chaotic and Random Processes," September 2004.

[29] G. Shewalo, N. Shinde, A. Shirode et al., *Report on Chaos Theory*, Sardar Patel Institute of Technology, Mumbai, India, 2012.

[30] E. N. Lorenz, *The Essence of Chaos*, vol. 1, University of Washington Press, Seattle, Wash, USA, 1st edition, 1993.

[31] S. S. Tiberio, *Stochastic and Deterministic Differential Equation Modeling: The Accuracy of Recovering Dynamic Model Parameters of Change*, University of Notre Dame, 2004.

[32] B. Dictionary, Non-Linearity, Definition of Non-linearity, 2015.

[33] M. Leok and B. Tiong, "Estimating the attractor dimension of the equatorial weather system," *Acta Physica Polonica A*, vol. 85, pp. 27–35, 1994.

[34] B. B. Mandelbrot, *The Fractal Geometry of Nature*, vol. 2 of *Cataloging in Publication Data*, Library of Congress, Washington, DC, USA, 1977.

[35] B. Goertzel, *Chaotic Logic: Language, Thought, and Reality from the Perspective of Complex Systems Science*, vol. 9, Springer, 1994.

[36] J. Xue and Z. Shi, "Short-time traffic flow prediction based on chaos time series theory," *Journal of Transportation Systems Engineering and Information Technology*, vol. 8, no. 5, pp. 68–72, 2008.

[37] S. J. Clement, M. A. P. Taylor, K. V. K. Rao, T. V. Mathew, and P. J. Gundaliya, *From Chaotic Road Traffic to Cooperative Opportunistic Percolation Using Cellular Automata*, NSW Transport and Population Data Centre, 2005.

[38] C. Frazier and K. M. Kockelman, "Chaos theory and transportation systems: an instructive example," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1897, no. 1, 2004.

[39] L. W. Lan, F.-Y. Lin, and A. Y. Kuo, "Testing and prediction of traffic flow dynamics with chaos," *Journal of the Eastern Asia Society for Transportation Studies*, vol. 5, 2003.

[40] J. D. Farmer, E. Ott, and J. A. Yorke, "The dimension of chaotic attractors," in *The Theory of Chaotic Attractors*, pp. 142–149, Springer, 2004.

[41] D. H. Campos, "Non-linear time series analysis of the EEG during sleep," *Revista de la Academia Colombiana de Ciencias Exactas*, vol. 20, no. 78, pp. 491–501, 1996.

[42] C. J. Cellucci, A. M. Albano, and P. E. Rapp, *Statistical Validation of Mutual Information Calculations: Comparisons of Alternative Numerical Algorithms*, Bureau of Medicine and Surgery of the Navy, Washington, DC, USA, 2004.

[43] H. D. I. Abarbanel, *Analysis of Observed Chaotic Data*, Institute for Nonlinear Science, Springer, New York, NY, USA, 1996.

[44] Ixellence, "False nearest neighbors (FNN)," May 2015.

[45] M. S. Khan and T. A. Siddiqui, "Estimation of fractal dimension of a noisy time series," *International Journal of Computer Applications*, vol. 45, no. 10, 2012.

[46] L. W. Lan, F.-Y. Lin, and A. Y. Kuo, "Identification for chaotic phenomena in short-term traffic flows: a parsimony procedure with surrogate data," *Journal of the Eastern Asia Society for Transportation Studies*, vol. 6, pp. 1518–1533, 2005.

Advances in
Operations Research

Advances in
Decision Sciences

Journal of
Applied Mathematics

Algebra

Journal of
Probability and Statistics

The Scientific
World Journal

International Journal of
Differential Equations

International Journal of
Combinatorics

Advances in
Mathematical Physics

Submit your manuscripts at
http://www.hindawi.com

Hindawi

Journal of
Complex Analysis

Journal of
Mathematics

Mathematical Problems
in Engineering

Abstract and
Applied Analysis

Discrete Dynamics in
Nature and Society

International
Journal of
Mathematics and
Mathematical
Sciences

Journal of
Discrete Mathematics

Journal of
Function Spaces

International Journal of
Stochastic Analysis

Journal of
Optimization