Neural Networks 22 (2009) 861–863

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Neural networks letter

Effects of spectral radius and settling time in the performance of echo state networks

Ganesh K. Venayagamoorthy *, Bashyal Shishir

Real-Time Power and Intelligent Systems Laboratory, Missouri University of Science and Technology, Rolla, MO 65409, USA

ARTICLE INFO

Article history:
Received 10 January 2009
Accepted 22 March 2009

Keywords:
Echo state network
Function approximation
Settling time
Spectral radius
System monitoring
Time series prediction

ABSTRACT

Echo State Networks (ESNs) have tremendous potential on a variety of problems if successfully designed. The effects of varying two important ESN parameters, the spectral radius (α) and settling time (ST) are studied in this letter. Spectral radius of an ESN is the maximum of all eigenvalues of the reservoir weights whereas ST is measured by the number of iterations allowed in the reservoir after its excitation by an input and before the sampling of the ESN output. The influence of these parameters on the performance of an ESN is illustrated using three different types of problems. These problems include a function approximation, a time series prediction and a complex system monitoring/estimation. An α of 0.8 gives best result in all of these experiments and the performance of the ESN degrades when ST is increased. This degradation in the ESN's performance is due to the decaying of the echoes and attenuation in the reservoir. The increase in ST adversely affects the ESN performance and as such no long-term echoing arrangement is desired. Reducing ST greatly reduces the computational requirement making ESNs suitable even for tasks that require a high frequency of operation.

1. Introduction

Echo State Network (ESN) is a Recurrent Neural Network (RNN) learning architecture characterized by a large randomly connected, recurrent reservoir passively excited by the input signal such that the network can be trained using the readout weights that combine the desired output from the reservoir (Jaeger (2005)). ESNs provide a novel and easier to manage approach to supervised training of RNNs (Jaeger (2003)). Several learning algorithms are known that incrementally adapt the synaptic weights of a RNN to tune it towards the target system but such algorithms are rarely used due to slow convergence and suboptimal performance (Hass and Jaeger (2004)). The ESN differs from other algorithms in two important ways: firstly, a large reservoir (50–1000 neurons) is used and secondly, only the readout weights are updated during learning as opposed to other algorithms that tune all synaptic connections. These differences render ESN training into a simple linear regression task. Details on ESN training are provided in Jaeger (2002). The structure of an ESN is shown in Fig. 1.

The simplicity of ESN has led to its wide acceptance and several applications have been reported since its introduction in 2001. ESN has been used for motor control (Salmen and Ploger (2005)), nonlinear system identification (Jaeger (2003)), robotics (Jaeger, Hertzberg, and Schonherr (2002); Xu, Lan, and Principe (2005)), wide area monitoring of power systems (Venayagamoorthy (2007)), motion identification (Ishu, van der Zant, Becanovic, and Ploger (2004)), gait motion analysis (Noris et al. (2006)) and also for saving energy in wireless sensor networks (Hass and Jaeger (2004)). However, being a new technique, ESN has several issues regarding appropriate selection of ESN parameters (Jaeger (2005)) and practicality (Prokhorov (2005)). While ESN offers a simple method to deploy RNNs, it is agreed that it still requires further research to make the design process straightforward and to make ESN applicable in solving real-world problems.

H. Jaeger, the founder of ESN (Jaeger (2005)), identified the following parameters to be appropriately selected for developing an effective ESN: network size, spectral radius of reservoir weight matrix and the scaling of the input. As with all other artificial neural networks, choosing a right set of network parameters is important for successful design of an ESN. In this letter, two important parameters, the spectral radius and the settling time of the ESN reservoir are varied to develop an understanding of their respective importance in the design of ESN. The spectral radius (α) of an ESN is the maximum of all eigenvalues of the reservoir weights. As proposed in Jaeger (2002), the spectral radius of the ESN reservoir should be between 0 and 1 to ensure that the network has the echo state property. Settling Time (ST) is measured by the number of iterations (echoes) allowed in the reservoir after its excitation by an input and before the sampling of the output. A lower ST means

* Corresponding author. Tel.: +1 573 3416641; fax: +1 573 3414532.
E-mail address: gkumar@ieee.org (G.K. Venayagamoorthy).

0893-6080/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2009.03.021

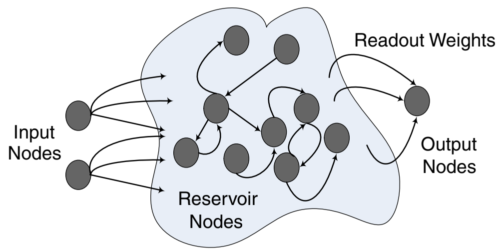


Fig. 1. A typical ESN.

that the echoing is short termed or truncated and the output is available immediately after the input is applied to the reservoir. A higher ST means more echoing in the network and the output is delayed.

By performing different case studies, the performance of ESN for different problems is studied to develop heuristics for designing ESN. In all of the studies discussed in this letter, the reservoir size is kept constant at 50 nodes. Sigmoidal activation functions are used in the reservoir nodes. The output nodes act only as summation units. The input, reservoir and readout weights are randomly initialized at the beginning of the training. The spectral radius of the ESN reservoir is varied between 0.2 and 0.8, in steps of 0.2. The readout weights are found using the gradient descent learning rule in 100 iterations. With four different values for α and two values for ST (one and five), a total of eight experiments are carried out in each case study. In each experiment, ESNs are trained and the Mean Square Error (MSE) vs. training iteration plot of over 10 trials is used to compare the learning performance of the ESN.

2. Case study 1: Function approximation

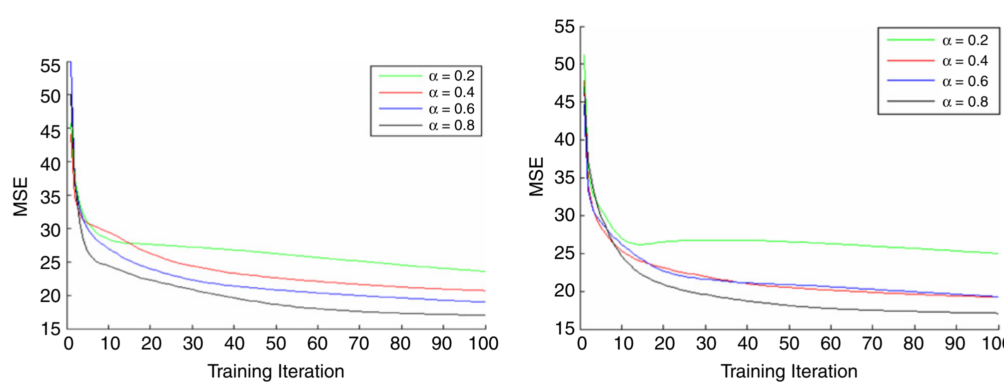
In this case study, ESN is trained to approximate a two-dimensional Sinc function as defined in (1)

$$f = \frac{\sin(x) \times \sin(y)}{x \times y}. \tag{1}$$

Here, the spectral radius (α) is varied from 0.2 to 0.8 in steps of 0.2. The ESN has two input nodes and one output node. The ESN input during training is limited to between -10 to 10 and is sampled at an interval of one. In addition to varying α of the reservoir, the ST of the ESN is also experimented with. The ESN is allowed to settle for certain number of iterations (one and five) before reading the output of the ESN. For both values of ST, higher values of α result in faster learning of the ESN. The increase in settling time did not improve the learning rate of the ESN. Fig. 2 shows the training error for different combinations of α and ST.

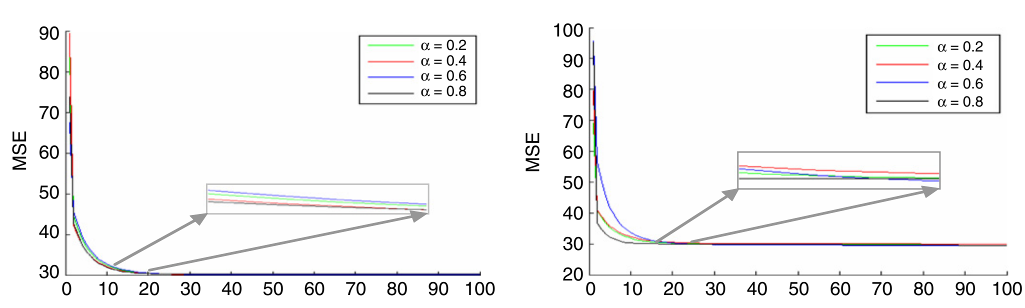
3. Case study 2: Time series prediction

In this study, a single input single output ESN is used to predict the standard Mackey–Glass time series. The task is to predict the $(\tau + 6)$th sample when the τth sample is the input to the ESN. The number of neurons in the reservoir is kept to 50 as in the first case study. Similarly, α is varied from 0.2 to 0.8 and ST is changed from one to five. Of the total 1200 time samples in the Mackey–Glass time series test set, only the first 1000 samples are used in the ESN training. The result of varying α and the ST is shown in Fig. 3. The ESN training behavior is found to be almost identical for all possible combinations of α and ST except for the fact that $\alpha = 0.8$ led to faster convergence in both cases. Fig. 4 shows the predicted series from the actual series when $\alpha = 0.8$ and ST=1.



(a) ST $= 1$. (b) ST $= 5$.

Fig. 2. ESN learning a two-dimensional Sinc function.



(a) ST $= 1$. (b) ST $= 5$.

Fig. 3. ESN learning the Mackey–Glass series.

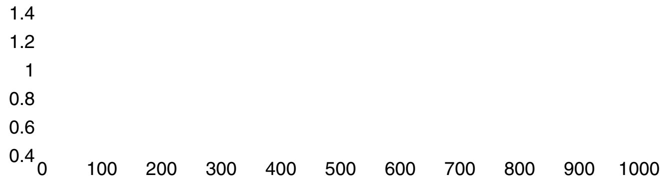


Fig. 4. Predicted and actual Mackey–Glass time series.

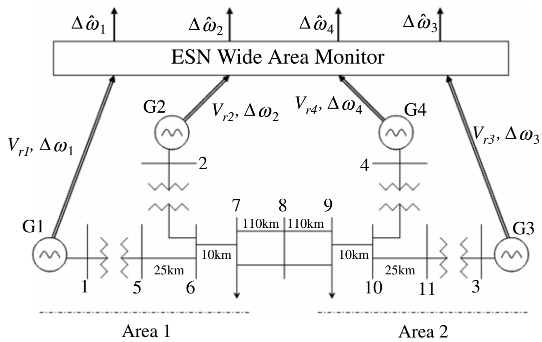


Fig. 5. ESN for system monitoring/speed estimation.

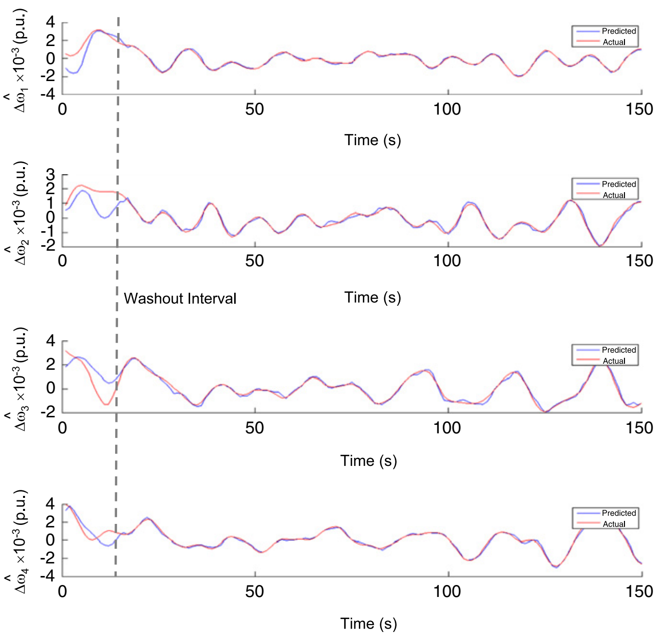


Fig. 6. Estimated speed deviations vs. actual speed deviations.

4. Case study 3: Complex system monitoring

In this study, the ESN is used for wide area system monitoring of the two-area power network as proposed in Venayagamoorthy (2007). The two-area power system consists of two fully symmetrical areas tied together by two transmission lines. Each area is equipped with two identical synchronous generators rated 20 kV/900 MVA. All the generators are equipped with two identical speed governors and turbines, and voltage regulators and exciters, as described in Venayagamoorthy (2007). Fig. 5 shows the two-area power system and the ESN inputs for the estimation of four generator speed deviations simultaneously. The power system is simulated in real time on the Real Time Digital Simulator (RTDS) (Forsyth, Maguire, and Kuffel (2004)). The ESN with $\alpha = 0.8$ and ST of one performed the best. Fig. 6 shows the performance of the ESN ($\alpha = 0.8$, ST=1) in estimating the speed deviations for all the four generators simultaneously.

5. Conclusions

The results obtained in all the three case studies above are comparable and the performance of the echo state network improves with higher spectral radius. Of the four values experimented for spectral radius, α, 0.8 gives the best result in all the cases. The increase in settling time however degraded the ESN performance in all of the case studies. ESNs retain the impact of past inputs as the echoes in the reservoir and thus are able to perform well in prediction problems. When the ST is increased, the impact of the past inputs decays and hence the learning of the network is compromised. The results show that α of 0.8 is optimal for use in designing echo state networks and also that the increase in ST does not help the network in learning but rather degrades the learning capability of ESNs. This relieves the ESNs from the additional computational complexity that would otherwise be required. The consistent results obtained by using the same set of parameters on a variety of problems instills confidence in these conclusions made.

References

Forsyth, P., Maguire, T., & Kuffel, R. (2004). Real time digital simulation for control and protection system testing. In *Proc. of IEEE 35th annual power electronics specialists conference*: *Vol. 1* (pp. 329–335).

Hass, H., & Jaeger, H. (2004). *Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless telecommunication*. International University Bremen.

Ishu, K., van der Zant, T., Becanovic, V., & Ploger, P. (2004). Identification of motion with echo state network, OCEANS '04. MTS/IEEE TECHNO-OCEAN '04, Volume 3.

Jaeger, H. (2002). Tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the 'echo state network' approach, Sankt Augustin: GMD Forschungszentrum Informationstechnik, 48S, GMD Report, 159.

Jaeger, H. (2003). *Adaptive nonlinear system identification with echo state networks*. Frauenhofer Institute for Autonomous Intelligent Systems.

Jaeger, H. (2005). Reservoir riddles: Suggestions for echo state network research, *IJCNN '05 proceedings*. Vol. 3 (pp. 1460–1462).

Jaeger, H., Hertzberg, J., & Schonherr, F. (2002). Learning to ground fact symbols in behavior-based robots, In F. van Harmelen, (Ed) *Proceedings of the 15th European conference on artificial intelligence*.

Noris, B., Nobile, M., Piccinini, L., Berti, M., Molteni, M., & Berti, E. et al. (2006). Gait analysis of autistic children with echo state networks, In *Workshop on echo state networks and liquid state machines*.

Prokhorov, D. (2005). Echo state networks: Appeal and challenges neural networks, 2005, In *IJCNN '05. Proceedings*. Volume 3, (pp. 1463–1466).

Salmen, M., & Ploger, P.G. (2005). Echo state networks used for motor control, robotics and automation, In *Proceedings of the 2005 IEEE international conference* (pp. 1953–1958).

Venayagamoorthy, G. K. (2007). Online design of an echo state network based wide area monitor for a multimachine power system. *Neural Networks*, *20*(3), 404–413. Echo State Networks and Liquid State Machines.

Xu, D., Lan, J., & J.C., Principe (2005). Direct adaptive control: An echo state network and genetic algorithm approach Neural Networks, In *Proceedings of the IJCNN '05* Vol. 3 (pp. 1483–1486).