Identification of Non-Linear Structures using Recurrent Neural Networks

Kirkegaard, Poul Henning; Nielsen, Søren R. K.; Hansen, H. I.

Publication date:
1994

Document Version
Early version, also known as pre-print

INSTITUTTET FOR BYGNINGSTEKNIK
DEPT. OF BUILDING TECHNOLOGY AND STRUCTURAL ENGINEERING
AALBORG UNIVERSITET • AUC • AALBORG • DANMARK

FRACTURE & DYNAMICS
PAPER NO. 55

To be presented at the 13th International Modal Analysis Conference, Nashville, Tennessee, February 13-16, 1995

P. H. KIRKEGAARD , S. R. K. NIELSEN & H. I. HANSEN
IDENTIFICATION OF NON-LINEAR STRUCTURES USING RECURRENT NEURAL NETWORKS
OCTOBER 1994 ISSN 0902-7513 R9432

The FRACTURE AND DYNAMICS papers are issued for early dissemination of research results from the Structural Fracture and Dynamics Group at the Department of Building Technology and Structural Engineering, University of Aalborg. These papers are generally submitted to scientific meetings, conferences or journals and should therefore not be widely distributed. Whenever possible reference should be given to the final publications (proceedings, journals, etc.) and not to the Fracture and Dynamics papers.

FRACTURE AND DYNAMICS PAPERS

PAPER NO. 27: Jakob Laigaard Jensen, Poul Henning Kirkegaard & Rune Brincker: *Model and Wave Load Identification by ARMA Calibration.* ISSN 0902-7513 R9035.

PAPER NO. 28: Rune Brincker, Steen Krenk & Jakob Laigaard Jensen: *Estimation of Correlation Functions by the Random Decrement Technique.* ISSN 0902-7513 R9041.

PAPER NO. 29: Poul Henning Kirkegaard, John D. Sørensen & Rune Brincker: *Optimal Design of Measurement Programs for the Parameter Identification of Dynamic Systems.* ISSN 0902-7513 R9103.

PAPER NO. 30: L. Gansted & N. B. Sørensen: *Introduction to Fatigue and Fracture Mechanics.* ISSN 0902-7513 R9104.

PAPER NO. 31: R. Brincker, A. Rytter & S. Krenk: *Non-Parametric Estimation of Correlation Functions.* ISSN 0902-7513 R9120.

PAPER NO. 32: R. Brincker, P. H. Kirkegaard & A. Rytter: *Identification of System Parameters by the Random Decrement Technique.* ISSN 0902-7513 R9121.

PAPER NO. 33: A. Rytter, R. Brincker & L. Pilegaard Hansen: *Detection of Fatigue Damage in a Steel Member.* ISSN 0902-7513 R9138.

PAPER NO. 34: J. P. Ulfkjær, S. Krenk & R. Brincker: *Analytical Model for Fictitious Crack Propagation in Concrete Beams.* ISSN 0902-7513 R9206.

PAPER NO. 35: J. Lyngbye: *Applications of Digital Image Analysis in Experimental Mechanics.* Ph.D.-Thesis. ISSN 0902-7513 R9227.

PAPER NO. 36: J. P. Ulfkjær & R. Brincker: *Indirect Determination of the $\sigma - w$ Relation of HSC Through Three-Point Bending.* ISSN 0902-7513 R9229.

PAPER NO. 37: A. Rytter, R. Brincker & P. H. Kirkegaard: *An Experimental Study of the Modal Parameters of a Damaged Cantilever.* ISSN 0902-7513 R9230.

PAPER NO. 38: P. H. Kirkegaard: *Cost Optimal System Identification Experiment Design.* ISSN 0902-7513 R9237.

PAPER NO. 39: P. H. Kirkegaard: *Optimal Selection of the Sampling Interval for Estimation of Modal Parameters by an ARMA-Model.* ISSN 0902-7513 R9238.

PAPER NO. 40: P. H. Kirkegaard & R. Brincker: *On the Optimal Location of Sensors for Parametric Identification of Linear Structural Systems.* ISSN 0902-7513 R9239.

PAPER NO. 41: P. H. Kirkegaard & A. Rytter: *Use of a Neural Network for Damage Detection and Location in a Steel Member.* ISSN 0902-7513 R9245

PAPER NO. 42: L. Gansted: *Analysis and Description of High-Cycle Stochastic Fatigue in Steel.* Ph.D.-Thesis. ISSN 0902-7513 R9135.

FRACTURE AND DYNAMICS PAPERS

PAPER NO. 43: M. Krawczuk: *A New Finite Element for Static and Dynamic Analysis of Cracked Composite Beams.* ISSN 0902-7513 R9305.

PAPER NO. 44: A. Rytter: *Vibrational Based Inspection of Civil Engineering Structures.* Ph.D.-Thesis. ISSN 0902-7513 R9314.

PAPER NO. 45: P. H. Kirkegaard & A. Rytter: *An Experimental Study of the Modal Parameters of a Damaged Steel Mast.* ISSN 0902-7513 R9320.

PAPER NO. 46: P. H. Kirkegaard & A. Rytter: *An Experimental Study of a Steel Lattice Mast under Natural Excitation.* ISSN 0902-7513 R9326.

PAPER NO. 47: P. H. Kirkegaard & A. Rytter: *Use of Neural Networks for Damage Assessment in a Steel Mast.* ISSN 0902-7513 R9340.

PAPER NO. 48: R. Brincker, M. Demosthenous & G. C. Manos: *Estimation of the Coefficient of Restitution of Rocking Systems by the Random Decrement Technique.* ISSN 0902-7513 R9341.

PAPER NO. 49: L. Gansted: *Fatigue of Steel: Constant-Amplitude Load on CCT-Specimens.* ISSN 0902-7513 R9344.

PAPER NO. 50: P. H. Kirkegaard & A. Rytter: *Vibration Based Damage Assessment of a Cantilever using a Neural Network.* ISSN 0902-7513 R9345.

PAPER NO. 51: J. P. Ulfkjær, O. Hededal, I. B. Kroon & R. Brincker: *Simple Application of Fictitious Crack Model in Reinforced Concrete Beams.* ISSN 0902-7513 R9349.

PAPER NO. 52: J. P. Ulfkjær, O. Hededal, I. B. Kroon & R. Brincker: *Simple Application of Fictitious Crack Model in Reinforced Concrete Beams. Analysis and Experiments.* ISSN 0902-7513 R9350.

PAPER NO. 53: P. H. Kirkegaard & A. Rytter: *Vibration Based Damage Assessment of Civil Engineering Structures using Neural Networks.* ISSN 0902-7513 R9408.

PAPER NO. 54: L. Gansted, R. Brincker & L. Pilegaard Hansen: *The Fracture Mechanical Markov Chain Fatigue Model Compared with Empirical Data.* ISSN 0902-7513 R9431.

PAPER NO. 55: P. H. Kirkegaard, S. R. K. Nielsen & H. I. Hansen: *Identification of Non-Linear Structures using Recurrent Neural Networks.* ISSN 0902-7513 R9432.

PAPER NO. 56: R. Brincker, P. H. Kirkegaard, P. Andersen & M. E. Martinez: *Damage Detection in an Offshore Structure.* ISSN 0902-7513 R9434.

PAPER NO. 57: P. H. Kirkegaard, S. R. K. Nielsen & H. I. Hansen: *Structural Identification by Extended Kalman Filtering and a Recurrent Neural Network.* ISSN 0902-7513 R9435.

Department of Building Technology and Structural Engineering
Aalborg University, Sohngaardsholmsvej 57, DK 9000 Aalborg
Telephone: +45 98 15 85 22 Telefax: +45 98 14 82 43

INSTITUTTET FOR BYGNINGSTEKNIK
DEPT. OF BUILDING TECHNOLOGY AND STRUCTURAL ENGINEERING
AALBORG UNIVERSITET • AUC • AALBORG • DANMARK

FRACTURE & DYNAMICS
PAPER NO. 55

To be presented at the 13th International Modal Analysis Conference, Nashville, Tennessee, February 13-16, 1995

P. H. KIRKEGAARD , S. R. K. NIELSEN & H. I. HANSEN
IDENTIFICATION OF NON-LINEAR STRUCTURES USING RECURRENT NEURAL NETWORKS
OCTOBER 1994 ISSN 0902-7513 R9432

IDENTIFICATION OF NON-LINEAR STRUCTURES USING RECURRENT NEURAL NETWORKS

P. H. Kirkegaard, S.R.K. Nielsen & H.I. Hansen
Department of Building Technology and Structural Engineering
Aalborg University
Sohngaardsholmsvej 57, 9000 Aalborg, Denmark

ABSTRACT

Two different partially recurrent neural networks structured as Multi Layer Perceptrons (MLP) are investigated for time domain identification of a non-linear structure. The one partially recurrent neural network has feedback of a displacement component from the output layer to a tapped-delay-line (TDL) input layer. The other recurrent neural network based on the Innovation State Space model (ISSM) has feedback of the state space vector from the output layer to the input layer. The recurrent neural network approaches are validated with respect to prediction and simulation of a non-linear process by application to simulated data from a viscous damped oscillator with hysteresis of the curve-linear type described by the Bouc-Wen model. The oscillator is subjected to amplitude modulated Gaussian white noise filtered through a Kanai-Tajimi filter. It is found that the two neural network models can act as actual system identifiers, predictors and simulators. The recurrent neural network with a TDL seems to be a better simulator than the ISSM network.

NOMENCLATURE

x_{lj} : Output of the jth node in the lth layer
θ_{lj} : A threshold of the jth neuron in the lth layer
N_0 : Number of network input
N_k : Number of network output
$f(\cdot)$: Activation function
J : Error function
y_j : Desired output
$\hat{y}_j$: Actual output
$w_{lj,i}$: Connection weight from node i to node j
$\mathbf{w}$: Vector including weights and thresholds
$\hat{X}(t)$: Estimated scalar response
$F(t)$: Scalar excitation
$G(\cdot)$: Nonlinear function
δ_{ij} : Dirac delta
$\mathbf{E}(t)$: Prediction error vector
$\mathbf{H}$: Observation matrix
$\mathbf{I}$: Unit matrix
ζ : Structural damping ratio
ω_0 : Circular eigenfrequency
$Z(t)$: Hysteretic component
$g(\cdot)$: Restoring term
α : Elastic fraction of restoring force
γ : Hysteretic parameter in Bouc-Wen model
β : Hysteretic parameter in Bouc-Wen model
n : Hysteretic parameter in Bouc-Wen model
z_0 : Yield displacement
ζ_s : Damping ratio of subsoil layer model
ω_s : Circular eigenfrequency of subsoil layer model
$\mathbf{X}(t)$: State vector
$W(t)$: Unit white noise
$r(t)$: Modulation function
r_0 : Strength of excitation
Δt : Time interval
$E[\cdot]$: Expectation operator
g : Acceleration of Gravity

1. INTRODUCTION

Often, it is not possible to represent adequately, system characteristics such as e.g. non-linearity, time delay, time-varying parameters and overall complexity in a mathematical model. System identification methods for analysing time-invariant linear dynamic systems are effective and relatively easy to use. The contrast with analysis of e.g. non-linear-dynamics is sharp. The steps for system identification for non-linear systems starts with detection of the existence of non-linearities, followed by quantitative assessment of their magnitude and modelling, e.g. by estimation of the corresponding parameters in a possible model.

According to Natke et al. [1] no unique approach seems to be possible. Many system identification methods for non-linear systems such as e.g. Volterra series tend to be useful in certain applications and to give only partial or approximate information and to be cumbersome, see e.g. Tomlinson et al. [2]. Further, system identification techniques for non-linear problems are sensitive to noise, some require accurate initial conditions, and they often need to assume the non-linearity form a priori, or by using a sufficient number of terms in a series to obtain a given accuracy while the physical insight often is lost in the process. In order to avoid some of these problems much research has been done with respect to model non-linear dynamic structures using artificial neural networks.
The potential value of neural networks in system identification is their ability to approximate non-linear functions using a generic function. Most of the work has been based on multilayered feedforward networks trained by the backpropagation algorithm, see e.g. Masri et al. [3], Qi et al. [4], Riva et al. [5]. However, such static neural networks cannot by themselves represent dynamic systems. The static neural network can only be used to model dynamic systems if the inputs and outputs to the network are selected based on the understanding of the system to be modelled. Generally, in order to model non-linear dynamic systems, it is necessary to use the so-called recurrent neural networks which incorporate feedback, see e.g. Hush et al. [6], Hertz et al. [7], Seidl et al.[8], Pham et al. [9] and Hansen et al. [10].
The objective of this paper is to investigate two different ways to incorporate feedback into a neural network in order to model a non-linear structural system. The two different recurrent neural network approaches are described in section 3. In section 4 the approaches are being used to model a viscous damped oscillator with hysteresis of the curve-linear type described by the Bouc-Wen model. The oscillator is subjected to amplitude modulated Gaussian white noise filtered through a Kanai-Tajimi filter. At last, in section 5 and 7 conclusions and references are given, respectively.

2. ARTIFICIAL NEURAL NETWORKS

Artificial neural networks are computational models inspired by the neuron architecture and operation of the human brain. The pioneering work in this field is usually attributed to McCulloch et al. [11]. They developed a simplified model of a neuron. The brain is composed of neurons of many different types, see e.g. McCulloch et al. [11]. For a more detailed description of neural networks, see e.g. Hertz et al. [7] and Hush et al. [6]. Since McCulloch and Pitts in 1943 there have been many studies of mathematical models of neural networks. Many different types of neural networks have been proposed by changing the network topology, node characteristics and learning procedures. Examples of those are e.g. the Hopfield network, see Hopfield [12], the Kohonen network, see Kohonen [13], and the so-called multilayered perceptron (MLP). The MLP is currently given the greatest attention by application developers, see e.g. Rumelhart et al. [14].

2.1 Multilayer Perceptron

The multilayered perceptron network belongs to the class of layered feed-forward nets with supervised learning. A multilayered neural network is made up of one or more hidden layers placed between the input and output layers, see fig. 1. Each layer consists of a number of nodes connected in the structure of a layered network. The typical architecture is fully interconnected, i.e. each node in a lower level is connected to every node in the higher level. Output units cannot receive signals directly from the input layer. During the training phase activation flows are only allowed in one direction, a feed-forward process, from the input layer to the output layer through the hidden layers. The input vector feeds each of the first layer nodes, the outputs of this layer feed into each of the second layer nodes and so on

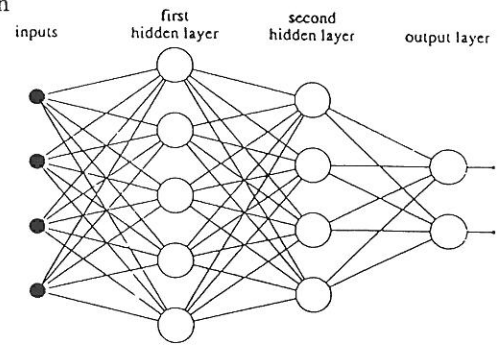

Fig. 2.1: Principle of a multilayer perceptron neural network.

Associated with each connection between node i in the preceding layer $l-1$ and node j in the following layer l is a numerical value $w_{lj,i}$ which is the strength or the weight of the connection. At the start of the training process these weights are initialized by small random values. Signal pass through the network and the jth node in layer l computes its output according to

$$x_{lj} = f\Big(\sum_{i=1}^{N_{l-1}} w_{lj,i} x_{l-1,i} - \theta_{lj}\Big) \qquad (1)$$

for $j = 1, .., N_l$ and $l = 1, .., k$, where x_{lj} is the output of the jth node in the lth layer. θ_{lj} is a bias term or a threshold of the jth neuron in the lth layer. The kth layer is the output layer and the input layer must be labelled as layer zero. N_l is the number of neurons in the lth layer, N_0 and N_k refer to the numbers of network inputs and outputs, respectively.

The function $f(\cdot)$ is called the node activation function and is assumed to be differentiable and to have a strictly positive first derivative. For the nodes in the hidden layers, the activation function is often chosen to be a so-called sigmoidal function

$$f(\beta) = \frac{1}{1+e^{-\beta}} \tag{2}$$

The activation function for the nodes in the input and output layers is often chosen as linear.
The first stage of creating an artificial neural network to model an input-output system is to establish the appropriate values of the connection weights $w_{lj,i}$ and thresholds θ_{lj} by using a learning algorithm. A learning algorithm is a systematic procedure for adjusting the weights and thresholds in the network to achieve a desired input/output relationship, i.e. supervised learning. The most popular and successful learning algorithm used to train multilayer neural networks is currently the so-called back-propagation routine, see e.g. Rumelhart [14]. The back-propagation algorithm normally employs a gradient descent search technique for minimizing an error, normally defined as the mean square difference between desired $y_j(t)$ and estimated outputs $\hat{y}_j(t) = \hat{x}_{kj}$ of output node j at all the times $t = 1, .., N$. I.e. the error J is given as

$$J = \frac{1}{2}\sum_{t=1}^{N}\sum_{j=1}^{N_k}(y_j(t) - \hat{y}_j(t))^2 \tag{3}$$

where N_k is the number of network output nodes. The thresholds are adjusted in the same way as the weights. The process of computing the gradient and adjusting the weights and thresholds is repeated until a minimum of the error J (or a point sufficiently close to the minimum) is found. However, it is generally true that the convergence of the back-propagation algorithm is fairly slow. Attempts to speed learning include variations on simple gradient search, line search techniques and second order techniques, see e.g. Hertz et al. [7], Billings et al. [15].
The MLP described above belongs to the static class of neural networks. Such network is only capable of modelling dynamic systems if the inputs and outputs to the network are established from knowledge of the system to be identified. Therefore, in general, it is necessary to use artificial recurrent neural networks with feedback to model dynamic systems, see Hertz et al. [7]. This class of artificial neural networks with feedback, and therefore, inherently recursive is commonly referred to as partially or fully recurrent neural networks. A fully recurrent neural network has feedback between all the nodes.

3. MODELLING OF NON-LINEAR DYNAMIC STRUCTURES USING ARTIFICIAL RECURRENT NEURAL NETWORKS

In the following two approaches to incorporate feedback into an MLP in order to model a non-linear dynamic single degree-of-freedom system are described. Both approaches are partially recurrent neural networks as output from the nets are used as input in the next step.

3.1 Tapped-Delay-Line (TDL) Recurrent Neural Network

A partially MLP neural network where there is a feedback of the output of the network through a TDL is very general, and is capable, in theory, of modelling any system which can be expressed as a non-linear function $G(\cdot)$

$$\hat{X}(t) = G\big(\hat{X}(t-\Delta t), .., \hat{X}(t-n\Delta t), ...$$

$$...., F(t-\Delta t), .., F(t-m\Delta t), \mathbf{w}\big) \tag{4}$$

where $\hat{X}(t)$ and $F(t)$ are the estimated scalar response and the scalar excitation at the discrete time t, respectively. Δt is the sampling time interval and $\mathbf{w}$ is a vector including the weights and thresholds. The function $G(\cdot)$ can be modelled by an MLP artificial neural network with n displacement input nodes, m excitation input nodes and a given number of hidden layers and one output node.
One difficulty with recurrent networks is developing meaningful teaching algorithms. Since the output of the nodes is a recursive function of the output of the nodes on the previous step, the calculation of the gradient must also be a recursive computation. The gradient of the error J with respect to the weights for a recurrent neural network modelling (4) is estimated by

$$\frac{\partial J}{\partial w_{lj,i}} = -\sum_{t=1}^{N}(y(t) - \hat{y}(t))\frac{\partial \hat{y}(t)}{\partial w_{lj,i}} \tag{5}$$

where the desired output $y(t) = X(t)$. The last derivative in (5) is found by differentiating (1). Since $\hat{y}(t) = x_{k1}$ this derivative becomes

$$\frac{\partial x_{k1}}{\partial w_{lj,i}} = f'(\cdot)[\delta_{1j}\delta_{kl}x_{k-1,i} + \sum_{m=1}^{N_{k-1}} w_{k1,m}\frac{\partial x_{k-1,m}}{\partial w_{lj,i}}] \tag{6}$$

where $f'(\cdot)$ is the derivative of the activation function. The gradient of the error J with respect to the thresholds is estimated in the same way as described above.

3.2 Recurrent Neural Network Modelling of an Innovation State Space Model

In this section a neural network model is formulated by a Nonlinear Innovation State Space Model (ISSM) where p state variables are observed, see Sørensen [16].

$$\hat{\mathbf{X}}(t) = G\left(\hat{\mathbf{X}}(t-\Delta t), F(t-\Delta t), \mathbf{E}(t-\Delta t), \mathbf{w}\right) \tag{7}$$

$$\mathbf{Y}(t) = \mathbf{H}\hat{\mathbf{X}}(t) + \mathbf{E}(t) \tag{8}$$

$\hat{\mathbf{X}}(t)$ is the estimate of the state vector at the discrete time step t. $\mathbf{E}(t)$ is the prediction error vector of order p at the discrete time step t. $\mathbf{Y}(t)$ is the measured observation vector of order p at the discrete time step t. $\mathbf{H}$ is the observation matrix of order $p \times n$. Incomplete State Information may occur, i.e. $\hat{\mathbf{X}}(t)$ is not completely measurable. The matrix $\mathbf{H}$ can be chosen to $\mathbf{H} = [\mathbf{I}_{p,p}\mathbf{0}_{p,n-p}]$, where $\mathbf{I}_{p,p}$ is a $p \times p$ unity matrix and $\mathbf{0}_{p,n-p}$ is a $p \times (n-p)$ zero matrix. In this way the elements in $\mathbf{H}\hat{\mathbf{X}}(\mathbf{t})$ are equal to the first p elements of the obsevation vector $\hat{\mathbf{Y}}(t)$.
In order to train an MLP to learn the Innovation State space model in (7) and (8) an MLP with the state vector, the scalar excitation and the prediction error vector are used as input and the state vector as output. This means that the desired outputs of the network $y_j(t)$ are the elements of $\hat{\mathbf{Y}}(t)$. The gradient of the error J with respect to the weight vector $\mathbf{w}$ for the recurrent neural network is estimated by

$$\frac{dJ}{d\mathbf{w}^{\mathbf{T}}} = -\sum_{t=1}^{N} \mathbf{E}^T(t)\mathbf{H}\frac{d\hat{\mathbf{X}}(t)}{d\mathbf{w}^T} \tag{9}$$

The gradients of the estimated state variables $\hat{\mathbf{X}}$ with respect to the weights $\mathbf{w}$ can be updated from the recursive equation

$$\mathbf{\Psi}(t) = \frac{d\hat{\mathbf{X}}(t)}{d\mathbf{w}^T} = \mathbf{\Lambda}(t) + \left[\hat{\mathbf{\Phi}}(t) - \hat{\mathbf{K}}(t)\mathbf{H}\right]\mathbf{\Psi}(t-1) \tag{10}$$

where

$$\mathbf{\Lambda}(t) = \frac{\partial\hat{\mathbf{X}}(t)}{\partial\mathbf{w}^T} \tag{11}$$

$$\hat{\mathbf{\Phi}}(t) = \frac{\partial\hat{\mathbf{X}}(t)}{\partial\mathbf{X}^T(t-\Delta t)}, \hat{\mathbf{K}}(t) = \frac{\partial\hat{\mathbf{X}}(t)}{\partial\mathbf{E}^T(t-\Delta t)} \tag{12}$$

$\mathbf{\Lambda}(t)$, $\hat{\mathbf{\Phi}}(t)$, and $\hat{\mathbf{K}}(t)$ can be found by differentiation of the output of the network in (7) with respects to the weight vector and input as in the same way it is done in (6) for the TDL neural network. $\hat{\mathbf{\Phi}}(t)$ and $\hat{\mathbf{K}}(t)$ can be interpreted as the dynamic transfer matrix of an equivalent linear system and the corresponding Extended Kalman Filter (EKF) gain, respectively. I.e. that learning the Innovation State Space model corresponds to solving the EKF problem. In Kirkegaard et al. [17] this problem is investigated.

4. EXAMPLE

In this numerical example the recurrent neural network approaches described in section 3 are validated and compared with application to simulated data from a viscous damped oscillator with hysteresis.

4.1 The Equation of Motion

The equation of motion for a single degree-of-freedom hysteretic oscillator may be written by the stochastic differential equation

$$\ddot{X}(t) + 2\zeta\omega_0\dot{X}(t) + \omega_0^2 g(X(t), Z(t)) = F(t) \tag{13}$$

where $X(t)$, $\dot{X}(t)$ and $\ddot{X}(t)$ are displacement, velocity and acceleration, respectively. $F(t)$ is the external excitation. ζ and ω_0 are the damping ratio and circular eigenfrequncy, respectively. g represents the hysteretic restoring term which is a functional of the preceding deformation history $X(t)$. Generally, g can be decomposed into a linear, non-hysteretic term and a hysteretic component. Thus

$$g = \alpha X(t) + (1-\alpha)Z(t) \tag{14}$$

The non-dimensional factor α is the ratio of post-yield stiffness to pre-yield (elastic) stiffness. It measures the relative contribution of the non-hysteretic term, i.e. $\alpha = 1$ corresponds to purely elastic restoring. $Z(t)$ is the hysteretic component of the restoring force. Here $Z(t)$ is of the curve-linear type described by the Bouc-Wen model, proposed by Bouc as extended by Wen, see e.g Wen [18]. The model is available to represent a large class of hysteretic systems.

$$\dot{Z} = \left(1 - \beta sign(\dot{X})\frac{|Z|^{n-1}Z}{z_0^n} - \gamma\frac{|Z|^n}{z_0^n}\right)\dot{X} \tag{15}$$

γ, β and n are "loop parameters". The parameters γ and β define the shape of the hysteresis loop. n is the smoothness of the transition from elastic to plastic response. A large value of n, e.g. n =5, corresponds to almost an elastic-plastic system. If $\beta+\gamma=1$, z_0 can be interpreted as the yield displacement of the oscillator.

The oscillator is subjected to modulated Gaussian white noise filtered through a Kanai-Tajimi filter. The excitation process $F(t)$ is then obtained from the stochastic differential equations, see Tajimi [19]

$$F(t) = 2\zeta_s\omega_s\dot{U}(t) + \omega_s^2 U(t) \qquad (16)$$

$$\ddot{U}(t) + 2\zeta_s\omega_s\dot{U}(t) + \omega_s^2 U(t) = -r(t)W(t) \quad (17)$$

$F(t)$ can then be interpreted as the negative of the ground surface acceleration, and ζ_s and ω_s are the damping ratio and the circular eigenfrequency of a single degree-of-freedom shear model of the underlying subsoil. $\{W(t), t \in]0, \infty[\}$ is unit intensity white noise with the auto-spectral density function $\frac{1}{2\pi}$. The deterministic modulation function is given as follows, see Saragoni et al. [20].

$$r(t) = r_0 \exp\left(-\ln(2) \frac{\frac{t}{t_1} - \ln\frac{t}{t_1} - 1}{\frac{t_2}{t_1} - \ln\frac{t_2}{t_1} - 1} \right) \qquad (18)$$

The amplitude r_0 defines the strength of the excitation. t_1 and t_2 are respectively the instants of the time of maximum intensity and the time where the intensity has dropped to half value. White noise sequences are generated by the broken line process model of Ruiz and Penzien, see Clough et al. [21].
Assuming the excitation $F(t)$ is constant at the value $F_{k-1} = F(t_{k-1})$ throughout the interval $[t_{k-1}, t_k[$, where $t_k = k\Delta t$, (13) is integrated on state space form using a 4th Runge-Kutta scheme. The state vector for (13), (14) and (15) is $\mathbf{X}(t) = (X(t), \dot{X}(t), Z(t))^T$. Normally, the state variables $X(t)$ and $\dot{X}(t)$ can be observed, whereas $Z(t)$ is non-observable.
The following parameters are used for simulating the hysteretic oscillator:

$$\zeta = 0.01, \quad \omega_0 = 2\pi s^{-1}, \quad \alpha = 0.05$$
$$\beta = \gamma = 0.5, \quad n = 1, \quad z_0 = 0.01m$$
$$\zeta_s = 0.5, \quad \omega_s = 30s^{-1}, \quad t_1 = 3s, \quad t_2 = 15s$$

z_0 has been selected so that $E[\max(|x(t)|)] \approx 2z_0$, where $\max(|x(t)|)$ is the peak displacement of the oscillator during a certain realization.

4.2 Training of Neural Networks

The data sets used for training of the TDL neural network are the displacement response $X(t)$ and the excitation $F(t)$, respectively. The ISSM neural network is also trained with the velocity $\dot{X}(t)$ component, i.e $Z(t)$ is not observed. Each simulated signal consists of 2000 points which are sampled with a sampling period of 0.01 s. In the training phase 663 points with a period of 0.03 s. are used for the TDL neural network. It was found that the TDL could not learn the time series with a smaller period than 0.03 s. The ISSM network was trained with all of the points. The TDL neural network is constructed by trial and error with 6 displacement nodes and 5 excitation nodes in an input layer, 9 nodes in a hidden layer and one node in the output layer. The ISSM neural network is constructed with the state vector, the prediction error vector and the scalar excitation in the input layer and 7 nodes in a hidden layer. The outputs from the ISSM net are the components in the state vector. In this paper the minimization of the error J is solved using the method of sequential quadratic programming, see Schittkowski [22]. The weights and thresholds are changed after the full-time interval $t = 1, .., N$ by summing the gradients for each time step. As stop criteria cross-validation is used.
The two neural networks were trained with the time series shown in figure 4.1

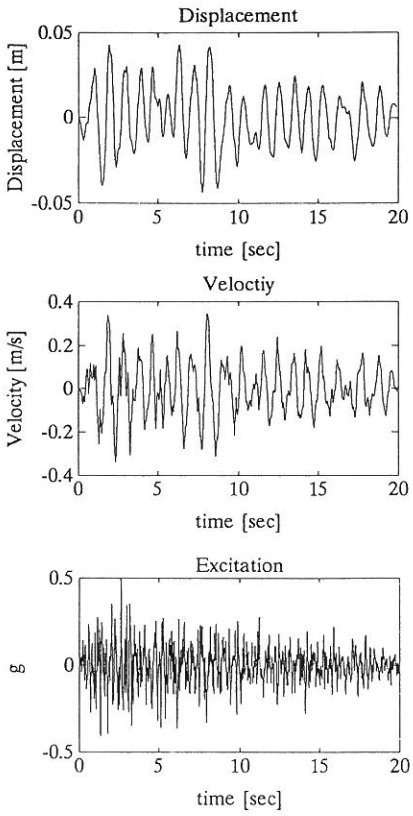

Fig. 4.1: Training data

4.3 Prediction

Figure 4.2 shows the comparison of the simulated displacement and the displacements predicted using the TDL and ISSM neural networks as an one step ahead predictors. I.e. the input to the networks are assumed to be measured for each time step. The initial displacement values for the TDL network and the initial state vector for the ISSM network are fixed to zero, respectively. The predicted displacements are obtained for other time series than used for training. This means that the generalization capability of the trained network can be checked.

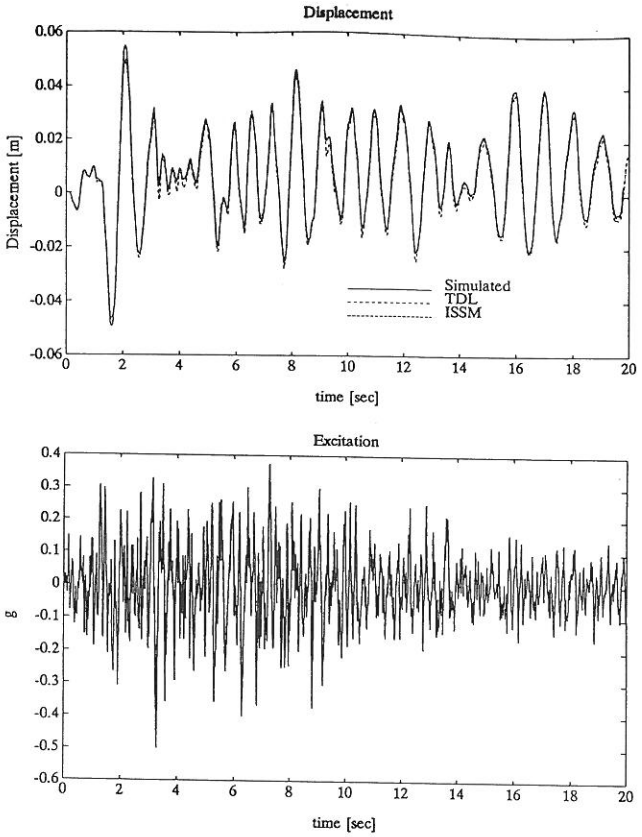

Fig. 4.2: Comparison of neural network predicted displacements and Runge-Kutta simulated displacements.

The results presented in figure 4.2 show that the trained recurrent neural networks can represent the dynamic of the hysteretic oscillator. It is observed that the neural networks are very successful at estimating the time and magnitude of the peak displacements. Further, the amplitude and frequency contents are also acceptably represented. Especially, it is interesting that the trained networks subjected to a time series different from the training data are capable of modelling the oscillator.
Figure 4.3 and figure 4.4 show the autocorrelation of the prediction errors and the cross-correlation of the prediction errors and input, respectively. The dotted straight lines in the figures show the 95% confidence limits. The confidence limits for the autocorrelation are based on χ^2 test for whiteness, see e.g. Söderström et al. [23]. The prediction errors (residuals) are the difference between the simulated and the predicted displacements. The autocorrelation test shows if there is any correlation in the prediction errors. In an ideal identification the prediction errors would be identical to a white-noise sequence. The cross-correlation test shows if there is any correlation between the prediction errors and the input. The existence of any correlation for positive correlation lags suggests that there are unidentified frequency components in the prediction errors, and the model seems to be underparameterized. The existence of any correlation for negative lags, however, indicates that the input of the structure is influenced by its output. In other words, there is a feedback in the structure, see e.g. Söderström et al. [23]. From figure 4.3 it is concluded that the prediction errors are almost white within the 95% confidence interval, and that within the same confidence interval it is concluded from figure 4.4 that the errors do not correlate with input.

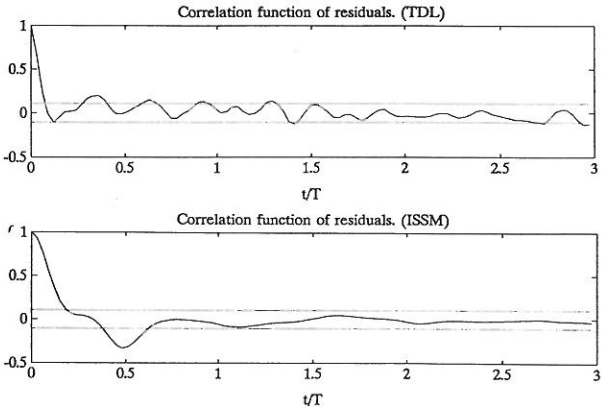

Fig. 4.3: Correlation functions of residuals

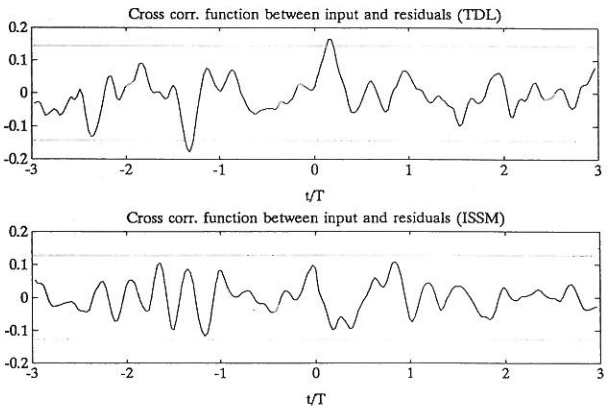

Fig. 4.4: Cross-correlation function between residuals and input.

In order to compare the predicted displacements from the TDL and ISSM neural network 25 simulations are performed. The averaged mean value μ_{N_s} and averaged standard deviation σ_{N_s} of the prediction errors for the ISSM and the TDL network, respectively are shown in figure 4.5.

$$\mu_{N_s} = \frac{1}{N_s}\sum_{j=1}^{N_s}\frac{1}{N\Delta t}\int_0^{N\Delta t}\epsilon_j(t)dt \qquad (19)$$

$$\sigma_{N_s} = \sqrt{\frac{1}{N_s}\sum_{j=1}^{N_s}\frac{1}{N\Delta t}\int_0^{N\Delta t}(\epsilon_j(t)-\mu_{N_j})^2 dt} \qquad (20)$$

$\epsilon_j(t)$ is the prediction errors of the jth simulation N_j of the N_s simulations. (19) and (20) are estimated for the TDL and ISSM network, respectively.
Figure 4.5 shows only small deviations between the TDL and ISSM network obtained average mean values and the standard deviation values, respectively. This means that there is no large difference between the predicted displacements obtained by the TDL and the ISSM neural network.

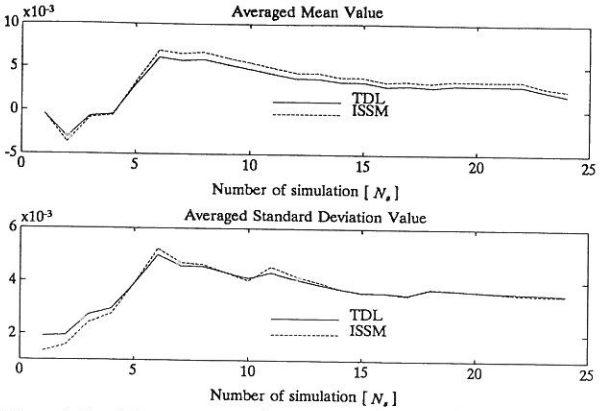

Fig. 4.5: The averaged mean value and the standard deviation value of the prediction errors for the neural networks, respectively.

4.4 Simulation

The capability of the TDL and ISSM recurrent neural networks to act as simulators will now be investigated. Figure 4.6 shows the results of the neural networks simulations compared with a Runge-Kutta simulation. Simulation with the TDL neural network is done by subjecting the excitation time series to the network. Further, it is assumed that all the displacement input nodes are zero at the start. Simulation with the ISSM neural network is performed by subjecting the excitation time series to the network and keeping the prediction error vector fixed to zero for all time steps. The initial state vector is fixed to zero.

The results presented in figure 4.6 show that the trained recurrent neural networks can give an acceptable simulation of the displacements of the hysteretic oscillator. It is observed, as for the predictions, that the neural networks are very successful at estimating the time and the magnitude of the peak displacements. Further, the amplitude and frequency contents are again acceptably represented.

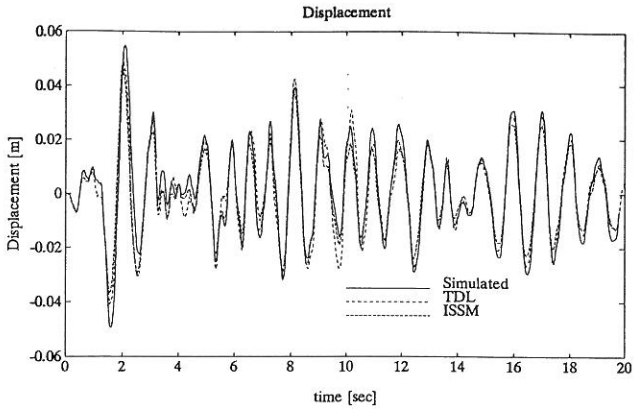

Fig. 4.6: Comparison of neural network simulated displacements and Runge-Kutta simulated displacements.

The simulations obtained with the neural networks are compared in the way as the predictions. This means that the averaged mean value and standard value of the prediction errors for the neural networks have been estimated for 25 simulations, respectively.

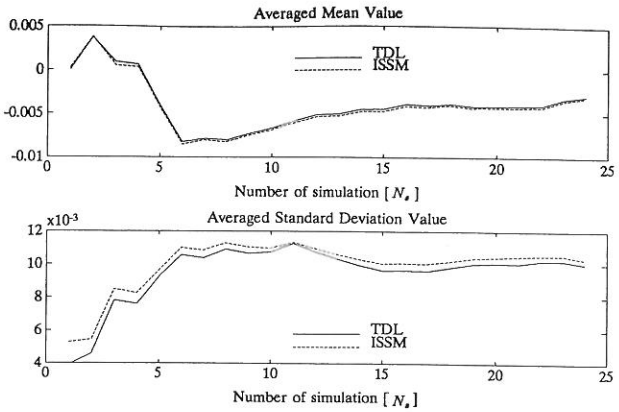

Fig. 4.7: The averaged mean value and the standard deviation value of the prediction errors for the neural networks, respectively.

From figure 4.7 it is seen that the TDL neural network gives a smaller averaged standard deviation of the prediction errors than the ISSM network which implies that the TDL network seems to be a better simulator than the ISSM network.

5. CONCLUSIONS

This paper presents two different partially recurrent neural networks structured as Multi Layer Perceptrons (MLP) for time domain identification of a non-linear structure. The results show that the two recurrent neural networks are capable of modelling a viscous damped oscillator with hysteresis subjected to stochastic dynamic loading. Of particular interest, the results, obtained by using the networks for prediction and simulation, show that the time and magnitude of the peak displacement, as well as the amplitude and frequency content are modelled very well. From the results where the two neural networks are used as a simulator it seems as the TDL neural network is a better simulator than the ISSM network.

6.ACKNOWLEDGEMENTS

This paper is partially supported by the research project "Safety and Reliability" sponsored by the Danish Technical Research Council which is greatly acknowledged.

7. REFERENCES

[1] Natke, H. G., G. Tomlinson & J.T.P Yao: Safety Evaluation Based on Identification Approaches Related to Time-Variant and Nonlinear Structures. Vieweg and Sohn, 1993.

[2] Tomlinson, G. & S.J. Gifford: Recent Advances in the Application of Functional Series to Non-linear Structures. Journal of Sound and Vibration, Vol. 135, No. 2, 1989.

[3] Masri, S.F., A.G. Chassiakos & T.K. Caughey: Structure-unknown Non-Linear Dynamic Systems: Identification Through Neural Networks. Smart Mater. Struct. Vol. 1, pp. 45-56, 1992.

[4] Qi, G.Z., H.M. Chen, K.H. Tsai & J.C.S. Yang: Structural Dynamic Model Identification Through Neural Network. Proceedings of the 12th Int. Modal Analysis Conf., pp. 1167-1172, 1994.

[5] Riva, A., L. Garibaldi, E. Giorcelli & A. Fasana: Dynamic System Identification by means of Neural Network. Proceedings of the 11th Int. Modal Analysis Conf., pp. 928-933, 1993.

[6] Hush, D.R. & B.G. Horne: Progress in Supervised Neural Networks. IEEE Signal Processing Magazine, January, 1993.

[7] Hertz, J., A. Krogh & R.G. Palmer: Introduction to the Theory of Neural Computation. Addison-Wesley, Redwood City, CA, 1991.

[8] Seidl, R.D. & R.D. Lorenz: A Structure by which a Recurrent Neural Network Can Approximate a Nonlinear Dynamic System. IJCNN, Seattle, 1991.

[9] Pham,D.T. & S.J. Oh: A Recurrent Backpropagation Neural Network for Dynamic System Identification. Journal of Systems Engineering, Vol. 2, 1992.

[10] Hansen, H.I., P.H. Kirkegaard & S.R.K. Nielsen: Modelling of Deteriorating RC-Structures under Stochastic Dynamic Loading by Neural Networks. Structural Reliability Theory, Paper No. 125. ISSN 0902-7513 R9409 Proc. of the 2nd Int. Conf. on Stochastic Computational Mechanics, Athens, Greece, June 13-15, 1994

[11] McCulloch, W.S. & W. Pitts: A Logical Calculus of the Ideas Immanent in Nervous Activity. Bull. Math. Biophys. Vol. 5, pp. 115-133, 1943.

[12] Hopfield, J.J: Neural Networks and Physical Systems with Emergent Collective Computational Abilities. Proc. Nat. Acad. Sci. Vol. 79, pp. 2554-2558, 1982.

[13] Kohonen, T.: Self-Organization and Associative Memory. Springer-Verlag, 1984.

[14] Rumelhart, D.E. & J.L. McClelland: Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1 Foundations. Cambridge, MIT Press, 1986.

[15] Billings, S.A., H.B. Jamaluddin & S. Chen: A Comparison of the Backpropagation and Recursive Prediction Error Algorithms for Training Neural Networks. Mechanical Systems and Signal Processing, Vol. 5, pp. 233-255, 1991.

[16] Sørensen, O.: Neural Networks in Control Applications. Ph.D-thesis from Aalborg University, 1994.

[17] Kirkegaard, P.H, S.R.K. Nielsen & H.I. Hansen: Structural Identification by Extended Kalman Filtering and a Recurrent Neural Network. Paper from Aalborg University, 1994.

[18] Wen, Y.K: Methods for Random Vibration of Hysteretic Systems. Journal of the Engineering Mechanics Division, ASCE, Vol. 102, No. EM2, pp.249-263, 1976.

[19] Tajimi, H. : A Standard Method of Determining the Maximum Response of a Building Structure during an Earthquake. Proc. 2nd World Conference on Earthquake Engineering, Vol. 2, 1960.

[20] Saragoni, G.R. & G.C. Hart: Simulation of Artificial Earthquakes. Earthquake Eng. Struct. Dyn., Vol. 2, pp. 249-267,1974.

[21] Clough, R.W. & J. Penzien: Dynamics of Structures. McGraw-Hill, New York, 1975.

[22] Schittkowski, K.: NLPQL: A FORTRAN Subroutine Solving Constrained Non-Linear Programming Problems. Annals of Operations Research, Vol. 5, 1985.

[23] Söderström, T. & P. Stoica: System Identification. Prentice Hall, 1987.