Modeling non-linear communication systems using neural networks

Rachid Fateh[1], Hassan Ougraz[2], Youness Belaadel[3], Mathieu Pouliquen[4], Miloud Frikel[4], Said Safi[2]
[1]Department of Computer Sciences, Faculty of Applied Sciences, Ibn Zohr University, Agadir, Morocco
[2]LIMATI Laboratory, Sultan Moulay Slimane University, Beni Mellal, Morocco
[3]TIAD Laboratory, Sultan Moulay Slimane University, Beni Mellal, Morocco
[4]LIS Laboratory-UR 7478, ENSICAEN, Normandie University, UNICAEN, Caen, France

Article Info

Article history:

Received Apr 29, 2024
Revised Oct 22, 2024
Accepted Oct 30, 2024

Keywords:

Artificial neural networks
Hammerstein system
Mean squared error
Nonlinear systems
Telecommunications

ABSTRACT

Nonlinear systems present significant modeling challenges due to their complex dynamics and often unpredictable behavior. Traditional mathematical approaches can struggle to represent such systems accurately. In recent years, neural networks have emerged as promising tools to address this challenge. This article explores the use of neural networks to model nonlinear systems, focusing specifically on the application of the Hammerstein system. We examine network architecture and training methodologies suited to the complexity of nonlinear dynamics. Additionally, we explore strategies to improve the interpretability of neural network models in this context, enabling a better understanding of the underlying behavior of the system. Through a case study and empirical evaluations, we demonstrate the effectiveness of neural network-based approaches for estimating the behavior of nonlinear systems. Our work highlights the potential of neural networks as a versatile and powerful tool for modeling complex nonlinear phenomena.

Corresponding Author:

Rachid Fateh
Department of Computer Sciences, Faculty of Applied Sciences, Ibn Zohr University
Agadir, Morocco
Email: fateh.smi@gmail.com

1. INTRODUCTION

Telecommunications are crucial in our everyday lives, impacting our communication, jobs, and interaction with the world. Telecommunications, from basic phone calls to worldwide streaming multimedia material, are pervasive and indispensable in our contemporary society. The rise of digital technologies and wireless networks has significantly increased the significance of telecommunications, making it a crucial foundation of our global technical infrastructure [1]. Nevertheless, the rapid expansion of telecommunications is accompanied with a rising intricacy of the underlying systems. Contemporary telecommunications networks have several issues, such as optimizing resource allocation, reducing interference, guaranteeing service quality, handling large-scale data flows [2], and channel estimation [3]. The difficulties are worsened by the intrinsically non-linear properties of telecommunications networks, in which phenomena like signal fluctuations [4], distortions [5], identifications [6], equalization [7], and interference may result in highly complex and unexpected behaviors [8]. The complicated structure of non-linear systems presents a substantial obstacle in diverse domains, including engineering, physics, economics, and biology [9]. Linear systems can be described by simple cause-and-effect linkages, while non-linear systems display complex and non-proportional interactions among variables, resulting in dynamic behaviors. These interactions result in a diverse range of phenomena, notably oscillations, bifurcations, chaotic dynamics, and emergent patterns. The intricate nature of this problem is a barrier to our conventional comprehension and requires the use of advanced modeling approaches. Nevertheless, the intrinsic unpredictability and development of non-linear systems provide a significant challenge in accurately representing them solely using standard analytical methods. As a result, researchers have started using increasingly sophisticated techniques, such as computer modeling employing neural networks and Kernel methods, to accurately understand the intricate behaviors and emerging characteristics of these intricate systems [10], [11].

The modelling and identification of nonlinear systems pose substantial difficulties in systems engineering, especially when dealing with active distribution networks and intricate dynamic systems. Hammerstein-Wiener systems, which consist of nonlinear and linear blocks connected in series, provide a robust and efficient framework for modelling various dynamic behaviors. Hammerstein-Wiener models have proven effective in building dynamic counterparts of active distribution channels, which are important for analyzing and estimating long-term issues in electrical networks. This methodology enables enhanced control and optimization of renewable energy sources, hence expanding the stability and efficiency of distribution networks [12]. It is essential to have an advanced understanding of these dynamic equivalents to effectively tackle present and future energy management obstacles. Robust estimating approaches have been developed to improve the accuracy and reliability of models for identifying Hammerstein systems using quantized data, even when there are limitations in sensor precision. These strategies are necessary for offering optimum performance in situations when there is an insufficient or absence of accurate data [13]. The precision and dependability of these estimates are especially important in industrial and technical settings where quantized sensors are often used. Additionally, the process of modeling and identifying Hammerstein systems when there are inaccuracies in the output measurements presents considerable difficulties. Proposed advanced strategies aim to minimize the impact of noise, consequently ensuring both the precision and dependability of accepted models for a wide range of practical applications [14]. These approaches are essential for supporting the integrity of models in contexts where measurements are more susceptible to disturbances and interferences.

Traditional analytical models, which depend on differential equations or mathematical functions, frequently have limitations when it relates to accurately represent the wide range of behaviors found in non-linear systems. These models sometimes depend on oversimplified assumptions that may not accurately represent reality, especially in situations where the relationships between variables are intricate and not well comprehended. As a result, there is a growing need to prioritize the search for modeling methods that are more flexible and adaptive in various domains. In addition to techniques based on neural networks, kernel methods such as support vector machines (SVMs) and Kernel regression have also emerged as strong tools for modeling non-linear systems. These techniques exploit Kernel functions to transform data into feature spaces with larger dimensions, enabling the capturing of non-linear interactions with greater precision. Artificial neural networks (ANNs) have become a significant field of research in machine learning and artificial intelligence (AI), ANNs have been widely used in several domains such as classification, regression, prediction, pattern recognition, robotics, and signal processing within the last twenty years [15]–[21]. They provide a powerful technique for handling signals and information in telecommunications, distinguished by a high degree of adaptability and flexibility that distinguishes them from conventional methods [22], [23]. Due to their capacity to acquire intricate patterns from unprocessed data, they are very advantageous tools in several domains of signal processing [24]. Neural networks has the ability to detect non-linear patterns and trends in signals [25], hence facilitating more accurate and resilient analysis of telecommunications data [26]. Since there is their innate flexibility, they are able to modify their behavior to accommodate changing communication channel conditions, resulting in improved system security and efficiency. Moreover, neural networks possess the ability to acquire knowledge from large quantities of data, a critical aspect in communication fields characterized by extensive and intricate datasets. The application of them in applications including demodulation [27], signal detection [28], noise suppression [29], and channel prediction [30] has resulted in significant enhancements in terms of service quality, data transmission rate, and resistance to interference. In this paper, we present a novel approach that leverages ANNs as an alternate and complementary way of modeling nonlinear communication systems, specifically nonlinear Hammerstein systems. By exploiting the capabilities of ANNs, inspired by the complex working of the human brain, we want to overcome the constraints associated with previous methodologies and offer a more efficient and accurate way to capture the dynamics of complex communication networks.

The article is organized as follows: section 2 gives a brief description of the theoretical basis and challenging settings. In section 3, we present a full description of our proposed technique for modeling non-linear systems in the telecommunications domain using neural networks. The findings of our experiments are reported in section 4 followed by an examination of the importance of our technique. Section 5 concludes the paper.

2. THEORETICAL FOUNDATIONS AND PROBLEM CONTEXT

In this section, we introduce some notations and assumptions that will be used throughout the paper. The Hammerstein system, a distinctive nonlinear model, is frequently employed in the realm of system identification. System identification aims to create precise mathematical models that mirror the behavior of real-world systems using observed input-output data. In this context, we focus on the Hammerstein system depicted in Figure 1. This system comprises a nonlinear static function followed by a finite impulse response (FIR) filter with a known order. This structure is chosen for its ability to effectively represent both nonlinear and linear dynamics in a system, offering flexibility and interpretability during the identification process.

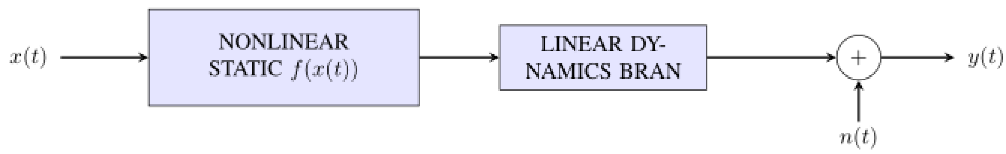


Figure 1. Block diagram of Hammerstein system

This nonlinear system can be expressed mathematically by (1).

$$y(t) = H_{\text{BRAN}} * [f(x(t))] + n(t) \tag{1}$$

Where:
$y(t)$ denotes the system's output at time t.
$x(t)$ represents the system's input at time t.
H_{BRAN} symbolizes the linear channel of the system, which is defined by its coefficients.
$f(\cdot)$ is the nonlinear function applied to the input.
$n(t)$ is the noise introduced to the output.
By putting out the following hypotheses, we hope to create a framework that encapsulates the key elements of the issue and makes it easier to create efficient estimation techniques.

- Hypothesis 1: to test a wide range of scenarios and make sure that the estimation model is resilient across various input conditions, we randomly select input values $x(t)$ within a specific range.
- Hypothesis 2: the noise, denoted as $n(t)$, is recommended to be additive white gaussian noise (AWGN) and independent of both $x(t)$ and $y(t)$.
- Hypothesis 3: considering that the linear channel H_{BRAN} is a BRAN (broadband radio access network) channel developed by the European Telecommunications Standards Institute (ETSI) [31], [32]. With the help of this hypothesis, we can concentrate on learning how noise and nonlinearities affect the behavior of the system without having to deal with the extra complications that come with time-varying channels.
- Hypothesis 4: the adoption of an ANN model presents an adaptable and efficient model for approximating the complicated interaction between input and output in the system.
- Hypothesis 5: the hyperbolic tangent function $\tanh(.)$ is an example of a continuous and differentiable nonlinear function that can be chosen to guarantee smooth and well-behaved transformations of the data.

All of these theories serve as the foundation for researching and resolving the difficulties involved in employing neural networks to estimate the output of Hammerstein systems. By carefully examining these features, we may create robust and reliable estimate approaches that have wide-ranging applications in engineering and signal processing areas. The fundamental purpose of this research is to use neural networks to estimate the output of Hammerstein systems created by (1).

Remark 1: the learning process of a neural network involves minimizing a loss function such as mean squared error (MSE). The optimization of network parameters $\mathbf{W}$ and $\mathbf{b}$ is achieved using optimization algorithms like stochastic gradient descent. Mathematically, the weight updates can be formulated as:

$$\mathbf{W}_{(t+1)} = \mathbf{W}_{(t)} - \eta \frac{\partial \text{MSE}}{\partial \mathbf{W}}, \quad \mathbf{b}_{(t+1)} = \mathbf{b}_{(t)} - \eta \frac{\partial \text{MSE}}{\partial \mathbf{b}}$$

where η is the learning rate and t is the iteration of the optimization algorithm.

Theorem 1 (convergence theorem [33]): under certain regularity conditions on the data and network architecture, neural networks converge to an optimal solution when minimizing a loss function like MSE through an optimization process. Mathematically, for a neural network with parameters θ and a loss function $J(W)$, we have:

$$\lim_{k \to \infty} \nabla J(W_k) = 0 \tag{2}$$

where W_k represents the network parameters at iteration k of the optimization algorithm.

3. MODELING APPROACH WITH ARTIFICIAL NEURAL NETWORKS

This section provides a comprehensive explanation of our methodology for estimating the output of the Hammerstein system using a multi-layer perceptron (MLP) neural network. The foundation of our methodology is based on using a MLP that is designed with a specific structure, acceptable activation functions, an effective optimizer, and an appropriate loss function to handle the identification issue. Initially, we chose a MLP because to its ability to record intricate connections between input and output data. The MLP was built with two hidden layers and one output layer. The hidden layers consist of 64 neurons each, whereas the output layer has a solitary neuron, which is ideal for our identification issue because we are estimating a single output value. The Figure 2 illustrates a standard MLP configuration consisting of two hidden layers and one output layer.

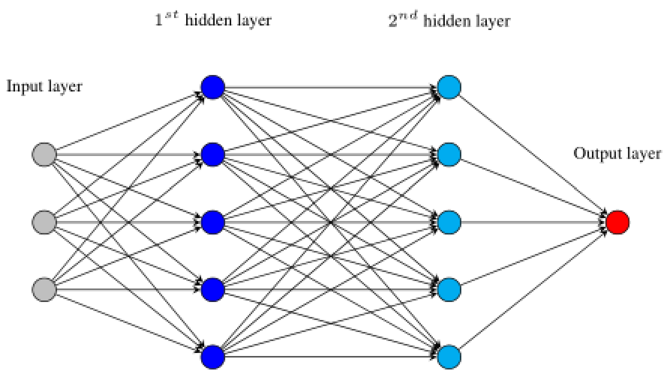


Figure 2. Multilayer perceptron

To inject non-linearity into the model's structure, the rectified linear unit (ReLU) activation function is utilized for the hidden layers. This activation is formally described as:

$$\mathcal{R}(x) = \max(0, x) \tag{3}$$

where x represents the weighted sum of inputs and biases for each neuron in the layer.

The optimizer is an algorithm for first-order gradient-based optimization of stochastic objective functions used to change the network's weights through training. This algorithm is a common optimization technique in the area of deep learning, able to adapt to the peculiarities of the task rapidly. Updates to the model parameters using this method are computed using (4).

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon} \cdot \hat{m}_t \tag{4}$$

where θ_t represents the model parameters at time t, η is the learning rate, and $\hat{m}_t$ and $\hat{v}_t$ are the estimates of the first and second order moments of the parameter gradients respectively.

Our learning technique is supervised, which means the neural network is trained using data that is labeled, which comprises input-output pairs. Through training, the network changes its parameters repeatedly to reduce the divergence in its predictions and the actual output values presented in the labeled dataset. Supervised learning is the ideal strategy for this issue, since it includes estimating the output of the Hammerstein system using data labeled. The method for estimating the output of Hammerstein system using neural networks is outlined as Algorithm 1.

Algorithm 1. Neural-based nonlinear system identification

Step 1: Input

- z_{mem}: Input data representing the memory of the nonlinear system
- y: Output data corresponding to the input data
- N: Total number of data points
- L: Length of the input vector
- *Percentage_split*: Percentage of data for training (e.g., 0.8 for 80% training, 20% testing)

Step 2: Main loop

while $\{x(t), y(t)\}_{t=1}^{N}$ available **do**

- Split the Data into Training and Test Sets

$$\begin{aligned}X_{\text{train}} &= z_{\text{mem}}[:\lfloor N \times 0.8 \rfloor]\\ X_{\text{test}} &= z_{\text{mem}}[\lfloor N \times 0.8 \rfloor :]\\ y_{\text{train}} &= y[:\lfloor N \times 0.8 \rfloor]\\ y_{\text{test}} &= y[\lfloor N \times 0.8 \rfloor :]\end{aligned}$$

- Create the Neural Network Model

$$\begin{aligned}&\text{Layer 1: } Z^{[1]} = \mathcal{R}(W^{[1]} \cdot X + b^{[1]})\\ &\text{Layer 2: } Z^{[2]} = \mathcal{R}(W^{[2]} \cdot Z^{[1]} + b^{[2]})\\ &\text{Output Layer: } \hat{y} = W^{[3]} \cdot Z^{[2]} + b^{[3]}\end{aligned}$$

end while

Step 3: Return the Trained Model $\hat{y}_i$

Where:

- $X \in \mathbb{R}^{N\times L}$ represents the input data matrix (memory of the output z), where N is the number of samples and L is the length of the memory,
- $W^{(1)} \in \mathbb{R}^{L\times 64}$ and $b^{(1)} \in \mathbb{R}^{64}$ are the weights and biases of the first layer,
- $W^{(2)} \in \mathbb{R}^{64\times 64}$ and $b^{(2)} \in \mathbb{R}^{64}$ are the weights and biases of the second layer,
- $W^{(3)} \in \mathbb{R}^{64\times 1}$ and $b^{(3)} \in \mathbb{R}$ are the weights and bias of the output layer.

In our study, we depend on many basic theorems that offer the theoretical basis for our methodology. These theorems prove the successful convergence of neural networks to an optimum solution given particular constraints of regularity on the data and network design, as well as their universal approximation capabilities for any continuous function over a compact interval. Additionally, they assert that the sum of independent and identically distributed random variables converges in distribution to a normal distribution as the number of samples rises. These theorems are especially significant for modeling situations involving noise since they give insights into the distribution of noise introduced to the signal in a model. By using these core theorems, we provide the theoretical basis of our study and highlight both the possibilities and limits of neural network-based modeling in the context of modeling Hammerstein systems.

Theorem 2 (universal approximation theorem [34], [35]): Let $f : [a, b] \to \mathbb{R}$ be a continuous function on a compact interval $[a, b]$. Then, for any $\epsilon > 0$, there exists a neural network with one or more hidden layers and nonlinear activation functions, such as $\mathcal{R}(x) = \max(0, x)$, capable of approximating f with arbitrary precision. Formally, for any continuous f on $[a, b]$ and $\epsilon > 0$, there exists a neural network $\hat{f}$ such that:

$$\|f - \hat{f}\|_\infty < \epsilon$$

where $\|\cdot\|_\infty$ represents the infinity norm over the interval $[a, b]$.

Theorem 3 (central limit theorem [36]): the sum of independent and identically distributed random variables, such as noise added to the signal in a model, approximately follows a normal distribution as the number of samples increases. Formally, for independent and identically distributed random variables y_i with mean $\mathbb{E}[y_i]$ and variance σ^2, the standardized sum $\frac{1}{\sqrt{N}} \sum_{i=1}^{N} (y_i - \mathbb{E}[y_i])$ converges in distribution to a normal distribution $\mathcal{N}(0, \sigma^2)$ as $N \to \infty$.

4. SIMULATION AND RESULTS

In this section, we performed Monte Carlo simulations to evaluate the efficiency of our machine learning-based method for estimating nonlinear systems. The linear channel was characterized by the impulse response vector H_{BRAN}, which contains the values in Table 1. The BRAN radio channel impulse response is described as:

$$h(n) = \sum_{i=0}^{p-1} H_i \delta(n - \tau_i),\ p = 18. \tag{5}$$

where $\delta(n)$, τ_i and $H_i \in N(0, 1)$ denote, the Dirac function, the paths i time delay and path i magnitude respectively.

Table 1. Delay and magnitudes of 18 targets of BRAN A channel

Delay $\tau_i(ns)$	Magnitudes $H_i(dB)$	Delay $\tau_i(ns)$	Magnitudes $H_i(dB)$
0	0	90	-7.8
10	-0.9	110	-4.7
20	-1.7	140	-7.3
30	-2.6	170	-9.9
40	-3.5	200	-12.5
50	-4.3	240	-13.7
60	-5.2	290	-18
70	-6.1	340	-22.4
80	-6.9	390	-26.7

Afterward, the data was partitioned into distinct training and testing sets. The model of neural networks was created to predict the output of the nonlinear system. This model is composed of two hidden layers, with each layer containing 64 neurons. The model employs the ReLU activation function. In addition, the output layer is a dense layer with a single unit. The model was trained on the training data and measured on the data from the tests using the MSE as the selection criteria, explained as:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i - y_i)^2 \tag{6}$$

here, $\hat{y}_i$ represents the model's prediction for the i-th data point, while y_i corresponds to the true value. This performance metric allows us to quantify the discrepancy between our predictions and the actual values of the nonlinear system. The input sequence, depicted in Figure 3, was randomly generated using a uniform distribution over the range $[-5, 5]$. This sequence corresponds to a dataset of size $N = 1024$. The resulting output scenario, shown in Figure 4, is based on this input sequence and illustrates the generated scenario.

4.1. Estimating the nonlinear system output

The three figures illustrate the prediction of the output of Hammerstein system for varying sizes of input datasets (N): 512, 1024, and 2048. Each figure exhibits two separate curves: the blue curve represents the real output of the system, while the orange curve shows the output approximated by our neural network model. From the analysis of these figures, it is clear that the estimated result of the Hammerstein system differs depending on the data length. When N is equal to 512 as Figure 5, there may be some minor differences, but the estimated performance seems to be rather accurate since the orange curve nearly follows the blue curve.

When the value of N is increased to 1024 as Figure 6, there is a small but noticeable improvement in the estimate, resulting in improved alignment between the two curves. Finally, when N equals 2048 as Figure 7, the estimate seems to be the most accurate, since the orange curve closely aligns with the blue curve. This indicates an improved capacity to make accurate predictions as the amount of the input dataset is increased.

This gain in estimate accuracy with an increase in N may be due to numerous variables, including the influence of the optimization technique utilized, such as gradient descent, as well as the complexity involved with training the neural network model. Gradient descent is performed to minimize the loss of function during model training. With a higher N, the model has a connection to more training data, which may allow more efficient exploration of the parameter space and convergence towards a more global minimum of the loss function. Consequently, the model can better represent the complex communications between the input and output data, resulting in a more accurate assessment of the Hammerstein system's output. However, it's worth mentioning that raising N may also contribute to an increase in computing complexity involved with training the neural network model. A bigger dataset demands more computer resources to analyze and train the model, which might increase training time and computational power requirements. Nevertheless, despite this extra complexity, the gains in terms of estimating accuracy acquired with a bigger dataset frequently outweigh this additional cost in terms of resources.

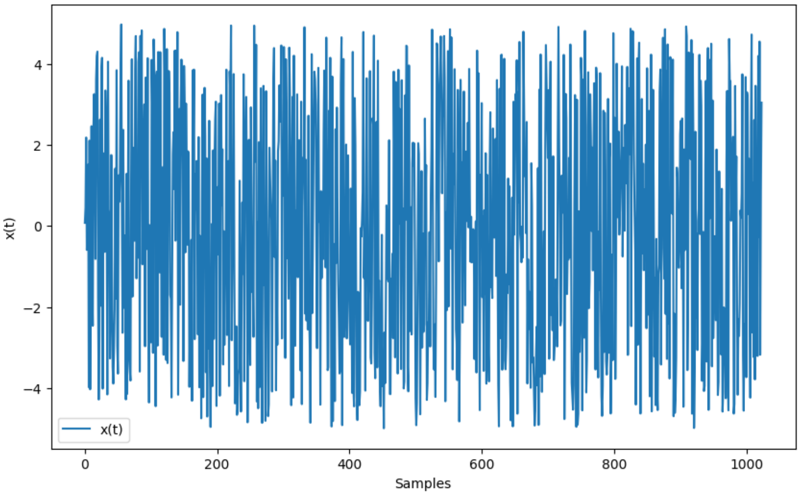


Figure 3. The input sequence, with a data length of $N = 1024$

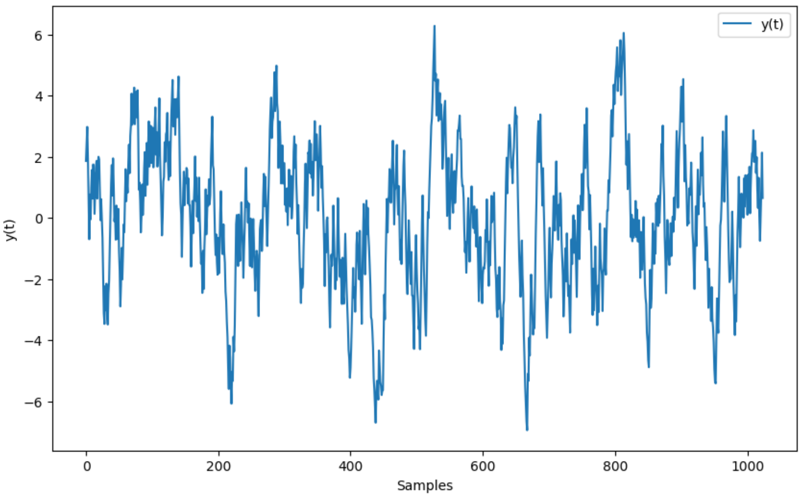


Figure 4. The output data with $SNR = 30$ and $N = 1024$

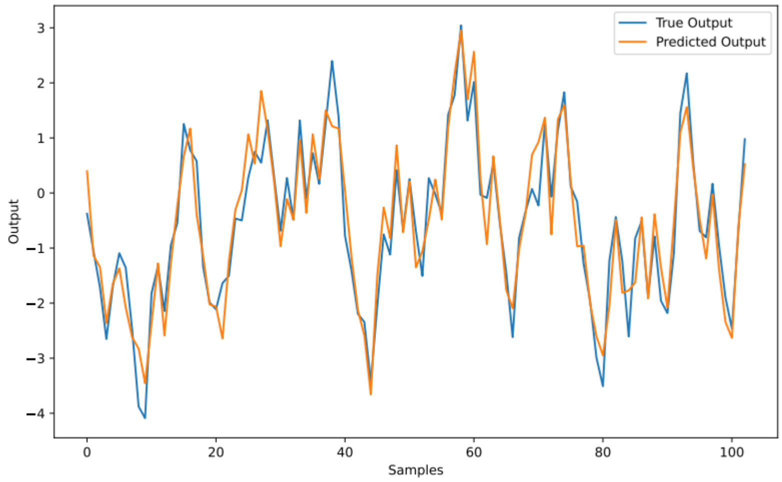


Figure 5. Estimation of the Hammerstein system output for a data length of $N = 512$

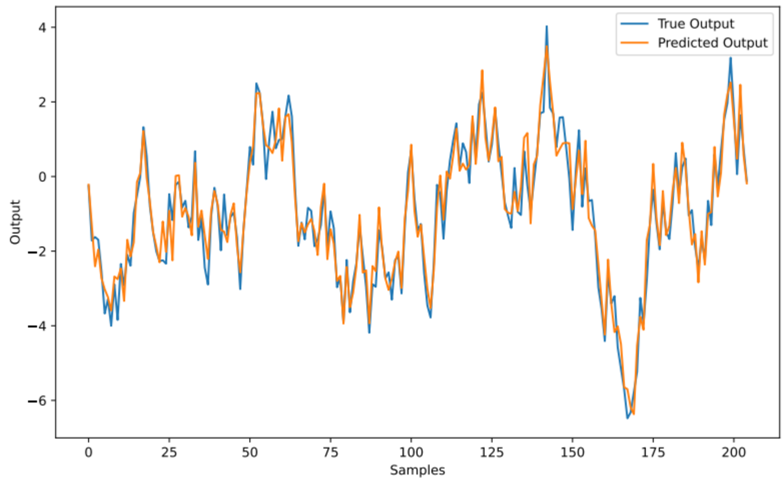


Figure 6. Estimation of the Hammerstein system output for a data length of $N = 1024$

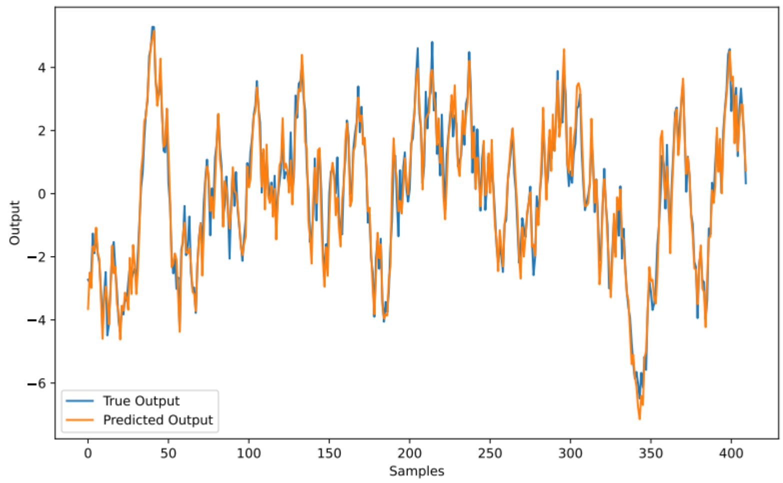


Figure 7. Estimation of the Hammerstein system output for a data length of $N = 2048$

4.2. Performance in noisy environment

Figure 8 illustrates the change in the model's MSE as the value of N varies over the training epochs. The MSE curve demonstrates the variation in error when the model is trained on the data for each N number (512, 1024, and 2048). Firstly, while reviewing the MSE curves, a general tendency is observed: the error lowers as the size of the dataset rises. For $N = 512$, the starting MSE is 3.5, steadily reducing to 0.1031 at the conclusion of training. For $N = 1024$, the initial MSE is 3.5, followed by a continuous drop to 0.0451. Lastly, for $N = 2048$, the starting MSE is 2.3, and it reduces to 0.0233 at the conclusion of training. Secondly, comparing the curves for various values of N shows changes in the convergence rate and the final degree of error. Particularly, with a bigger N, such as $N = 2048$, the final MSE is often lower, showing superior performance of the model with a larger dataset. Finally, the study of the MSE curves also enables for the discovery of any overfitting or underfitting concerns. If the MSE on the test set grows while that on the training set falls, it signals overfitting of the model. Conversely, if the MSE stays high or declines very slowly, it may suggest underfitting.

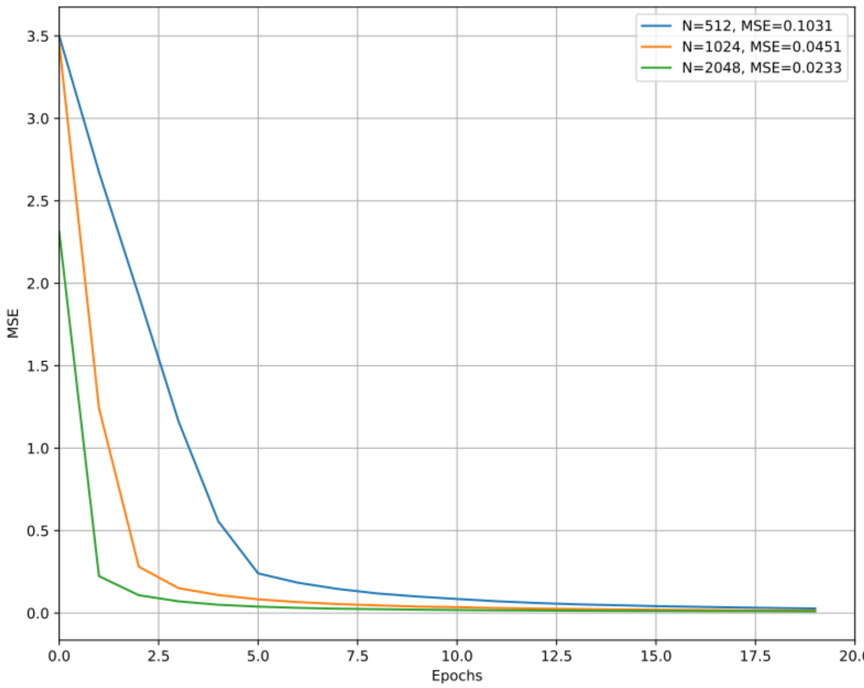


Figure 8. The MSE loss versus the epoch number for different data lengths N

5. CONCLUSION

This paper thoroughly investigates the challenge of identifying nonlinear systems, specifically focusing on Hammerstein systems. The core methodology employs ANN while incorporating gaussian noise into the output. The paper provides a comprehensive evaluation of the ANN algorithm's effectiveness and substantiates its claims through a practical demonstration using a numerical example. For instance, the numerical example highlights the remarkable performance of the approach, with notably low MSE values. Particularly noteworthy is the case with $N = 2048$ samples, where the MSE decreases impressively to 0.02 at the final iteration. Additionally, the alignment between the performance metrics obtained through simulation and empirical observations underscores the robustness and reliability of the model in accurately capturing the system dynamics. Looking ahead, the paper outlines plans for future research, which entail exploring a variety of neural network architectures and integrating regularization techniques to enhance model generalization, especially in the context of complex nonlinear systems. This forthcoming investigation will have a specific focus on applications within the telecommunications domain.

REFERENCES

[1] J. Wang, X. Ma, J. Zhang, and X. Zhao, "Impacts of digital technology on energy sustainability: China case study," *Applied Energy*, vol. 323, p. 119329, Oct. 2022, doi: 10.1016/j.apenergy.2022.119329.

[2] Y. Teekaraman, H. Manoharan, A. R. Basha, and A. Manoharan, "Hybrid optimization algorithms for resource allocation in heterogeneous cognitive radio networks," *Neural Processing Letters*, vol. 55, no. 4, pp. 3813–3826, 2023, doi: 10.1007/s11063-020-10255-2.

[3] M. Lakshmanan, P. S. Mallick, and L. Nithyanandan, "Joint timing synchronization and channel estimation using perfect sequence in uplink time domain synchronous OFDMA," *International Journal of Electronics and Telecommunications*, vol. 62, no. 3, pp. 225–229, 2016, doi: 10.1515/eletel-2016-0030.

[4] Y. Li, S. Jiao, and B. Geng, "Refined composite multiscale fluctuation-based dispersion Lempel–Ziv complexity for signal analysis," *ISA Transactions*, vol. 133, pp. 273–284, 2023, doi: 10.1016/j.isatra.2022.06.040.

[5] H. Hollien and W. Majewski, "Speaker identification by long-term spectra under normal and distorted speech conditions," *The Journal of the Acoustical Society of America*, vol. 62, no. 4, pp. 975–980, 1977, doi: 10.1121/1.381592.

[6] R. Fateh, A. Darif, A. Boumezzough, S. Safi, and M. Frikel, "A novel Kernel algorithm for finite impulse response channel identification," *Journal of Telecommunications and Information Technology*, no. 2, pp. 84–93, Jun. 2023, doi: 10.26636/jtit.2023.169823.

[7] R. Fateh, A. Darif, and S. Safi, "Performance evaluation of MC-CDMA systems with single user detection technique using Kernel and linear adaptive method," *Journal of Telecommunictions and Information Technology*, no. 4, pp. 1–11, Dec. 2021, doi: 10.26636/jtit.2021.151621.

[8] C. J. Lynch, H. Kavak, R. Gore, and D. Vernon-Bido, "Identifying unexpected behaviors of agent-based models through spatial plots and heat maps," in *Understanding Complex Systems*, 2019, pp. 129–142.

[9] C. Liu, L. Zhu, and M. Belkin, "Loss landscapes and optimization in over-parameterized non-linear systems and neural networks," *Applied and Computational Harmonic Analysis*, vol. 59, pp. 85–116, 2022, doi: 10.1016/j.acha.2021.12.009.

[10] A. Perrusquía and W. Yu, "Identification and optimal control of nonlinear systems using recurrent neural networks and reinforcement learning: an overview," *Neurocomputing*, vol. 438, pp. 145–154, 2021, doi: 10.1016/j.neucom.2021.01.096.

[11] R. Fateh, A. Darif, and S. Safi, "The influence of gaussian kernel width on indoor and outdoor radio channels identification from binary output measurements," *International Journal of Information and Communication Technology*, vol. 23, no. 4, pp. 327–345, 2023, doi: 10.1504/IJICT.2023.134853.

[12] L. D. P. Ospina, V. U. Salazar, and D. P. Ospina, "Dynamic equivalents of nonlinear active distribution networks based on hammerstein-wiener models: an application for long-term power system phenomena," *IEEE Transactions on Power Systems*, vol. 37, no. 6, pp. 4286–4296, 2022, doi: 10.1109/TPWRS.2022.3153117.

[13] L. Li, F. Wang, J. Zhang, and X. Liu, "Hammerstein system identification using robust estimator based on quantized observation," in *2023 IEEE 12th Data Driven Control and Learning Systems Conference (DDCLS)*, May 2023, pp. 95–99, doi: 10.1109/DDCLS58216.2023.10166678.

[14] Q. Zha, F. Li, and R. Liu, "Identification of the Hammerstein nonlinear system with noisy output measurements," *Control Theory and Technology*, vol. 22, no. 2, pp. 203–212, May 2024, doi: 10.1007/s11768-024-00196-9.

[15] R. Linggard, D. Myers, and C. Nightingale, *Neural networks for vision, speech and natural language*. Springer Science Business Media, 2012.

[16] M. Rezaeianzadeh, H. Tabari, A. Arabi Yazdi, S. Isik, and L. Kalin, "Flood flow forecasting using ANN, ANFIS and regression models," *Neural Computing and Applications*, vol. 25, no. 1, pp. 25–37, 2014, doi: 10.1007/s00521-013-1443-6.

[17] J. Schmidhuber, "Deep learning in neural networks: an overview," *Neural networks*, vol. 61, pp. 85–117, 2015, doi: 10.1016/j.neunet.2014.09.003.

[18] Khan *et al.*, "Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks.," *Nature Medicine*, vol. 7, no. 6, pp. 673–679, 2001.

[19] M. Farzaneh-Gord, B. Mohseni-Gharyehsafa, A. Arabkoohsar, M. H. Ahmadi, and M. A. Sheremet, "Precise prediction of biogas thermodynamic properties by using ANN algorithm," *Renewable Energy*, vol. 147, pp. 179–191, 2020, doi: 10.1016/j.renene.2019.08.112.

[20] B. Maleki, M. Ghazvini, M. H. Ahmadi, H. Maddah, and S. Shamshirband, "Moisture estimation in cabinet dryers with thin-layer relationships using a genetic algorithm and neural network," *Mathematics*, vol. 7, no. 11, p. 1042, Nov. 2019, doi: 10.3390/math7111042.

[21] S. A. Vahora and N. C. Chauhan, "Deep neural network model for group activity recognition using contextual relationship," *Engineering Science and Technology, an International Journal*, vol. 22, no. 1, pp. 47–54, 2019, doi: 10.1016/j.jestch.2018.08.010.

[22] J. Jagannath, N. Polosky, A. Jagannath, Francesco Restuccia, and T. Melodia, "Neural networks for signal intelligence: theory and practice," *Machine Learning for Future Wireless Communications*, pp. 243–264, 2020, doi: 10.1002/9781119562306.ch13.

[23] D. Auge, J. Hille, E. Mueller, and A. Knoll, "A survey of encoding techniques for signal processing in spiking neural networks," *Neural Processing Letters*, vol. 53, no. 6, pp. 4693–4710, 2021, doi: 10.1007/s11063-021-10562-2.

[24] P. Freire, S. Srivallapanondh, A. Napoli, J. E. Prilepsky, and S. K. Turitsyn, "Computational complexity evaluation of neural network applications in signal processing," *arXiv preprint arXiv:2206.12191*, Jun. 2022, [Online]. Available: http://arxiv.org/abs/2206.12191.

[25] S. Behera, R. Misra, and A. Sillitti, "Multiscale deep bidirectional gated recurrent neural networks based prognostic method for complex non-linear degradation systems," *Information Sciences*, vol. 554, pp. 120–144, 2021, doi: 10.1016/j.ins.2020.12.032.

[26] L. Wu *et al.*, "Artificial neural network based path loss prediction for wireless communication network," *IEEE Access*, vol. 8, pp. 199523–199538, 2020, doi: 10.1109/ACCESS.2020.3035209.

[27] N. Daldal, M. Nour, and K. Polat, "A novel demodulation structure for quadrate modulation signals using the segmentary neural network modelling," *Applied Acoustics*, vol. 164, p. 107251, Jul. 2020, doi: 10.1016/j.apacoust.2020.107251.

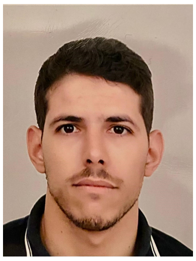
[28] M. Khani, M. Alizadeh, J. Hoydis, and P. Fleming, "Adaptive neural signal detection for massive MIMO," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5635–5648, 2020, doi: 10.1109/TWC.2020.2996144.

[29] C. K. A. Reddy *et al.*, "The INTERSPEECH 2020 deep noise suppression challenge: datasets, subjective testing framework, and challenge results," in *Interspeech 2020*, Oct. 2020, vol. 2020-Octob, pp. 2492–2496, doi: 10.21437/Interspeech.2020-3038.
[30] W. Jiang and H. D. Schotten, "Deep learning for fading channel prediction," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 320–332, 2020, doi: 10.1109/OJCOMS.2020.2982513.
[31] "Broadband radio access networks (BRAN); HIPERLAN type 2; requirements and architectures for wireless broadband access," European Telecommunications Standards Institute (ETSI): Sophia Antipolis, France, 1999.
[32] "Broadband radio access networks (BRAN); HIPERLAN type 2; physical layer," European Telecommunications Standards Institute (ETSI): Sophia Antipolis, France, 2001.
[33] J. Sum, C. S. Leung, and K. Ho, "Convergence analyses on on-line weight noise injection-based training algorithms for MLPs," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 11, pp. 1827–1840, 2012, doi: 10.1109/TNNLS.2012.2210243.
[34] B. Csáji, "Approximation with artificial neural networks," MSc Thesis, Faculty of Sciences Eötvös Loránd University Hungary, 2001.
[35] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of control, signals and systems*, vol. 2, no. 4, pp. 303–314, 1989.
[36] J. Sirignano and K. Spiliopoulos, "Mean field analysis of neural networks: a central limit theorem," *Stochastic Processes and their Applications*, vol. 130, no. 3, pp. 1820–1852, 2020, doi: 10.1016/j.spa.2019.06.003.

BIOGRAPHIES OF AUTHORS

Rachid Fateh received his B.Sc. in Mathematics, Computer Science, and Applications from Ibn Zohr University, Ouarzazate, Morocco, in 2017. He pursued an M.Sc. in Telecommunications Systems and Computer Networks at Sultan Moulay Slimane University, Beni Mellal, Morocco, graduating in 2019. He is currently completing his Ph.D. in Information Technology and Telecommunications at Sultan Moulay Slimane University, specializing in Computer Science and Telecommunications. In 2024, he held the position of temporary teaching and research associate at IUT Grand Ouest Normandie in Caen, France, with expertise in Computer Engineering, Automation, and Signal Processing. Since September 2024, he has been an associate professor at the Faculty of Applied Sciences, Ibn Zohr University, Morocco. His research interests include system identification, equalization of communication channels using reproducing Kernel Hilbert spaces, as well as signal processing and estimation. He can be contacted at email: fateh.smi@gmail.com.

Hassan Ougraz was born in Kelaat Mgouna, Tinghir, Morocco, in 1997. He received the B.Sc. in Mathematics, Computer Science and Applications from Polydisciplinary Faculty of Ouarzazate, Ibn Zohr University (Agadir, Morocco) in 2018, and the M.Sc. in Telecommunications Systems and Computer Networks from Polydisciplinary Faculty of Beni Mellal, Sultan Moulay Slimane University (Beni Mellal, Morocco) in 2020. He is currently working towards a Ph.D. in Computer Science and Mathematics from Polydisciplinary Faculty of Beni Mellal, Sultan Moulay Slimane University. His research interests include array signal processing, wideband DOA estimation, DOA estimation, and wireless communication. He can be contacted at email: hassan.ougraz.24@gmail.com.



Youness Belaadel was born in Ait Faska, Ait Ourir, Morocco, in 1997. He received his B.Sc. in Electronics from Semlalia Faculty of Marrakech, Cadi Ayad University (Marrakech, Morocco) in 2018, and his M.Sc. in Telecommunications Systems and Computer Networks from the Polydisciplinary Faculty of Beni Mellal, Sultan Moulay Slimane University (Beni Mellal, Morocco) in 2020. He is currently working towards a Ph.D. in Computer Science and Mathematics from the Faculty of Science and Technology of Beni Mellal, Sultan Moulay Slimane University. His research interests include the security of IoT and the exploitation of artificial intelligence in this field. He can be contacted at email: younesbelaadel49@gmail.com.

Mathieu Pouliquen received the engineering degree from the Ecole Nationale Superieure d'Ingenieurs de Caen, France, in 2000, then the Ph.D. in Automatic Control from the University of Caen, in 2003. He was assistant professor at the University of Caen between 2004 and 2021. Since September 2021, he is full professor. His research interests include linear and nonlinear system identification. He can be contacted at email: mathieu.pouliquen@unicaen.fr.

Miloud Frikel is an associate professor at National Graduate School of Engineering and Research Center (ENSICAEN), and head of SATE's Department (Embedded Systems and Control), and he is the deputy-director of the Systems Engineering Lab of Normandy (LIS). Since 2008, he is an associate professor at National Graduate School of Engineering and Research Center (ENSICAEN), and head of SATE's Department. From 2003 to 2008, he was with the R&T (Networks and Telecommunications) Department of the IUT of Caen (Normandy University). From 1998 to 2000, he was with the Signal Processing Lab, Institute for Systems and Robotics, Institute Superior Tecnico, Lisbon, as a researcher in the field of wireless location and statistical array processing. And from 2000 to 2003, he worked in the Institute for Circuit and Signal Processing of the Technical University of Munich, Germany. He is member of German Foundation : Alexander von Humboldt Stiftung. He can be contacted at email: miloud.frikel@ensicaen.fr.

Said Safi received his B.Sc. in Physics (option: Electronics) from Cadi Ayyad University, Marrakech, Morocco in 1995, M.Sc. and Ph.D. degree from Chouaib Doukkali University and Cadi Ayyad University, in 1997 and 2002, respectively. He was a professor of information theory and telecommunication systems at the National School for Applied Sciences, Tangier, Morocco, from 2003 to 2005. Since 2006, he has been a professor of applied mathematics and programming at the Polydisciplinary Faculty, Sultan Moulay Slimane University, Beni Mellal, Morocco. In 2008, he received a Ph.D. in Telecommunication and Informatics from Cadi Ayyad University. In 2015, he received the degree of professor in Sciences at Sultan Moulay Slimane University. His general interests span the areas of communications and signal processing, estimation, time series analysis, and system identification subjects. His current research topics focus on transmitter and receiver diversity techniques for single and multi-user fading communication channels, and wide band wireless communication systems. He can be contacted at email: safi.said@gmail.com.