# Efficient Multistep Nonlinear Time Series Prediction involving Deterministic Chaos based Local Reconstruction methodologies and Multilayer Perceptron Neural Networks in Diode Resonator Circuits

D.A. KARRAS, M.P. HANIAS
Department of Automation Engineering
Sterea Hellas Institute of Technology,
GREECE
dakarras@teiste.gr, dimitrios.karras@gmail.com, dimitrios.karras@ieee.org,

*Abstract:* - A novel non-linear signal prediction method is presented using non linear signal analysis and deterministic chaos techniques in combination with an improved local reconstruction methodology and multilayer neural networks for a diode resonator chaotic circuit generated time series forecasting. Multisim is used to simulate the circuit and show the presence of chaos as well as to generate the time series data. The Time series analysis is performed by the method proposed by Grasberger and Procaccia, involving estimation of the correlation and minimum embedding dimension as well as of the corresponding Kolmogorov entropy. These parameters are used to construct the preprocessing step of a first stage of a one step / multistep predictor. This first stage involves, in the sequel a local reconstruction based approach. More specifically, it is suggested that by extracting a class of informative features coming from second order information, involving the topology of their neighbouring state vectors, from the state vectors of the local reconstruction approach then, significantly better results could be obtained with respect to chaotic time series reconstruction. In the second stage of the proposed method a multilayer neural network, trained with the conjugate gradient algorithm, is employed in order to provide the proper topology preserving error characteristics for the associated time series prediction. One of the novelties of the proposed two stage predictor lies on that the ANN involved could be employed as second order predictors, that is as error predictors of the non-linear signal analysis based forecasted values. This novel two stage chaotic signal forecasting technique is evaluated through an extensive experimental study.

*Keywords:-* time series forecasting, non-linear signal analysis, diode, chaos, time series, correlation dimension, prediction, error prediction, neural networks, local reconstruction, Backpropagation error

## 1 Introduction

Time series forecasting, or time series prediction, takes an existing series of data and forecasts the future data values. The goal is to observe or model the existing data series to enable future unknown data values to be forecasted accurately

A novel two-stage time series prediction method is presented in this paper and is applied to the prediction of a chaotic signal produced by a diode resonator chaotic circuit. This circuit, being quite simple, illustrates how chaos can be generated. We have selected Multisim [1] to simulate circuits since it provides an interface as close as possible to the real implementation environment. In addition, complete circuits implementation and oscilloscope graphical plots are all presented. While non-linear signal analysis methods have been quite extensively studied and applied in several systems presenting chaos, chaotic time series prediction for electronic circuits is a field not too deeply investigated so far. Chaos has already been recognized to be present in electronic circuits [2]-[5]. Some preliminary investigations on such time series prediction have been performed by the authors in [6]. The present paper aims at developing efficient predictors for such chaotic time series. To this end, the classical nonlinear signal analysis (i.e [7]-[8] ) has been involved as a first stage of the proposed predictor, while back-propagation neural networks have been employed in the second stage to enhance first stage results, being a second order predictor for the first

time in then relevant literature. An extensive experimental study shows that the proposed predictor is very favourably evaluated in terms of accuracy with the classical nonlinear signal analysis methodology.

## 2 The Non Autonomous Driven RL Diode Circuit

A non autonomous chaotic circuit referred to as the driven RL-diode circuit (*RLD*) [2- 4] shown in Fig 1
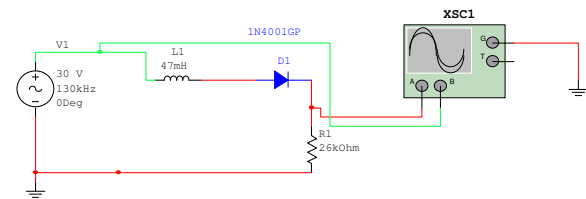


**Fig – 1** RL-Diode chaotic circuit

It consists of a series connection of an ac-voltage source, a linear resistor $R_1$, a linear inductor L1 and a diode $D_1$ type 1N4001GP, that is the only nonlinear circuit element. An important feature of this circuit is that the current i (or the voltage across the resistor R) can be chaotic although the input voltage V1 is non-chaotic. The usual procedure is to choose a parameter that strongly affects the system. We found that for $V_1$=30V RMS and input frequency f=130 KHz, inductance L1=47mH, the response is a chaotic one. The results of the Multisim simulation are shown in Fig. 2. The RL-diode was implemented and the voltage oscillations across the resistor $V_{R1}$ and its phase portrait V1 vs $V_{R1}$ are shown in Fig.2
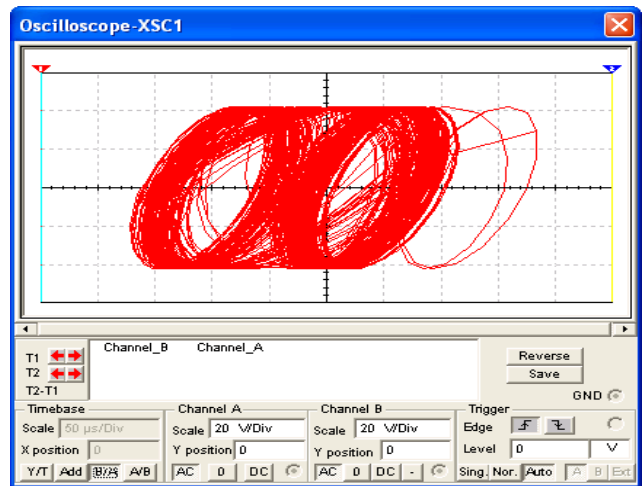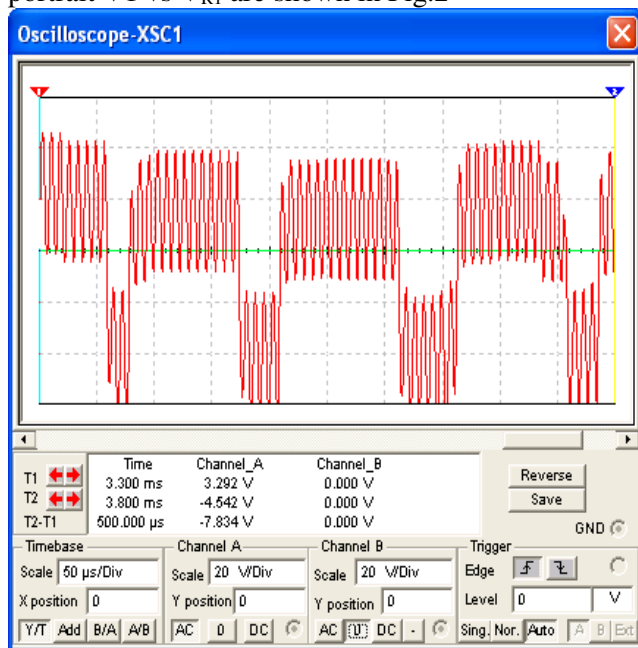




**Fig. 2** Time series $V_{R1}$ (t) (left column) Phase portrait of $V_1$versus VR1 (right column)

## 3. The Proposed Novel Prediction Methodology

### 3.1 First Stage Preprocessing: The Non Linear Signal Analysis Preprocessing of the State Vectors

Time series prediction takes an existing series of data

$$x_{t-n},\ldots,x_{t-2},x_{t-1},x_t \qquad (1)$$

and forecasts the future

$$x_{t+1},x_{t+2},\ldots \qquad (2)$$

data values. Taking into account this point of view we could interpret the data produced by the RLD circuit as a non-linear chaotic time series. The goal is to observe or model the existing data series to enable future unknown data values to be forecasted accurately. To evaluate the resulted time series, the method proposed by Grasberger and Procaccia [7,8] and successfully applied in similar cases [9-11] has been applied in order to define the first stage of the proposed predictor. According to Takens theory [12] the measured time series were used to reconstruct the original phase space. For this purpose, we calculated the correlation integral, for the simulated signal, defined by the following relation [13].

$$C_m(r)=\lim\frac{2}{(N)(N+1)}\sum_{i=1}^{N}\sum_{j=i+1}^{N}H\{r-(\sum_{k=1}^{m}\left[x_{i+k}-x_{j+k}\right]^2)^{\frac{1}{2}}\} \quad (3)$$

for $\lim r \rightarrow \infty$ , where

N…………………..is the number of points,

H…………………is the Heaviside function,

m is the embedding dimension

In the above equation N is the number of the experemental points here N=16337, $X_i$ is a point in

the m dimensional phase space with $X_i$ given by the following relation [12]

$$X_t=\{V_{R1}(t_i), V_{R1}(t_i+\tau), V_{R1}(t_i+2\tau)....V_{R1}(t_i+(m-1)\tau)\} \quad (4)$$

The vector
$X_t=\{V_{R1}(t_i), V_{R1}(t_i+\tau), V_{R1}(t_i+2\tau).....V_{R1}(t_i+(m-1)\tau)\}$,
represents a point to the m dimensional phase space in which the attractor is embedded each time, where $\tau$ is the time delay $\tau=i\Delta t$ determined by the first minimum of the time delayed mutual information ,$I(\tau)$ [13-16]. In our case, because of sample rate $\Delta t=4.8\times10^{-7}$ s, the mutual information function exhibits a local minimum at $\tau=6$ time steps as shown at Fig -3.
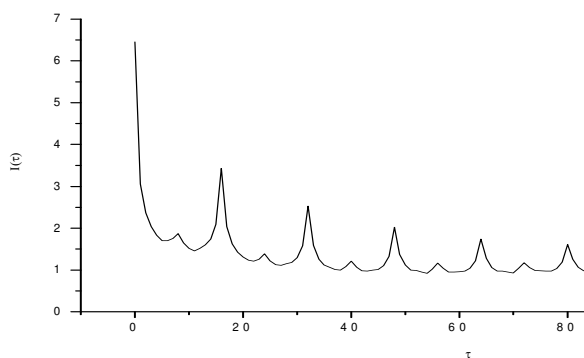


**Fig 3** Average Mutulal Information vs time delay $\tau$

We used this value for the reconstruction of phase space. With (3) dividing this space into hypercubes with a linear dimension r we count all points with mutual distance less than r. It has been proven [7-8] that if our attractor is a strange one, the correlation integral is propotional to $r^\nu$ where $\nu$ is a measure of the dimension of the attractor, called the correlation dimension. The correlation integral C(r) has been numerically calculated as a function of r from formula (3), for embedding dimensions m=1..10. In Fig 4 (upper insert) the slopes $\nu$ of the lower linear parts of these double logarithmic curves give information characterizing the attractor
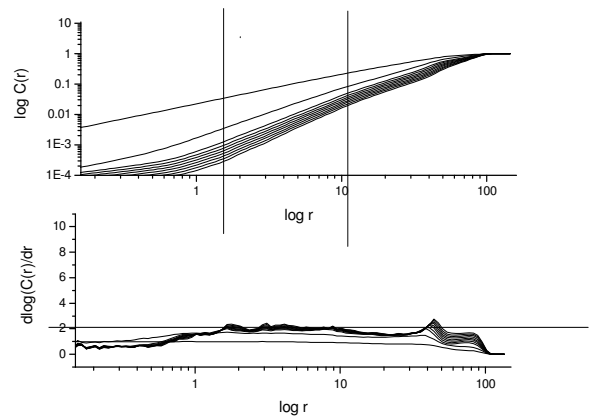


**Fig 4** The correlation intergral C(r) vs logr, for different embedding dimensions m (upper insert). The corresponding slopes and the scaling region (lower insert).

In fig 4 (lower insert) the corresponding average slopes $\nu$ are given as a function of the embedding dimension m It is obvious from these curves that $\nu$ tends to saturate, for higher m's , at non integer value $\nu=2.11$ with this value of $\nu$ the minimum embedding dimension could be $m_{min}=3$ [13]. So the minimum embedding dimension of the attractor for one to one embedding is 3.

In order to get more precise measurements of the strength of the chaos present in the oscillations we have introduced the Kolmogorov entropy. According to [13] the method described above also gives an estimate of the Kolmogorov entropy, i.e. the correlation integral C (r) scales with the embedding dimension m according to the following relation

$$C(r) \sim e^{-m\tau K_2} \quad (5)$$

Where $K_2$ is a lower bound to the Kolomogorov entropy. From the plateau of fig 5 we estimate $K_2=0.11$ bit/s
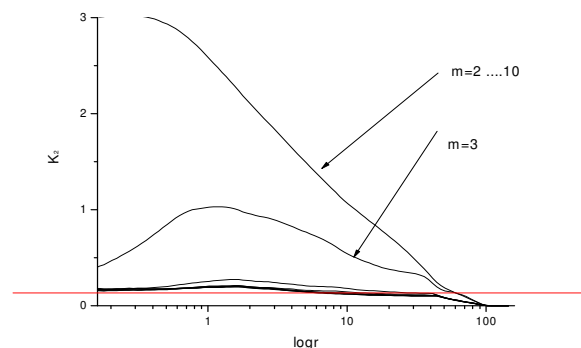


**Fig 5** The Kolmogorov entropy vs log r for different embedding dimensions

## 3.2 First stage Main Processing: The proposed Enhanced Local Reconstruction Methodology for Time Series Reconstruction and Forecasting

In this paper we present a novel methodology to attempt predictions in chaotic time series based on the local modeling approach, known and mainly employed in chaotic time series forecasting [2], as discussed above. Several important modifications in the stages of the original local modeling technique are suggested so as to improve its efficiency, involving a second order approach.

First, concerning the stage of determining the components of the neighboring state vectors of the local modeling algorithm, we augment them by incorporating second order information embedded in these state vectors but not currently used. In the sequel, we use this second order information for forecasting prediction errors aiming at improving our predictions.

More specifically, let us consider a time series $x(t)$ with state vectors $X(t) = (x\_1(t), x\_2(t), x\_3(t),... x\_n(t)) = (x(t), x(t-l), x(t-2l),... x(t-(n-1)l))$, where $l$ is the delay time. Also, it is assumed that the dimensionality $n$ is quite large, which is true especially for a chaotic time series where the minimal requirement is $n >= D$, with $D$ its attractor dimension, discussed in the previous section. The next step is to assume a functional relationship between the current state vector $X(t)$ and the future one $X(t+T)$, $X(t+T) = f\_T(X(t))$. The problem is to find a predictor $F\_T$ approximating $f\_T$. The local approximation methodology is one of the most effective approaches for solving it by involving nearby only states to make predictions. To predict $x(t+T)$ we first impose a metric on the state space, denoted by $\| \|$, and find the $k$ nearest neighbors of $X(t)$, i.e., the $k$ states $X(t')$ with $t' < t$ that minimize $\|X(t)-X(t')\|$. The most widely accepted such metric is the known Euclidean distance. This is the first stage of the algorithm. Although this original approach seems natural, however, the state vectors include second order information that is not exploited so far.

To be more precise, each state vector $X(t)$ along with its $k$ nearest neighbors $X(t')$ form a group with a representative autocorrelation matrix which expresses second order characteristics of these state vectors. It is well known that such an autocorrelation matrix is formed by involving the following known relationship.

$$R(X(t), X(t^{(1)}), X(t^{(2)}),...... X(t^{(k)})) = (1/(k+1)) (\sum X(t^{(J)}) X^{Transpose}(t^{(J)})), \qquad (6)$$

where $J=0, 1, 2, \ldots k$, that is, the summation includes the current state vector $X(t^{(0)}) = X(t)$, as well as its $k$ nearest neighbors. Finally, $X^{Transpose}(t^{(J)})$ is the transpose matrix of the $J$ neighbor state vector $X(t^{(J)})$ of the current one.

Therefore, if for each state vector $X(t)$ of the time series its corresponding autocorrelation matrix $R(X(t))$ is derived from its associated $k$ nearest neighbors $X(t^{(J)})$ then, we could have a quantitative statistical description of the second order properties of this state vector, that is of the relations between its components. By applying Principal Component Analysis to these autocorrelation matrices their eigenvectors $e_r$, for $r = 1,2,\ldots n$ (dimension of current state vector $X(t)$), can be extracted. Since, as it is well known from signal processing theory, the eigenvectors corresponding to the largest eigenvalues might be associated with the time series inherent noise, only the $p$ eigenvectors corresponding to the average in magnitude eigenvalues are kept in the following steps of the suggested procedure. These $p$ eigenvectors are used to provide additional informative components to their associated state vector $X(t)$ and thus, form the new augmented state vector $X'(t)$.

More specifically, the informative features these eigenvectors provide to their corresponding state vector $X(t)$ are simply the cosines of their angles with the vector $X(t)$. Therefore, $m$ additional such components are incorporated into each state vector $X(t)$, conveying a kind of second order information about $X(t)$ neighborhood.

The above discussion leads to the definition of the augmented state vectors $X'(t)$ and the second order vectors $X''(t)$ as follows,

$$X'(t) = [ X(t), <X(t), e_1>/(\|X(t)\|*\| e_1\|),\ldots <X(t), e_r>/(\|X(t)\|*\| e_r\|), \ldots <X(t), e_p>/(\|X(t)\|*\| e_p\|)]$$

$$X''(t) = [ <X(t), e_1>/(\|X(t)\|*\| e_1\|),\ldots <X(t), e_r>/(\|X(t)\|*\| e_r\|), \ldots <X(t), e_p>/(\|X(t)\|*\| e_p\|)] \qquad (7)$$

This second vector $X''(t)$, defined completely above, will serve as input to the Neural Networks of the second stage in the next section, as second order error predictors.

In the previous definitions, r=1,2,…p are the above reported p eigenvectors having average in magnitude eigenvalues. The parameter p is a user defined parameter and depends on the dimension n of the current state vector X(t). In the experiments involved in the next section we always consider p=n/3.

The above specified second order feature extraction based on the local construction methodology will be applied in the next section for the chaotic time series forecasting as a second order error prediction approach.

### 3.3 Second Stage: Back-Propagation ANNs as Second Order Error Predictors

Based on the above stage, the proposed novel algorithm to enhance non-linear signal analysis prediction is as follows:

1. To predict point $V_{i+1}$, we determine the last known state of the system as represented by vector X = [$V_i$, $V_{i-\tau}$, $V_{i-2\tau}$, $V_{i-(m-1)\tau}$], where m is the embedding dimension and $\tau$ is the time delay.

2. With optimum values of delay time and embedding dimension m we then search the time series to find k similar states that have occurred in the past, where "similarity" is determined by evaluating the distance between vector X and its neighbour vector X' in the m-dimensional state space. So k close states (usually nearest neighbours of X) of the system that have occurred in the past are found, by computing their distances from X.

3. We used a fixed size of nearest neighbours K (calculated for optimizing prediction performance in the training phase). if a state X'= [$V'_i$, $V'_{i-\tau}$, $V'_{i-2\tau d}$, $V'_{i-(m-1)\tau}$] in the neighbourhood of X resulted in the observation $V'_{i+1}$ in the past, then the point $V_{i+1}$ which we want to predict must be somewhere near $V'_{i+1}$. This is the main concept of nonlinear signal analysis of first order approximation.

4. It is reasonable to calculate $V_{i+1} = (\Sigma q_k V'_k)/ \Sigma q_k$ , where $q_k$ the distance between current state X and neighboring state $X'_k$, whereas $V'_k$ the corresponding prediction from $X'_k$ vector (from the training set). The above sum is considered for all neighbors of X.

5. Our proposition to enhance prediction results is to write down $V_{i+1} = (\Sigma q_k V'_k)/ \Sigma q_k$ + error_$V_{i+1}$, where $(\Sigma q_k V'_k)/ \Sigma q_k$ is the first order prediction and error_$V_{i+1}$, is the prediction error to be minimized provided it is calculated properly. Therefore, it is a second order approximation proposal to predict such an error. This error_$V_{i+1}$, could be calculated through a suitable neural network as an error predictor [18]. This is precisely the main concept of the proposed novel methodology, which is based on the remarks of the previous section involving a second order feature extraction methodology using the local construction approach.

6. Suppose err_k the corresponding prediction error measured through the above procedure for each neighboring state $X'_k$ of given current state X above (out of the K neighbours of X). This err_k is known through the training set, since for each $X'_k$ in the training set we can calculate its corresponding K neighbours from the training set, and then, estimate, using step 5 above, the associated err_k. It could be reasonably considered, then, that this err_k prediction error has a strong correlation with the second order information outlined in the equations (7) of the previous section. This approach improves considerations and results of a preliminary attempt of the same authors for the same problem [19,20]. In that work, another approach was considered for the second order predictor. Namely, for each k state vector of the training set we had constructed all K such err_k. More specifically, each w of the K neighbours of the k state vector $X_k$ is associated with an error_k , let's say w_error_k. If we consider these K values for the relevant w_error_k error values corresponding to the $V_{i+1}$ prediction of step 4 above, then, we form an input vector of K inputs. In the sequel, in [19,20] we fed these K values as inputs to a back-propagation neural network of K-L1-L2-1 architecture. It was reasonable to say that these K relevant w_error_k error values where somehow associated with the sought prediction error error_k. In the present herein approach we formulate an input vector of the p second order characteristics of the previous section. These features could be directly considered second order since they reveal special topology characteristics of the K neighbour-

ing state vectors corresponding to the state vector $X_k$. On the other hand, the previous works [19,20] based K relevant w_error_k error values were indirectly associated with the second order characteristics of the considered state vectors. Such a more direct association might be beneficial for the problem at hand. Therefore, we herein consider a K*p-L3-L4-1 Backprogation error ANN architecture. That is, for each k state vector $X_k$ we find its K neighbours, state vectors $X'_{kw}$, where w=1..K. For each $X'_{kw}$ its associated p mid value eigenvector characteristics, of section 3.2, are extracted. In order to analyse the space of each $X'_{kw}$ we have considered only its two closest nearest neighbouring state vectors, where K>2. Therefore, a set of p*K values have been extracted following this procedure. This K*p-L3-L4-1 Backprogation error ANN network has been trained with the conjugate gradient algorithm, due to the large training set, since it is known to be the best algorithm for large data sets and ANN architectures [18], should be able to predict state's X error error_$V_{i+1}$.

7. The training set needed for step 6 is constructed for each state X of the training set by estimating all corresponding p*K second order features, outlined in section 3.2 and in step 6 above, of its K neighbours and its associated error_$V_{i+1}$, which of course serves as the desired output of the corresponding p*K values input pattern

# 4. Experimental Study

We have used a simulated time series from RLD circuit with $V_1$=30V RMS and input frequency f=130 KHz and we predict the voltage V across the resistor. We use locally linear models to predict the one step and the multistep procedures. That is, instead of fitting one complex model with many coefficients to the entire data set, we fit many simple models (low order polynomials) to small portion of the data set depending on the geometry of the local neighborhood of the dynamical system [17]. The general procedure is the following: To predict point $V_{i+1}$, we determine the last known state of the system as represented by vector **X** = [$V_i$, $V_{i-\tau}$, $V_{i-2\tau}$, $V_{i-(m-1)\tau}$], where m is the embedding dimension and $\tau$ is the time delay.
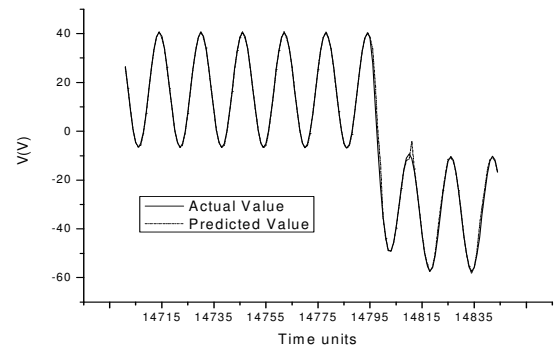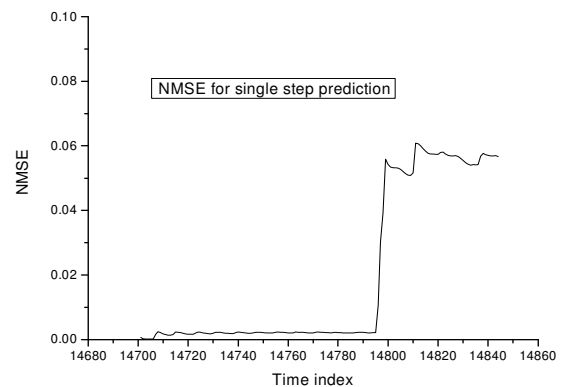


**Fig 6** One step prediction



**Fig 7** Mean squared error of our predictor normalized by the mean squared error of the random walk predictor for one step prediction

So we use as a delay time the value of $\tau$=6 as before. From previous analysis the correlation dimension for RLD circuit is found v=2.11. With optimum values of delay time and embedding dimension m=3 we then search the time series to find k similar states that have occurred in the past, where "similarity" is determined by evaluating the distance between vector X and its neighbour vector X' in the m-dimensional state space. So k close states (usually nearest neighbours of **X**) of the system that have occurred in the past are found, by computing their distances from **X** as explained in section 3 above

The idea is to fit a map which extrapolates **X** and its k nearest neighbours to determine the next value. If the observable signal was generated by some deterministic map M($V_i$,$V_{i-\tau}$, $V_{i-2\tau}$, $V_{(i-(m-1)\tau)}$) = $V_{i+\tau}$, that map can be recovered (reconstructed) from the data by looking at its behaviour in the neighbourhood of X. Using this map, an approximate value of $V_{i+1}$ can be obtained. We used a fixed size of nearest neighbours k=36. Also, concerning the second order procedure outlined in

section 3, p=3. Now we can use this map to predict $V_{i+1}$ In other words, we make an assumption that M is fairly smooth around X, and so if a state X'= [$V'_i$, $V'_{i-\tau}$, $V'_{i-2\tau d}$, $V'_{i-(m-1)\tau}$] in the neighbourhood of X resulted in the observation $V'_{i+1}$ in the past, then the point $V_{i+1}$ which we want to predict must be somewhere near $V'_{i+1}$. [17]. We have employed both the one step and multistep ahead prediction methods. In the one step ahead prediction, after each step in the future is predicted, the actual value is utilized for the next one – step prediction. In contrast, the multistep prediction is based only on the initial k states.

The calculated performance is otherwise known as the Normalized Mean Squared Error (NMSE) is calculated by (5-1),

$$NMSE = MAX\left(\frac{\sum_{i=1}^{NP}\left(\tilde{V}_i - V_i\right)^2}{\sum_{i=1}^{NP}\left(\overline{V}_i - V_i\right)^2}, \frac{\sum_{i=1}^{NP}\left(\tilde{V}_i - V_i\right)^2}{\sum_{i=1}^{NP}\left(V_{i-1} - V_i\right)^2}\right) \quad (5-1),$$

where $\tilde{V}_i$ is the predicted value, $V_i$, the actual value, $\overline{V}$ is the average actual value, and NP is the range of values in the prediction interval.

From (5-1), it can be seen that NMSE is the mean squared error of our predictor normalized by the mean squared error a random walk predictor. By definition, the minimum value of NMSE is 0. At that value, there is the exact match between the actual and predicted values. The higher NMSE, the worse is our prediction as compared to the trivial predictors. If NMSE is equal to 1, our prediction is as good as the prediction by the trivial predictor. If NMSE is greater than 1, our prediction worsens. With values of $\tau=6$, m=3 we achieved the minimum NMSE.

The second stage back-propagation ANN trained with conjugate gradient technique is of a 3*36-64-64-1 = 108-64-64-1 architecture [18].

We used 14700 data points and predicted the evolution for 889 succeeding dimensionless time steps. The results are shown at fig 6 where the one step ahead predicted values are coming from prediction out-of-sample set, where we pretend that we know the data only up until this point, and we try to predict from there, while the one step ahead predicted values are coming from prediction out-of-sample set. The NMSE is shown at fig 7 for the one step prediction.

We use the same procedure as before but with multistep ahead predictions. The results are shown at Fig - 8 The NMSE is shown at Fig - 9 for the multi step prediction
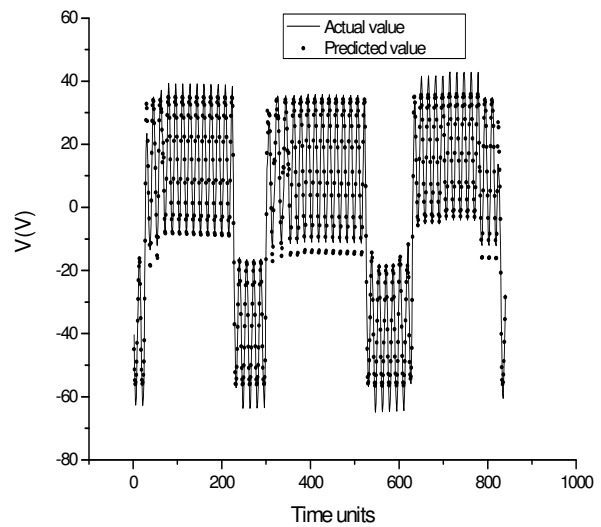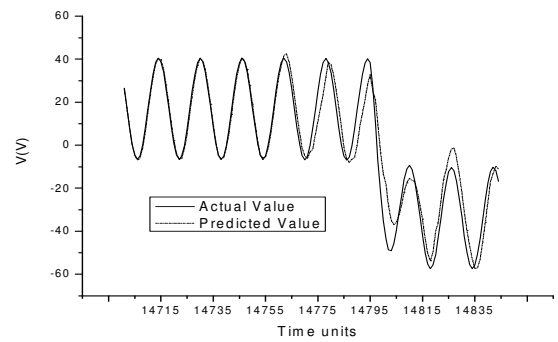




**Fig 8** Multistep prediction, Actual and predicted time series for 10 time steps ahead for the total set of points and the unknown time series (lower insert, in detail)
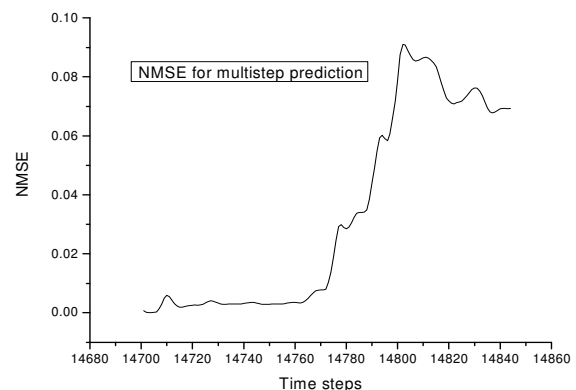


**Fig 9** Mean squared error of our predictor normalized by the mean squared error of the random walk predictor for multistep prediction

In comparison, when a first stage only predictor is used without the proposed neural network of stage 2, on average, for the 889 unknown data points we have achieved *6.3% worse performance in the one-step prediction for the NMSE and 5.8% worse performance in the multistep prediction experiments*. In the same experiments, performed in the previous approach of the same authors [19], there has been achieved more than *2% improvement concerning the 1-step prediction and 2.5% improvement concerning the multistep prediction performance*. Therefore, this new proposed methodology is worth evaluating it further in much larger scale experiments.

## 5. Conclusions and Future Trends

We have proposed a novel two-stage chaotic time series prediction scheme based on nonlinear signal analysis methods an improved second order based local reconstruction methodology and a novel error prediction back propagation ANN trained with the conjugate gradient algorithm. Applying the methods of non linear analysis in the time series produced by the chaotic simple RLD circuit we found that the strange attractor that governs the phenomenon is a Lorenz type attractor with a correlation dimension v=2.11 who is stretching and folding in a three dimensions phase space. This is also evident from the one step ahead and multistep successful predictions with the use of the correspondence strange attractor invariants as input parameters, and the efficient ANN model introduced in the second stage of the proposed forecasting system.

We believe that for a detailed understanding of chaos in the *RLD* circuits these results must be combined with the reverse-recovery effect and all of its nonlinearities. The proposed two stage local reconstruction based improved prediction methodology might be applied successfully in other chaotic and non-chaotic time series too, since it is quite general. This remains, also, a future target of the authors.

*References:*

[1] Sajid Iqbal ; Xizhe Zang ; Yanhe Zhu ; Xinyu Liu ; Jie Zhao "Introducing Undergraduate Electrical Engineering Students to Chaotic Dynamics: Computer Simulations with Logistic Map and Buck Converter", Modelling Symposium (AMS), 2014 8th Asia, Date of Conference: 23-25 Sept. 2014,

[2] T. Matsumato, L. Chua and S. Tanaka, "Simplest Chaotic Nonautonomous Circuit," *Phys. Rev A,* Vol. 30, pp.1155-1157, August 1984.

[3] A. Azzouz and M. Hasler, "Orbits of the R-L-Diode Circuit,*" IEEE Transaction on Circuits and Systems,*Vol.37, pp.1330-1339, November 1990.

[4] Riva, M., Belloni, F., Giustina, D. Della, "Teaching non linear systems: The educational approach of power electronics", Proceedings of the 2011 14th European Conference on Power Electronics and Applications pp. 1 - 7

[5] Renato Mariz de Moraes and Steven M. Anlage, "Unified model and reverse recovery nonlinearities of the driven diode resonator" Phys. Rev E. 68, 026201 (2003)

[6] M. P.Hanias, G. Giannaris , A. Spyridakis, A. Rigas, " Time series Analysis in chaotic diode resonator circuit", Chaos Solitons & fractals , v.27, #2, 569-573, 2006

[7] P. Grassberger and I. Procaccia, "Characterization of strange attractors" Phys. Rev Lett., 50, 346-349 (1983)

[8] P. Grassberger and I. Procaccia, "Measuring the strangeness of strange. attractors", Physica D 9,189 (1983)

[9] M.P. Hanias, J.A. Kalomiros,Ch. Karakotsou,A.N. Anagnostopoulos, J. Spyridelis, "Quasi-Periodic and Chaotic Self - Excited Voltage Oscillations in TlInTe$_2$ ", Phys. Rev B. 49,16994 (1994).

[10] Mozdy, E., Newell, T.C., Alsing, P.M., Kovanis, V., Gavrielides, A. (1995), Synchronization and control in a unidirectionally coupled array of chaotic diode resonators.", Physical Review E., 51(6), 5371-5376

[11] Abarbanel, HDI. "Analysis of Observed Chaotic Data", Springer-Verlag, New York 1996.

[12] F. Takens, Lecture Notes in Mathematics 898 (1981)

[13] H. Kantz and T.Schreiber , "Nonlinear Time Series Analysis", Cambridge University Press, Cambridge (1997)

[14] T. Aasen, D. Kugiumtzis, S.H.G. Nordahl, "Procedure for Estimating the Correlation Dimension of Optokinetic Nystagmus Signals", Computers and Biomedical Research 30, 95-116 (1997)

[15] Fraser AM, Swinney HL, "Independent coordinates for strange attractors from mutual in-

formation", Phys Rev A 33, pp. 1134-1140 (1986)

[16]   Fraser AM, "Information and entropy in strange attractors" IEEE transaction of information Theory 35, 245 (1989)

[17]   Abarbanel Henry, "Analysis of observed chaotic data", Springer Science & Business Media, 2012.

[18]   Haykin, Simon S., et al. Neural networks and learning machines. Vol. 3. Upper Saddle River, NJ, USA:: Pearson, 2009.

[19]   Hanias, M.P., Karras, D.A. , "On efficient multistep non-linear time series prediction in chaotic diode resonator circuits by optimizing the combination of non-linear time series analysis and neural networks", Engineering Applications of Artificial Intelligence, Volume 22, Issue 1, February 2009, Pages 32-39

[20]   Hanias, M. P., and D. A. Karras. "Efficient non linear time series prediction using non-linear signal analysis and neural networks in chaotic diode resonator circuits." Industrial Conference on Data Mining. Springer Berlin Heidelberg, 2007.