Fuzzy-Integrated Modular Neural Networks for Accurate Prediction of On-Time-In-Full Supply Chain Performance

Mariem Mrad[1], Mohamed Amine Frikha[2*], Younes Boujelbene[3], Soufiene Ben Othman[4]
Applied College, King Faisal University, 31982 Al-Ahsa, Saudi Arabia,[2,4]
Faculty of Economics and Management of Sfax, Tunisia[1,3]

***Abstract*—Accurate prediction of supply chain performance is essential for improving operational efficiency and enabling proactive, data-driven decision-making under dynamic and uncertain conditions. Conventional forecasting methods often struggle to capture the nonlinear relationships between operational factors and performance outcomes. This paper proposes an improved neural modeling framework for predicting supply chain performance based on artificial neural networks (ANNs). The proposed approach compares a mono-network model (global ANN) with a modular multi-network architecture composed of several local neural models integrated through a fuzzy fusion mechanism. Unlike existing studies that focus on isolated performance metrics, this work targets the prediction of the key composite indicator On-Time-In-Full (OTIF). Simulation experiments conducted on a nonlinear dynamic supply chain system demonstrate that the modular ANN approach achieves a significant reduction in learning error, dropping from 0.0223 in the global model to as low as 0.0004 in local modules. Furthermore, the total training time was reduced from 1631.58 seconds to an average of approximately 311 seconds per module. These results confirm that fuzzy-integrated modular architectures offer superior generalization and computational efficiency for advanced predictive analytics in complex supply chain management (SCM) environments.**

Keywords*—Artificial neural networks; supply chain management; performance prediction; predictive analytics; modular-network modeling; demand forecasting

I. INTRODUCTION

Reliable prediction of supply chain performance has emerged as a key requirement for organizations operating in volatile, complex, and highly interconnected markets, where competitiveness and sustainability depend on proactive decision-making [1]. As noted by Chopra and Meindl [4], as supply chains grow more digitized and globally distributed, their performance exhibits strong non-linearities, multi-criteria interdependencies, and time-varying dynamics. These are characteristics that traditional predictive approaches often cannot capture effectively. Classical techniques such as time-series modeling, econometric regression, and analytical optimization offer important insights but generally assume linear structures, static relationships, or single-output formulations, limiting their applicability to modern, data-rich supply chain environments.

Advances in artificial intelligence, particularly neural network–based methods, have significantly transformed predictive analytics in SCM [2]. Neural networks are capable of approximating highly nonlinear functions and capturing complex temporal dependencies within multidimensional data streams. More recently, research has evolved beyond conventional multilayer perceptron's toward hybrid deep learning architectures, such as CNN-LSTM, GNN-LSTM, and attention-based models, that integrate spatial, temporal, and relational information [3]. These approaches leverage the rich data ecosystems enabled by modern smart manufacturing technologies, such as interconnected sensors, intelligent automation systems, and digital twin frameworks, thereby supporting adaptive, real-time, and data-driven decision-making processes.

The concept of "supply chain performance" has evolved beyond traditional operational and financial metrics. Contemporary frameworks increasingly integrate Environmental, Social, and Governance (ESG) dimensions, promoting sustainable and socially responsible operations. As a result, composite metrics combining service-level indicators such as On-Time-In-Full (OTIF), cost efficiency, and inventory measures with sustainability criteria have emerged as holistic representations of supply chain outcomes [4].

Despite significant progress, several gaps persist. Most models focus on single variables, neglecting the interdependencies among performance dimensions. Furthermore, existing studies commonly employ a single global neural network, which struggles to model distinct operating regimes, such as stable periods versus disruption episodes [5]. Consequently, this study addresses the following central research question: How can a fuzzy-integrated modular neural architecture enhance the predictive accuracy, generalization capability, and computational efficiency of multi-dimensional supply chain KPIs compared to a traditional monolithic neural network model?

In response to these issues, this paper proposes an improved neural modeling framework for simultaneously predicting a vector of core supply chain KPIs, namely OTIF, total cost, and inventory level [6]. The contribution is twofold. First, a modular multi-network architecture is developed, in which specialized neural models learn the behavior of different operational zones. Second, these local models are integrated using a fuzzy fusion mechanism that dynamically weights predictions according to the prevailing system state [7]. This design mirrors the inherent regime-dependent behavior of real supply chains and enables a

*Corresponding author.

more robust approximation of their underlying nonlinear dynamics.

This research aims to:

- Develop artificial neural network models to capture composite supply chain performance through nonlinear system approximation.
- Develop and implement a modular multi-network architecture based on specialized local models that capture zone-specific dynamics.
- Integrate local models using a fuzzy fusion mechanism that adapts to operational conditions.
- Perform a thorough empirical evaluation comparing the single-network (global) model and the proposed multi-network architecture in terms of accuracy, generalization, and training efficiency.

The results demonstrate that the multi-network architecture significantly outperforms the traditional mono-network model, offering a more robust and accurate approach for capturing the dynamic, nonlinear, and multidimensional aspects of supply chain performance. This framework contributes both methodologically and practically to advancing next-generation predictive analytics in complex supply chain environments.

The remainder of this paper is structured as follows. Section II reviews related work on neural network–based prediction methods and composite supply chain performance modeling. Section III presents the system modeling framework and formulates the performance prediction problem. Section IV describes the proposed neural modeling approach, including the single-network and multi-network architectures and the fuzzy fusion mechanism. Section V outlines the experimental setup, simulation results, and comparative analysis of the proposed models. Finally, Section VI concludes the paper and discusses potential directions for future research.

II. Related Work

Neural network applications in SCM have undergone a significant paradigm shift, transitioning from basic feedforward structures to modular and hybrid deep learning systems. This evolution is driven by the need to handle the increasing volume, velocity, and variety of data generated in modern logistics networks [7-8].

A. *Evolution Neural Architecture: fFrom MLP to Hybrid Models*

Early research in supply chain predictive analytics relied heavily on the Multilayer Perceptron (MLP) and standard Backpropagation algorithms [9]. While these models proved effective for simple function approximation, they often failed to capture the long-term temporal dependencies and spatial relationships inherent in global supply networks. To overcome these limitations, recent literature has pivoted toward hybrid architectures [10]. For instance, CNN-LSTM models have emerged as a robust solution, utilizing Convolutional Neural Networks (CNNs) to extract local spatial features while Long Short-Term Memory (LSTM) units handle temporal sequences.

Expanding beyond temporal data, Graph Neural Networks (GNNs) integrated with LSTMs (GNN-LSTM) have gained prominence for their ability to represent supply chains as complex graphs. These models allow for the modeling of multi-hop dependencies and inter-organizational relationships, providing a more accurate reflection of how disruptions propagate through a network [11-12]. Furthermore, the integration of attention mechanisms and transformer-based models has further enhanced the ability of neural systems to focus on critical operational drivers during periods of high volatility.

B. *Evolution of Composite Supply Chain Performance Metrics*

The conceptualization of supply chain performance has evolved from isolated operational indicators toward multidimensional composite performance frameworks. Traditional evaluation models, such as the Supply Chain Operations Reference (SCOR) model and the Balanced Scorecard (BSC), emphasize integration across operational efficiency, financial outcomes, and customer service dimensions [13]. These frameworks provide structured methodologies for measuring service reliability, responsiveness, cost control, and asset utilization.

More recently, Environmental, Social, and Governance (ESG) criteria have been incorporated into supply chain performance assessment. ESG indicators increasingly serve as leading predictors of supply chain resilience, risk exposure, and long-term sustainability [14]. This shift reflects growing recognition that operational performance cannot be evaluated independently of environmental impact, regulatory compliance, and stakeholder expectations.

However, despite the emergence of composite evaluation frameworks, most neural network–based predictive models continue to focus on single performance indicators. The modeling of interdependent KPIs, such as OTIF, total cost, and inventory levels, remains relatively underexplored. This gap highlights the need for multi-output and modular architectures capable of capturing interrelated performance dynamics across heterogeneous operational regimes.

C. *Training Optimization and Feature Engineering*

The effectiveness of any neural predictive model is deeply tied to its training optimization and the quality of input features. While the Levenberg–Marquardt algorithm is widely recognized for its superior convergence speed in nonlinear system identification, it remains susceptible to local optima in highly complex supply chain datasets. Consequently, hybrid optimization techniques, such as combining gradient descent with metaheuristics like Particle Swarm Optimization (PSO) or Genetic Algorithms (GA), have become essential for hyperparameter tuning [15].

Moreover, advanced feature engineering, including the use of autoencoders for dimensionality reduction and automated feature selection, has improved model expressiveness. These strategies allow models to filter noise in high-dimensional data environments, ensuring that the neural network learns the underlying operational dynamics rather than over-fitting to transient fluctuations in demand or lead times.

D. Integration with Industry 4.0 Technologies

The proliferation of Industry 4.0 technologies has transformed the data ecosystem in which neural models operate. The integration of Internet of Things (IoT) sensors, RFID tracking, and blockchain-enabled traceability provides a continuous stream of real-time data [16]. One of the most critical advancements in this domain is the emergence of the Digital Twin (DT).

Digital twins provide virtual replicas of physical supply chain processes, allowing neural models to perform "what-if" simulations and predict future performance under various stress scenarios. This synergy between DTs and modular neural architectures enables a transition from reactive monitoring to proactive, state-dependent decision-making. By leveraging these real-time streams, modular networks can dynamically switch between specialized local models to adapt to the prevailing operational regime of the supply chain.

Despite these advancements, the literature remains heavily skewed toward single-output learning tasks, highlighting the need for multi-output, modular neural architectures capable of predicting interdependent performance indicators.

III. Neural Modeling and Asynthesis

A. Neural Networks

ANNs are powerful computational models widely applied, e.g., in function approximation, pattern recognition, nonlinear system identification, and time-series forecasting [17]. Their ability to learn complex, nonlinear relationships from data makes them particularly suitable for modeling dynamic systems, such as modern supply chains, where performance outcomes, costs, service levels, and inventory depend on multiple interrelated operational drivers.

The development of an ANN model requires a careful architectural and algorithmic design. Different ANN architectures exist, recurrent networks, convolutional networks, and hybrid deep learning structures, but multilayer perceptron's (MLPs) remain the most commonly used for system modeling due to their universal approximation capability. This work employs the MLP architecture due to its flexibility and computational efficiency in modeling the dynamics of composite supply chain performance [18].

A critical step in neural modeling is the selection of an appropriate learning algorithm. Numerous algorithms have been proposed to improve convergence properties, reduce computational effort, and enhance prediction accuracy. The most widely used algorithm for training multilayer networks is the Backpropagation (BP) algorithm [19]. Despite its popularity, conventional BP may suffer from slow convergence and sensitivity to initial weights and learning-rate parameters [20]. To address these limitations, several improvements have been introduced, including gradient-based acceleration techniques [21], online learning schemes to handle nonstationary inputs [22], and second-order algorithms such as the Levenberg–Marquardt (LM) method, which offers faster convergence and has become a preferred choice in many prediction tasks [23].

Prediction tasks [24] increasingly rely on neural networks due to their flexibility and strong approximation capabilities. Consistent with prior research on universal approximation [20-22], this study employs an ANN architecture with a single hidden layer, which is adequate for capturing the nonlinear dynamics of supply chain performance. In this work, the evolution of supply chain performance is represented using a simplified nonlinear dynamic structure in which the next performance value depends on past operational inputs and past performance levels. The system dynamics are described by the following recurrent equation:

$$y(k+1) = F[y(k), y(k-1), \ldots\ldots, y(k-n+1), u(k), u(k-1), \ldots\ldots\ldots\ldots, u(k-m+1)] \quad (1)$$

where,

$y(k)$ is the performance output at time k, representing the selected Key KPI, here defined as OTIF.

$u(k)$ is the input vector of operational drivers, including variables such as customer demand, production rate, and supplier lead time;

n and m denote the orders of the autoregressive and input terms, respectively,

$F(\cdot)$ is an unknown nonlinear mapping learned by the neural network.

The ANN model provides an estimated output vector $y_m(k+1)$, which represents the predicted supply chain performance at time k+1 based on historical inputs and outputs.

In multilayer neural networks, each neuron receives weighted signals from all neurons in the preceding layer. The signals are aggregated via a weighted sum and processed through an activation function to produce the neuron's output. In this study, the hidden neurons employ the standard sigmoid activation function, defined as:

$$f(x) = \frac{1}{1+e^{-x}}$$

The sigmoid function is particularly suitable for nonlinear system modeling because it introduces smooth nonlinearity, allowing the network to learn complex performance relationships in dynamic supply chain environments.

B. Direct Neural Network (DNM)

The Direct Neural Model (DNM) constitutes the baseline architecture used in this study for global supply chain performance prediction. In this study, the term DNM refers to a single, monolithic neural network trained on the complete dataset encompassing all operational regimes. This architecture, also called mono-network neural modeling, seeks to approximate the global nonlinear mapping between operational drivers and the composite performance vector, as depicted in Fig. 1.

The mono-network structure is straightforward to synthesize and implement. However, when the underlying system exhibits heterogeneous dynamics across different operational zones (e.g., stable operations, disruptions, promotional surges), a single global model may struggle to represent local nonlinearities adequately. This motivates the exploration of more flexible and modular architectures.

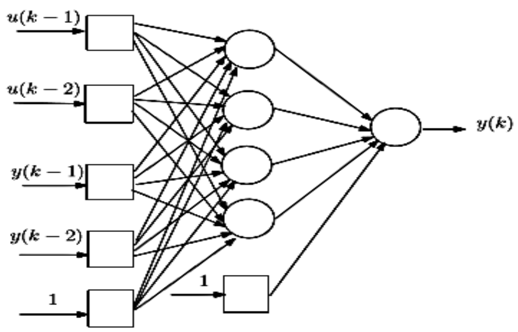

Fig. 1. Structure of the Direct Neural Model (Global ANN).

1) Learning process: ANNs operate as data-driven models whose behavior is shaped entirely through learning. Before training, the network contains no embedded knowledge; all synaptic weights must be iteratively adjusted based on examples drawn from the system. Learning, therefore, consists of modifying the network parameters so that the ANN's output approximates the desired performance response.

In supervised learning, the process requires:

a) Training dataset: A set of examples, each containing an input vector (operational drivers) and an output vector (desired KPIs). The dataset must be sufficiently rich to cover the range of operational conditions, essential for capturing supply chain dynamics influenced by demand variability, lead-time fluctuations, and cost shocks.

b) Cost function: A commonly used measure of the difference between the model output and the actual process output is defined as:$y_m(k+1)y_p(k+1)$

$$E_r = \frac{1}{2}[y_p(k+1) - y_m(k+1)]^2 \quad (2)$$

c) Optimization algorithm: A method for minimizing the cost function by iteratively adjusting the weight parameters.

The majority of ANN training approaches rely on nonlinear optimization, with the objective of reducing the total error over the entire set of examples. Gradient-based methods dominate this class of algorithms, with backpropagation serving as the standard reference technique.

2) The backpropagation algorithm: In an MLP, the estimated output is obtained by propagating the input forward through successive layers. For each neuron v, the local field and activation are computed as:

$$I_v = \sum_j W_{jv} O_j, O_v = f(I_v) \quad (3)$$

where:

$W_{jv}$is the weight connecting neuron $j$to neuron v,

$O_j$is the output of neuron j,

$f(\cdot)$ is the activation function, here, the sigmoid function:

$$f(x) = \frac{1}{1+e^{-x}}, \quad f'(x) = f(x)\big(1 - f(x)\big)$$

For each training example p, the instantaneous error is:

$$E_p(w) = \sum_{i=1}^{m} \frac{1}{2}(S_i^p - Y_i^p)^2 \quad (4)$$

The backpropagation algorithm uses the gradient of the error to update each synaptic weight:

$$W_{uv}(t+1) = W_{uv}(t) - \epsilon(t)\frac{\partial E^p}{\partial W_{uv}} \quad (5)$$

By defining:

$$d_v = \frac{\partial E^p}{\partial I_v}$$

The update rule becomes:

$$W_{uv}(t+1) = W_{uv}(t) - \epsilon(t)\, d_v\, O_u \quad (6)$$

Several factors influence the effectiveness of backpropagation [19]:

- Initialization of weights: Poor initialization may lead to slow or unstable convergence.
- Complexity and representativeness of the training dataset: Supply chain datasets with multiple regimes require careful sampling.
- Network architecture: Particularly the number of hidden neurons, which must balance approximation power and generalization.

3) Generalization: The ultimate goal of training is not merely to memorize training examples but to achieve generalization, meaning the network performs accurately on previously unseen data. Excessive training may lead to overfitting, where the model reproduces training data with high accuracy but performs poorly on new operational scenarios, a particularly hazardous issue in supply chains subject to unpredictable shifts.

Generalization quality depends on:

- Dataset diversity and coverage,
- Training algorithm stability,
- Model complexity,
- Proper use of validation datasets.

In this study, generalization is assessed using separate test data extracted from different operational zones (e.g., high-demand periods, disruptions), which allows us to evaluate how well each model adapts to regime changes.

Traditional mono-network models, such as the DNM, can capture global nonlinearities but often struggle in environments characterized by abrupt regime changes, supply disruptions, or demand surges. As supply chains increasingly exhibit such behaviors, single neural networks become less effective at representing local dynamics.

This limitation motivates the adoption of modular neural architectures, where multiple specialized local models are developed independently and later fused using intelligent aggregation techniques.

The following section introduces the multi-network modeling approach, designed to improve the accuracy, robustness, and interpretability of composite performance predictions in multi-level supply chain structures.

IV. Proposed Approach

Modern supply chain environments exhibit heterogeneous and regime-dependent behaviors driven by factors such as demand volatility, supply disruptions, promotional activities, and inventory imbalances. A single global neural network, as used in the DNM, attempts to approximate the entire nonlinear mapping overall operating conditions. However, such mono-network modeling often suffers from reduced accuracy, slow convergence, and limited generalization when the data distribution varies across operational zones.

In response to these limitations, a multi-network neural architecture is proposed in this study, where the global prediction task is decomposed into multiple local neural models, each specialized in learning the performance dynamics of a specific operational regime. The local models are later combined using a fuzzy fusion mechanism, allowing the final model to adapt smoothly to changing conditions while maintaining high predictive accuracy across all zones.

A. Motivation for Modular Neural Modeling

The characteristics of real-world supply chains provide strong motivation for adopting a modular neural modeling approach. Supply chains exhibit regime-dependent nonlinearities, behaving differently under stable, disrupted, or high-demand conditions. A single global ANN often struggles in this context, underfitting in some regions while overfitting in others. Furthermore, KPIs such as OTIF delivery, cost, and inventory demonstrate differing levels of sensitivity depending on the operational context. For instance, during promotional periods, inventory dynamics predominantly influence predictions, whereas lead-time uncertainties drive performance variance during disruptions, and cost optimization becomes the primary concern under stable conditions. In addition to improving predictive specificity, a modular approach enhances learning efficiency, as training multiple smaller models is computationally faster and can alleviate convergence issues. This specialization also improves generalization within each model's domain, outperforming a single global ANN that must account for the entire input space. Overall, the development of a multi-network system provides both stronger predictive accuracy and greater structural flexibility, making it particularly well-suited for multi-KPI performance forecasting in complex supply chains.

B. Partitioning of the Operational Input Space

In this study, the operational space is partitioned according to the behavior of the performance output $y(k)$, defined as the OTIF level. Since OTIF reflects the combined effect of demand patterns, production scheduling, and supply reliability, it provides a meaningful basis for distinguishing different operational regimes. The first regime corresponds to high-performance conditions, where OTIF remains consistently high, and the system operates with minimal variability. The second regime represents moderate or stable performance, characterized by balanced flows and predictable dynamics, where small fluctuations in OTIF occur but overall reliability is maintained. The third regime captures low-performance or disruption conditions, where OTIF declines due to supply delays, insufficient production capacity, or abrupt changes in demand. Partitioning the output space into these distinct regimes ensures that each local neural model is trained only on data representing the specific performance level it is designed to predict, allowing it to learn the nonlinear dynamics associated with that particular state of the supply chain.

C. Local Neural Models

Each zone is assigned a dedicated local ANN, trained independently using the subset of data relevant to its operational context. Despite being specialized, all local models share key characteristics:

- Architecture: A multilayer perceptron (MLP) with one hidden layer (universal approximator).
- Inputs: Lagged values of operational drivers (demand, lead time, production load) and past KPI values.
- A single performance Indicator:

$$y(k) = \mathrm{OTIF}(k)$$

- Training Strategy: Each model minimizes the prediction error only within its designated zone, resulting in faster convergence and lower modeling complexity.

This decentralization enables each ANN to capture finer nonlinearities that a global model would otherwise oversimplify.

D. Fuzzy Fusion Mechanism

After training the local models, their predictions are combined using a fuzzy validity function, which determines the relevance of each local model for the current system state. Instead of rigid zone boundaries, fuzzy membership functions allow smooth transitions between models, as illustrated in Fig. 2.

For the three local models, the fused output is:

$$y(k+1) = \frac{\sum_{i=1}^{3} \mu_i(k)\, y_i(k+1)}{\sum_{i=1}^{3} \mu_i(k)} \tag{7}$$

where,

- $y_i(k+1)$ is the predicted KPI vector from model i,
- $\mu_i(k)$ is the validity degree of the model i at time k, typically based on Gaussian membership functions that evaluate the proximity of the current input vector to the center of each operational zone.

Advantages of Fuzzy Fusion:

- Smoothly blends local predictions, preventing discontinuities.
- Automatically adapts to regime changes as operational conditions evolve.

- Enhances prediction stability and reduces error propagation in multi-KPI forecasting.

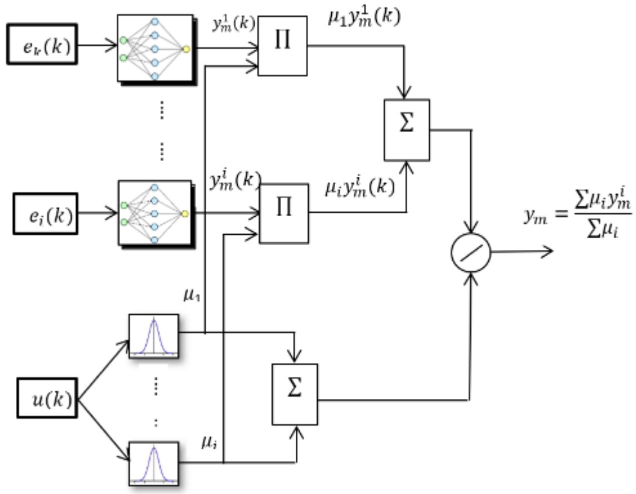

Fig. 2. Architecture of the multi-network neural model.

E. Expected Benefits of the Multi-Network Architecture

The proposed modular framework offers several key advantages over mono-network DNM modeling:

- Higher predictive accuracy: Local specialization reduces approximation error in each operational zone.
- Improved generalization: Local models avoid overfitting to irrelevant data and perform reliably in previously unseen scenarios.
- Reduced computational cost: Smaller models train faster and require fewer iterations to converge.
- Enhanced robustness during disruptions: Fuzzy fusion ensures reliable behavior even when the supply chain switches between regimes.
- Capability for multi-output prediction: The architecture effectively models interdependencies among OTIF, cost, and inventory.

V. Results

This section presents the simulation results using the proposed modeling approach, aimed at evaluating the ability of neural networks to represent the nonlinear dynamics of a dynamic supply chain demand system. An MLP with a single hidden layer and one output was employed. The sigmoid activation function was applied, and the network generated the estimated output $y_m(k+1)$.

A. Nonlinear Dynamic Supply Chain System

Following [1-2], the most influential factors affecting retail demand were identified based on expert evaluations. In this study, the unit sales price, a competitive factor that significantly influences customer purchasing behavior, particularly for independent retailers, is considered as the input variable. The demand quantity, representing customer orders, is treated as the output variable.

The simulation data were generated using the following nonlinear dynamic model:

$$y(k+1) = u(k) + \frac{y(k)}{1 + y^2(k-1) + u(k-1)} \quad (8)$$

where,

- $u(k)$ is the input signal,
- $y(k)$ is the system output,
- The dataset consists of 600 examples with $u(k) \in [0,200]$.

The regression vector used for modeling is:

$$x(k) = [u(k), u(k-1), y(k), y(k-1)]$$

All variables were normalized to match the operating range of the sigmoid neuron.

B. Global Neural Network Model (Mono-Network)

The global neural model was trained using the full input horizon of 600 examples, subdivided into three operating zones of 200 points each. Several experiments were performed to tune the learning parameters. The best configuration was obtained with:

- Learning rate $\varepsilon = 0.2$,
- Four neurons in the hidden layer.
- The backpropagation algorithm was used for training.

To evaluate the generalization capability, the trained network was tested with new input sequences not included in the training set. Three independent test sets of 150 samples each were generated, corresponding to the three operational zones.

Fig. 3 (sub-graphs a, b, and c) shows the evolution of the system output compared with the global neural model output. The results show that the global model generalizes well in Zones 2 and 3, but fails to capture the high nonlinearity in Zone 1.

The results show that the global model generalizes reasonably well in Zones 2 and 3, but not in Zone 1. This indicates that a single global model struggles to represent the nonlinear dynamics across all operating regimes.

C. Multi-Network Neural Modeling

In the second stage, the multi-network approach was implemented. The procedure involves computing the validity degree μ_i for each example, followed by simultaneous adjustment of all local model weights through an iterative training process.

1) Local Models: the Input Sequence Was Partitioned into Three Distinct Operating Zones: $u(k)$

- Zone 1: $u(k)$ fluctuates around 175 (±10%),
- Zone 2: $u(k)$ fluctuates around 35 (±10%),
- Zone 3: $u(k)$ fluctuates around 130 (±10%).

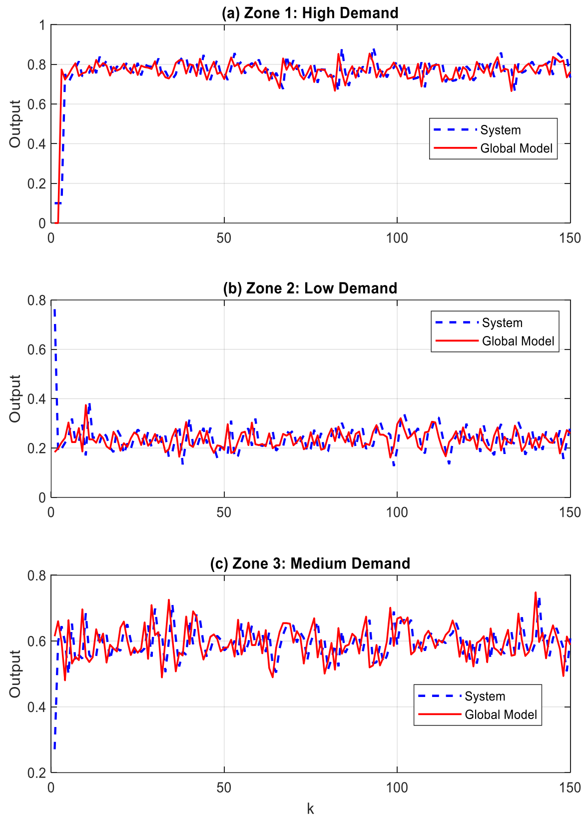

Fig. 3. Evolution of the global model output for the three generalization datasets.

Each subset was used to train a specialized local neural model corresponding to its operational regime. All local models were trained using the nonlinear system defined in (8), a hidden layer comprising three neurons, a learning rate of $\varepsilon = 0.2$, and the standard backpropagation algorithm.

Fig. 4 (sub-graphs a, b, and c) presents the generalization performance of the three local models. These local models exhibit excellent generalization performance in their respective operating zones, confirming the relevance of regime-based decomposition.

The local models exhibit excellent generalization performance in their respective operating zones, confirming the relevance of regime-based decomposition, as shown in Fig. 4.

2) Fused Multi-Network Model: Implementation Conditions for the Fused Model Include:

- Inputs covering all 600 examples.,
- Gaussian membership functions for fuzzy interpolation.

The fuzzy rules are defined as:

- If $u(k)$is A_1, then $y_1(k+1) = MN_1[y(k-1), u(k-1)]$
- If $u(k)$is A_2, then $y_2(k+1) = MN_2[y(k-1), u(k-1)]$
- If $u(k)$is A_3, then $y_3(k+1) = MN_3[y(k-1), u(k-1)]$

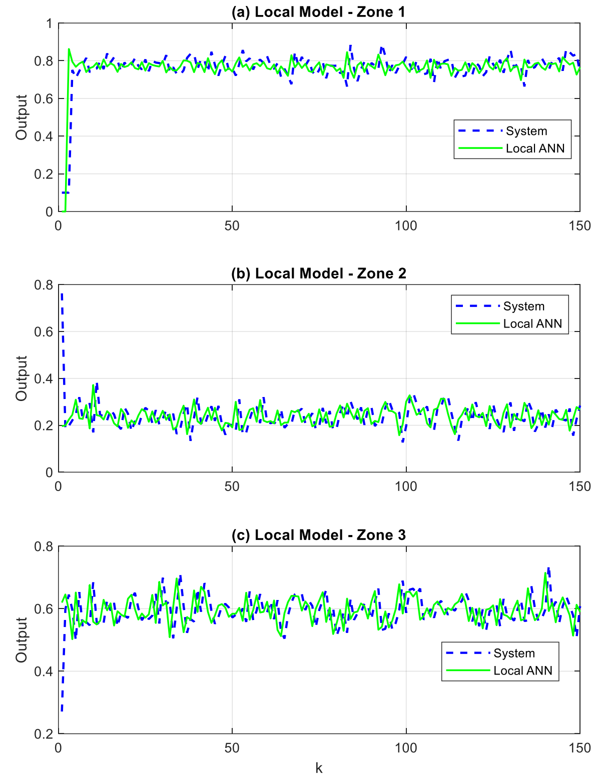

Fig. 4. Generalization results of the three local models.

The membership functions A_i are Gaussian functions centered at 180, 40, and 120, respectively:

$$\mu_i = e^{-\frac{(u(k)-c_i)^2}{2\sigma^2}} \text{with } \sigma = 30$$

The fused output is computed as:

$$y(k+1) = \frac{\sum_{i=1}^{3} \mu_i \, y_i(k+1)}{\sum_{i=1}^{3} \mu_i} \tag{10}$$

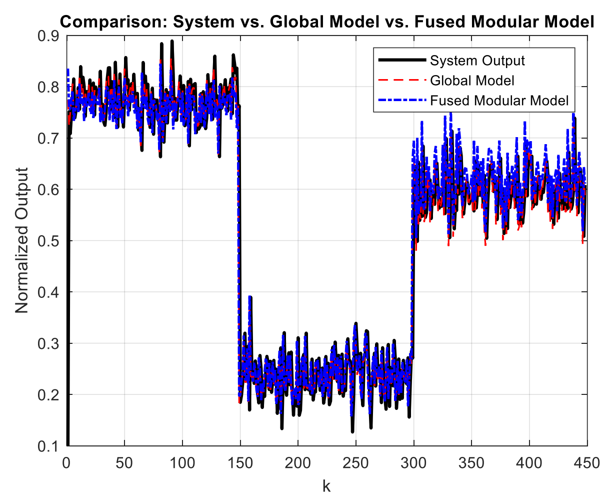

Fig. 5. System output, fused model output, and global model output.

Fig. 5 illustrates the fused model output compared with the system and the global model. The results in Table I show that the modular structure substantially improves accuracy and reduces training time. While hybrid state-of-the-art models like CNN-LSTM [10] or GNN-LSTM [11] provide high accuracy for time-series, they often require significantly more computational overhead and data for training compared to the proposed modular fuzzy-integrated architecture, which offers a balance of simplicity and regime-specific precision.

D. Comparison Between Global and Fused Models

The fused modular model consistently outperforms the global ANN, especially in the high-volatility operational regime, where nonlinear dynamics are more significant. In contrast to the global architecture, every local model demonstrates perfect generalization within its validity region.

The learning errors, test errors, computation time, and number of hidden neurons for the global and local neural models are summarized in Table I. The results show that the modular structure substantially improves accuracy and reduces training time compared to the global ANN.

TABLE I. PERFORMANCE COMPARISON OF GLOBAL AND LOCAL MODELS

Model	Learning Error	Test Error	Learning Time (s)	Hidden Neurons
Global Model	0.0223	0.0246	1631.58	5
First Local Model	0.0005	0.00055	304.62	3
Second Local Model	0.0013	0.0038	296.50	3
Third Local Model	0.0004	0.00049	332.63	3

The results clearly indicate that the total learning time of the fused approach, defined as the sum of the training times of all local models, is significantly lower than that of the global model. Furthermore, each local model achieves substantially lower training and testing errors. Consequently, the multi-network approach provides faster training, improved generalization, and a more accurate representation of nonlinear dynamics.

VI. CONCLUSION

This paper developed and validated a modular artificial neural network framework for predicting supply chain OTIF performance. The empirical results demonstrate that by decomposing a global nonlinear system into specialized local models integrated through fuzzy logic, prediction accuracy is significantly enhanced, reducing learning errors from 0.0223 to as low as 0.0004. Furthermore, the approach drastically optimized computational resources, reducing training time by more than 80% per module compared to the monolithic global network. This study proves that modularity and fuzzy fusion provide a superior balance of generalization and precision, offering a robust tool for decision-makers in complex, multi-regime supply chain environments. Future research will focus on integrating these models into real-time digital twin platforms to enable autonomous supply chain adjustments.

VII. LIMITATIONS AND FUTURE WORK

Despite the significant improvements in predictive accuracy and computational efficiency, this study has several limitations. First, the partitioning of operational zones is currently based on a manually predefined threshold derived from the output variable. While effective, this approach may not fully capture complex, nonlinear operational regimes. Future research should explore automated and data-driven clustering techniques, such as K-means, Gaussian Mixture Models, or Self-Organizing Maps, to dynamically define operational zones in a more adaptive and scalable manner. Second, although the modular framework reduces training time per individual module, maintaining and updating multiple fuzzy-integrated models in a real-time production environment may introduce architectural and computational overhead. Evaluating deployment feasibility in industrial settings, therefore, warrants further investigation. Third, the proposed framework has been benchmarked only against a global ANN baseline. Future studies should conduct comprehensive comparative analyses against state-of-the-art hybrid deep learning architectures, including CNN-LSTM, GNN-LSTM, and transformer-based models, to more rigorously assess relative performance gains. Finally, future work will focus on validating the framework using real-world industrial datasets and exploring the integration of attention-based and transformer architectures within each local module to enhance robustness and predictive stability, particularly during extreme or black-swan events.

FUNDING

This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Grant No. KFU262012].

REFERENCES

[1] M. T. Hagan and M. B. Menhaj, “Training feedforward networks with the Marquardt algorithm,” *IEEE Transactions on Neural Networks*, vol. 5, no. 6, pp. 989–993, 1994.

[2] M. Mrad and Y. Boujelbène, “Demand forecasting in the Tunisian pharmaceutical industry: A comparative study,” *Recent Patents on Biotechnology*, vol. 20, Oct. 2025, doi: 10.2174/0118722083394111250929111803.

[3] M. A. Frikha, and M. Mrad, “AI-enabled demand forecasting, technological capability, and supply chain performance: Empirical evidence from the global logistics sector,” *International Journal of Advanced Computer Science and Applications*, vol. 16, no. 9, Jan. 2025, doi: 10.14569/IJACSA.2025.0160917.

[4] S. Chopra, M. Sodhi, and F. Lücker, “Achieving supply chain efficiency and resilience by using multi-level commons,” *Decision Sciences*, vol. 52, no. 4, pp. 817–832, 2021. DOI: 10.1111/deci.12526.

[5] B. Yabana Kiremit and E. Dikmetaş Yardan, “A comparative study of neuro-fuzzy and neural network models in predicting length of stay in university hospital,” *BMC Health Services Research*, vol. 25, p. 446, 2025, doi: 10.1186/s12913-025-12623-x.

[6] M. Mrad and M. A. Frikha, “Robust control of cyber-physical teleoperation systems for synchronized healthcare supply chain management,” *International Journal of Advanced Computer Science and Applications*, vol. 16, no. 9, Jan. 2025, doi: 10.14569/IJACSA.2025.0160964.

[7] K. S. Narendra and K. Parthasarathy, “Gradient methods for the optimization of dynamical systems containing neural networks,” *IEEE Transactions on Neural Networks*, vol. 2, no. 2, pp. 252–262, 1991.

[8] M. Perau, H. Brings, S. Berninger, **et al.**, “Modular simulation model for operational supply chain management,” *ZWF–Zeitschrift für wirtschaftlichen Fabrikbetrieb*, 2024.

[9] M. Ashraf, A. Eltawil, and I. Ali, "Time-to-recovery prediction in a disrupted three-echelon supply chain using LSTM," *IFAC-PapersOnLine*, vol. 55, no. 13, pp. 1237–1242, 2022.

[10] Y. Xu, F. Li, and A. Asgari, "Prediction and optimization of heating and cooling loads in a residential building based on multi-layer perceptron neural network and different optimization algorithms," *Energy*, vol. 240, p. 122692, 2022, doi: 10.1016/j.energy.2021.122692.

[11] P. Thomas, M.-C. Suhner, and A. Thomas, "CART for supply chain simulation models reduction: Application to a sawmill internal supply chain," in *IFIP Adv. Inf. Commun. Technol.*, vol. 438, pp. 239–246, 2014.

[12] M. Mrad, M. A. Frikha, and Y. Boujelbène, "A comprehensive survey of artificial intelligence and robotics for reducing carbon emissions in supply chain management," Logistics, vol. 9, no. 3, p. 104, 2025, doi: 10.3390/logistics9030104.

[13] Q. Chen and X. Wang, "Supply chain performance evaluation model for integrated circuit industry based on fuzzy analytic hierarchy process and fuzzy neural network," *Journal of Intelligent Systems*, 2025.

[14] Y. Mai, M. H. Callefi, P. Grzona, R. Riedel, and M. Thürer, "Redefining supply chains with additive manufacturing: Insights from network modelling," *IFAC-PapersOnLine*, vol. 59, no. 10, pp. 1874–1879, 2025, doi: 10.1016/j.ifacol.2025.09.315.

[15] E. Badakhshan, N. Mustafee, and R. Bahadori, "Application of simulation and machine learning in supply chain management: A synthesis of the literature using the Sim-ML literature classification framework," Computers & Industrial Engineering, vol. 198, p. 110649, Dec. 2024, doi: 10.1016/j.cie.2024.110649.

[16] D. N. Pathak, R. K. Yadav, and H. Singh, "Analysis and traceability of supply chain management using standardized key performance indicators in the manufacturing unit," *International Research Journal of Multidisciplinary Scope*, 2025.

[17] T. Y. Choi, K. J. Dooley, and J. M. Rungtusanatham, "Supply networks and complex adaptive systems: Control versus emergence," *Journal of Operations Management*, vol. 19, no. 3, pp. 351–366, 2001.

[18] J. M. Swaminathan, S. F. Smith, and N. Sadeh, "Modeling supply chain dynamics: A multi-agent approach," *Decision Sciences*, vol. 29, no. 3, pp. 607–632, 1998.

[19] J. P. C. Kleijnen and R. G. Sargent, "A methodology for fitting and validating metamodels in simulation," *European Journal of Operational Research*, vol. 120, no. 1, pp. 14–29, 2000.

[20] S. Terzi and S. Cavalieri, "Simulation in the supply chain context: A survey," *Computers in Industry*, vol. 53, no. 1, pp. 3–16, 2004.

[21] J. P. C. Kleijnen and M. T. Smits, "Performance metrics in supply chain management," *Journal of the Operational Research Society*, vol. 54, no. 5, pp. 507–514, 2003.

[22] M. Mrad and Y. Boujelbène, "Multilayer perception neural networks method-based supply chain performance prediction by a new augmented SCOR® metrics: SCOR®4.0," in Proc. 2024 IEEE 15th International Colloquium on Logistics and Supply Chain Management (LOGISTIQUA), Sousse, Tunisia, May 2–4, 2024, doi: 10.1109/LOGISTIQUA61063.2024.10571519.

[23] Y. C. Correa-Martinez and M. Seck, "A generic representation of supply network resilience using simulation-based experimentation," Journal of Simulation, 2023.

[24] L. Zhong and Q. Li, "Intelligent diagnosis and prediction of turbine digital electro-hydraulic control system faults: Design and experimentation," PLoS ONE, 2023.