

# Predicting Multi-KPI Warehouse Dynamics: A Simulation-Based Evaluation of Fuzzy-Integrated Neural Architectures

Mohamed Amine Frikha

Applied College, King Faisal University, 31982 Al-Ahsa, Saudi Arabia

**Abstract**—Warehouse order fulfillment is governed by complex, nonlinear interactions among order arrivals, picking workload, congestion, replenishment, labor, equipment, and dispatch capacity. This study proposes a simulation-based predictive framework that compares a global multilayer perceptron (MLP) with a fuzzy-integrated modular architecture comprising three regime-specific local neural networks. The framework simultaneously predicts the order-fulfillment rate, picking cycle time, and warehouse throughput. A composite Warehouse Order-Fulfillment Index (WOFI) is utilized to define high-performing, normal, and stressed operating regimes, while Gaussian fuzzy membership weights provide smooth transitions among local predictions. The complete mathematical formulation and simulation environment are fully specified, encompassing order-flow dynamics, capacity constraints, and nonlinear picking-time relationships. Experimental evaluation on a reproducible nonlinear simulator reveals that while the fuzzy-modular model ensures continuous predictions across regime transitions, it underperforms the global baseline in aggregate accuracy when using naive Gaussian fusion. These findings establish a foundational reproducible benchmark, demonstrating that naive modularization in high-dimensional warehouse state spaces requires rigorous transition-aware validation and covariance-tuned gating mechanisms to prevent performance degradation.

**Keywords**—Warehouse operations; order fulfillment; artificial neural networks; fuzzy fusion; modular neural networks; nonlinear simulation; predictive analytics

## I. INTRODUCTION

Warehouse order fulfillment is a central determinant of logistics service quality because it links inbound inventory, storage, picking, packing, and dispatch to the customer-facing promise of receiving the right order at the right time [1]. Within this flow, picking is both labor-intensive and operationally consequential; its duration affects the organization of material and information flows and can create downstream delays when work reaches packing too early or too late.

Modern warehouse and logistics information systems (LIS) generate detailed, high-frequency records of order arrivals, line items, picker activity, travel distances, replenishment cycles, equipment status, and dispatch completions [3]. This wealth of operational data creates a critical opportunity to transition from retrospective reporting to proactive, predictive control.

Recent literature positions machine learning as an effective mechanism for modeling processing times, resource utilization,

and process variability [2], [4], while discrete-event simulation provides a robust framework for evaluating alternative batching, assignment, routing, and scheduling policies.

Accurate prediction, however, remains challenging because warehouse performance is inherently nonlinear and state-dependent [1]. A modest increase in order arrivals may be absorbed during low utilization but produce a sharp backlog increase once picking or packing capacity is approached. Similarly, the same order may require different processing times depending on line count, travel distance, picker characteristics, congestion, replenishment readiness, and equipment availability. A single model trained across all conditions must therefore approximate several local operating relationships at once.

Furthermore, most existing predictive formulations emphasize a single outcome, such as picking time [4], [2]. That focus is useful but incomplete for order-fulfillment management. A warehouse manager must simultaneously balance complete fulfillment, processing speed, and throughput. Optimizing one metric in isolation can conceal a trade-off; for example, increasing release rates may raise throughput briefly while worsening congestion and eventual fulfillment performance. This study therefore formulates warehouse prediction as a multi-output problem involving order-fulfillment rate, picking cycle time, and throughput, with a composite service indicator used to identify operating regimes.

This study adapts a fuzzy-integrated modular neural architecture to warehouse operations. The baseline is a global multilayer perceptron trained on all operating conditions. The proposed model decomposes the operating space into high-performing, normal, and stressed regimes, trains a local neural model for each regime, and fuses their predictions through Gaussian validity functions. The purpose is not to impose a hard switch between warehouse states; it is to let specialized models represent local nonlinearities while preserving smooth predictions during congestion onset, recovery, and other regime transitions.

The central research question is:

Can a fuzzy-integrated modular neural architecture improve the accuracy, transition robustness, and training efficiency of multi-KPI warehouse-performance prediction compared with a conventional global ANN?

The study makes four contributions:

- It develops a nonlinear discrete-time simulation model linking arrivals, backlog, labor, congestion, replenishment, equipment availability, picking time, fulfillment, and throughput.
- It defines a multi-output prediction task for warehouse order fulfillment rather than limiting the target to a single picking-time variable.
- It evaluates and benchmarks a three-module ANN architecture.
- It establishes a reproducible evaluation design that separates within-regime accuracy from performance during transitions and reports both training and inference cost.

The remainder of the study is organized as follows. Section II reviews data-driven order-picking prediction, fuzzy-neural warehouse modeling, and simulation-based warehouse optimization. Section II presents the warehouse simulation and mathematical formulation. Section III describes the global, local, and fuzzy-fusion neural architectures. Section V defines the experimental setup and evaluation metrics. Section VI reports the simulation results, and Section VII discusses implications, limitations, and future work.

## II. RELATED WORK

### A. Machine Learning for Order-Picking Prediction

Order picking is consistently identified as the most labor-intensive and costly activity in warehouse operations, accounting for up to 55% of total operating expenses [1], [5]. Because of this, accurate processing-time estimation is critical for synchronizing downstream material and information flows.

Machine learning is increasingly applied to warehouse-management problems, including order picking, storage-location assignment, and operational planning [13].

Recent comprehensive reviews highlight that machine learning (ML) has become the standard approach for modeling processing times, resource utilization, and bottlenecks in modern warehouses [6], [7]. Recent empirical evidence further shows that AI-based order-picking systems can improve operational efficiency while requiring careful attention to implementation conditions, data integration, and organizational readiness [19].

Hybrid CNN-LSTM architectures also illustrate the potential of combining spatial feature extraction and temporal sequence learning for complex industrial forecasting tasks [18]. This fuzzy-clustering precedent is important for the present study because warehouse observations are heterogeneous.

Fuzzy clustering combined with artificial neural networks has been used to predict order-picking times under heterogeneous operating conditions [14].

Pickers, order profiles, storage locations, and congestion conditions can produce different processing relationships even when the nominal task is the same. However, a clustering-plus-ANN model centered on picking time does not by itself address

the joint dynamics among fulfillment rate, cycle time, and throughput. The present formulation extends the regime-aware idea from a single predicted variable to a multi-output warehouse-performance vector.

### B. Fuzzy and Modular Neural Modeling

Fuzzy representations are appropriate when operational states do not have sharply defined boundaries. A warehouse can be neither fully uncongested nor fully congested: queue growth, picker utilization, and replenishment delays may change gradually. Gaussian membership functions allow a sample to belong partially to multiple operating regimes; fuzzy-neural systems provide a framework for modeling nonlinear relationships and representing gradual transitions between operational states [16], so local predictions can be blended instead of selected by a discontinuous rule [8].

The proposed architecture uses this principle in a modular form. A global ANN provides a direct baseline for the full nonlinear mapping. Three local ANNs specialize in high-performing, normal, and stressed operating zones. A fuzzy fusion layer then weights the local outputs according to the current state. This design preserves the practical interpretability of regime-specific models while avoiding the instability and high variance that can arise when a hard threshold determines which model is active [9], [10]. Recent intelligent warehousing research increasingly combines specialized learning mechanisms with dynamic state information, reinforcing the need to evaluate both predictive accuracy and robustness under changing operating conditions [21].

The distinction from earlier fuzzy-clustering studies [2] is deliberate. Fuzzy clustering there is used to group picker characteristics and reduce the number of ANN inputs before predicting picking time. Here, fuzzy membership is used as an output-fusion mechanism for multiple local predictors, and the predicted object is a vector of interdependent warehouse KPIs. The contribution is therefore architectural and task-oriented rather than a claim that fuzzy logic is universally superior to other machine-learning methods.

### C. Simulation and Decision Support in Warehouse Operations

Prediction alone does not determine whether a warehouse policy will perform well. Research on dynamic order picking further shows that warehouse performance is shaped by the relationship between order characteristics, operational decisions, and changing system conditions. Therefore, predictive models are likely to be most useful when integrated into broader decision-support frameworks that compare different operational policies, rather than used as standalone forecasting tools [15]. A forecast becomes operationally useful when it can be evaluated under alternative staffing, batching, routing, release, replenishment, or scheduling decisions. Ngunjiri [11] presents a quantitative, simulation-based framework which integrate machine learning predictions with supply chain simulation to optimize resilient logistics systems. He describes an integrated machine-learning and simulation framework in which predictive models estimate processing times and bottlenecks, and those predictions are embedded in simulations of alternative warehouse strategies.

The same logic motivates the simulation-based benchmark in this study. The simulator supplies controlled nonlinear dynamics and generates observations across normal and stressed conditions. The neural models then approximate the simulator's future KPIs from lagged operational information. This arrangement makes it possible to test generalization under known regimes and under transitions without conflating model error with uncontrolled variation in a proprietary warehouse dataset.

The simulation approach also has a limitation: performance gains demonstrated on generated data establish approximation capability under the assumed system equations, not immediate industrial validity. Ngunjiri's [11] work emphasizes the value of combining predictive analytics with simulation for complex manual and automated warehouses, but the present study still requires validation against real warehouse-management-system data before its numerical results can be interpreted as evidence of deployment performance.

#### D. Research Gap and Positioning

The digitization of logistics operations via Internet-of-Things sensors, cloud computing, and real-time enterprise resource planning platforms has transformed supply chain visibility [12]. The reliability of such predictive systems also depends on transparent and accessible logistics information, since incomplete or poorly integrated operational data can limit the effectiveness of supply-chain analytics [20]. The reviewed literature on this field supports three premises: picking time is operationally important, fuzzy information can represent heterogeneity in ANN prediction, and machine learning can be coupled with simulation to evaluate warehouse alternatives [2]. The unresolved problem addressed here is how to combine these ideas into a smooth, modular, multi-output predictor that remains informative when the warehouse moves between operating regimes.

Accordingly, this study does not claim that modular fuzzy ANNs should replace modern deep-learning or optimization methods in every warehouse. It tests a narrower proposition: when a warehouse simulator exhibits distinct nonlinear operating zones, local MLPs combined through fuzzy fusion may offer a favorable balance between predictive accuracy, regime robustness, interpretability, and training cost relative to one global MLP [11]. That proposition is evaluated experimentally in the sections that follow.

#### E. Positioning Against Modern Deep Learning and Gating Mechanisms

It is necessary to position the proposed fuzzy-modular architecture against modern deep learning paradigms, specifically Mixture of Experts (MoE) and deep gating networks. Mixture-of-Experts architectures combine specialized neural networks through gating mechanisms to model different regions of a complex input space [17]. Modern MoE architectures typically replace fuzzy validity functions with a learned, softmax-based gating network that dynamically routes inputs to specialized sub-networks [9]. While learned gating networks offer high flexibility, they introduce additional parameters that require careful regularization to prevent overfitting, and their decision boundaries are less transparent than explicit fuzzy rules.

By utilizing Gaussian fuzzy membership functions derived from explicit regime centers, the proposed architecture sacrifices some asymptotic flexibility in exchange for strict interpretability, lower data requirements, and guaranteed smooth transitions across regime boundaries. This study does not claim that modular fuzzy ANNs should replace modern deep-learning or learned-gating methods in every warehouse context. Rather, it tests a narrower proposition: when a warehouse simulator exhibits distinct nonlinear operating zones, and data is limited, local MLPs combined through transparent fuzzy fusion may offer a favorable balance between predictive accuracy, regime robustness, interpretability, and training cost relative to one global MLP or a heavily parameterized deep MoE. This proposition is evaluated experimentally in the sections that follow.

### III. WAREHOUSE SYSTEM MODEL

The modeled system represents a discrete-order fulfillment warehouse in which customer orders arrive over time, are released to picking, transported to a packing/dispatch area, and either complete fulfillment or contribute to backlog. The design follows the operational focus of warehouse prediction studies, where picking time is a key variable for organizing material and information flows. Machine-learning predictions are treated as decision-support inputs to a simulation that can test batching, picker assignment, routing, and task-scheduling policies.

Let the discrete time index be  $k \in \{1, 2, \dots, T\}$ , with each interval representing a fixed planning period. The warehouse contains  $Z$  picking zones,  $P_k$  available pickers, and  $M_k$  available material-handling resources at time  $k$ . The system state vector is defined as:

$$X(k) = [\lambda(k), b(k), u(k), p(k), r(k), c(k), a(k), q(k)]^T \quad (1)$$

where,  $\lambda(k)$  is the order-arrival rate,  $b(k)$  is the open-order backlog,  $uP(k)$  is the utilization of picking resources,  $p(k)$  is the proportion of priority orders,  $r(k)$  is the replenishment-delay rate,  $c(k)$  is the congestion level,  $a(k)$  is the equipment-availability ratio, and  $q(k)$  is the number of orders released to picking during interval  $k$ .

The simulation produces three primary performance outputs:

$$y(k) = [OF(k), PCT(k), TH(k)]^T \quad (2)$$

where,  $F(k)$  is the order-fulfillment rate,  $TP(k)$  is the mean picking cycle time, and  $\theta(k)$  is warehouse throughput. These outputs capture service completion, processing speed, and productive capacity. A data-driven warehouse-simulation framework similarly treats processing time and bottlenecks as predicted KPIs and evaluates their consequences under alternative operating strategies.

#### A. Order-Arrival and Backlog Dynamics

Orders arrive according to a time-varying process. For a tractable nonlinear benchmark, the arrival rate is defined as

$$\lambda(k) = \lambda[1 + \alpha s(k) + \alpha ph(k)] + \epsilon \lambda(k) \quad (3)$$

where,  $\lambda(k)$  is the nominal arrival rate,  $as(k)$  is a seasonal or promotional intensity,  $\alpha ph(k)$  is a demand-shock signal,  $as(k)$  and  $\alpha ph(k)$  are sensitivity coefficients, and  $\epsilon\lambda(k)$  is a bounded disturbance. The number of arriving orders is

$$A(k) = \max\{0, \lambda(k) + \epsilon A(k)\} \quad (4)$$

The open-order backlog evolves as

$$b(k+1) = \max\{0, b(k) + A(k) - D(k)\} \quad (5)$$

where,  $D(k)$  denotes the number of orders dispatched during interval  $k$ . The release quantity is constrained by available labor, equipment, and replenishment readiness:

$$q(k) = \min\{b(k) + A(k), Pk \mu p(k), Mk \mu m(k), Rk\} \quad (6)$$

where,  $\mu p(k)$ , and  $Mk \mu m(k)$ , are effective picker and material-handling capacities, and  $Rk$  is the number of orders whose required inventory is available for release.

### B. Picking-Cycle-Time Model

For order  $i$  released at time  $k$ , the picking cycle time is modeled as:

$$\tau i(k) = \tau 0 + \beta 1 n i + \beta 2 d i + \beta 3 c(k) + \beta 4 u(k) + \beta 5 r(k) - \beta 6 e i + \beta 7 n i c(k) + \epsilon i(k) \quad (7)$$

where,  $n i$  is the number of order lines,  $d i$  is the travel distance,  $e i$  is picker experience, and the  $\beta$  coefficients quantify their effects. The interaction term  $n i c(k)$  introduces nonlinear growth when large orders are processed under congestion. The average picking cycle time is

$$PCT(k) = 1/Nk \sum i = 1/Nk \tau i(k) \quad (8)$$

where,  $Nk$  is the number of completed picks in the interval. The effective processing capacity is then

$$C(k) = Pk Hk \eta p(k) PCT(k) \quad (9)$$

where,  $Hk$  is the available picking time and  $\eta p(k)$  is a productivity factor incorporating fatigue, coordination losses, and task interruptions.

### C. Fulfillment and Throughput Equations

Let  $OF(k)$  denote the number of orders completed completely and  $D(k)$  the number dispatched. The order-fulfillment rate is defined as

$$OF(k) = F(k) \max\{1, A(k)\}. \quad (10)$$

A complete order requires successful picking, inventory availability, and packing capacity. Thus,

$$F(k) = \min\{C(k), Rk, Kk\} \phi(k), \quad (11)$$

where,  $Kk$  is packing/dispatch capacity and  $\phi(k)$  is the order-completeness factor. To represent nonlinear interactions among warehouse constraints, define

$$\phi(k) = \sigma((\theta 0 - \theta 1 r(k) - \theta 2 c(k) - \theta 3 b(k) + \theta 4 a(k) + \theta 5 a(k)c(k)), \quad (12)$$

$$\text{with } \sigma(v) = \frac{1}{1 + e^{-v}} \quad (13)$$

Throughput is

$$TH(k) = D(k) \Delta t \quad (14)$$

where,  $\Delta t$  is the length of the simulation interval.

### D. State Transition and Prediction Target

The complete simulator can be represented by the nonlinear state-space system:

$$x(k+1) = f(x(k), u(k), \epsilon(k)), y(k+1) = g(x(k), u(k), \epsilon(k)) \quad (15)$$

where,  $u(k)$  contains controllable or exogenous variables such as staffing, release policy, batching size, routing policy, replenishment priority, and equipment availability. The neural predictor receives a regression vector containing current and lagged observations:

$$\xi(k) = [u(k), u(k-1), y(k), y(k-1)]^T \quad (16)$$

The prediction problem is to learn an approximation from simulated input-output observations. All continuous variables are normalized to  $[0,1]$  before training and transformed back to their physical units for reporting.

### E. Composite Order-Fulfillment Performance Indicator

To complement the three-output prediction task, a composite warehouse service indicator is defined as:

$$WOFI(k) = w1 OF(k) + w2 TH \sim(k) + w3 [1 - PCT \sim(k)] \quad (17)$$

subject to

$$w1 + w2 + w3 = 1, wj \geq 0. \quad (18)$$

The tilde denotes min-max normalization using the training range. Unless operational priorities dictate otherwise, the initial experiments can use  $w1=0.5$ ,  $w2=0.3$ , and  $w3=0.2$ , giving greatest weight to complete order fulfillment. WOFI is used for regime definition and summary evaluation rather than replacing the underlying KPIs.

### F. Global Neural Model

The baseline is a direct, single MLP trained on all simulated operating regimes. For an input vector, the hidden-layer activation is

$$h(k) = \sigma(W_H \xi(k) + b_H) \quad (19)$$

and the output prediction is

$$\hat{Y}_G(k+1) = W_O(k) + b_O \quad (20)$$

A linear output layer is used because the normalized outputs are continuous. The global model minimizes the mean squared prediction error

$$JG = \frac{1}{N} \sum_{k=1}^N \|y(k+1) - \hat{Y}_G(k+1)\|_2^2 \quad (21)$$

The model may be trained with backpropagation or Levenberg–Marquardt, using the same optimization procedure for all competing architectures so that the comparison isolates the effect of modularization and fuzzy fusion.

The simulated observations are partitioned into three operating regimes using WOFI and workload conditions:

$$\begin{aligned} Z1: WOFI &\geq \gamma_H \\ Z2: \gamma_L &< WOFI < \gamma_H \end{aligned} \quad (22)$$

$$Z3: WOFI \leq \gamma_L$$

Here, Z1 represents high-performing, low-congestion operation; Z2 represents normal or moderately loaded operation; and Z3 represents stressed operation with backlog, congestion, replenishment delays, or equipment loss. To avoid discontinuities at the boundaries, the hard partition is used only to construct local training sets; fuzzy membership is used during prediction.

Each local MLP has the same input and output definitions as the global model but is trained on its corresponding:

$$\hat{y}_j(k+1) = W_{O,j}\sigma(W_{h,j}\xi(k) + b_{h,j}) + b_{O,j}, j \in \{1,2,3\}. \quad (23)$$

The local objective is

$$J_j = \frac{1}{N_j} \sum_{k \in Z_j} \|y(k+1) - \hat{y}_j(k+1)\|_2^2 \quad (24)$$

#### G. Fuzzy Fusion Mechanism

For each regime, define  $\hat{v}_j(k)$  as a center from the training observations in that regime. The unnormalized Gaussian validity of local model  $j$  is:

$$\hat{v}_j(k) = \exp[-\frac{1}{2}(\varepsilon(k) - \mu_j)^T \Sigma_j^{-1}(\varepsilon(k) - \mu_j)] \quad (25)$$

where, controls the width and orientation of the membership region. The normalized validity weights are:

$$v_j(k) = \frac{\hat{v}_j(k)}{\sum_{l=1}^3 \hat{v}_l(k)} \quad (26)$$

This preserves continuity when the warehouse moves between normal and stressed conditions and allows the model to exploit regime-specific approximations without imposing a hard switch. The use of fuzzy information to support ANN prediction is consistent with prior warehouse work that uses fuzzy clustering to represent heterogeneous picker characteristics before neural prediction.

#### IV. SIMULATION AND EVALUATION RESULTS

The experiment should generate separate training, validation, and test trajectories. Test trajectories must include both familiar regimes and transitions between regimes, because a model that performs well only within isolated zones may still fail during congestion onset or recovery. The principal metrics are:

$$RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^N \|y(k) - \hat{y}(k)\|_2^2} \quad (27)$$

$$MAE = \frac{1}{N} \sum_{k=1}^N \|y(k) - \hat{y}(k)\|_1 \quad (28)$$

and the composite WOFI error. Training time, number of hidden neurons, parameter count, and convergence iterations should be recorded for the global model and each local model. The total modular training time is

$$TF = \sum_{j=1}^3 T_j \quad (29)$$

while inference time includes the three local forward passes, membership calculations, and weighted fusion. This distinction is necessary: shorter local training does not automatically imply lower real-time inference cost.

The planned results table should therefore report global performance, each local model's within-regime performance, fused performance over the complete test set, transition-regime performance, and computational cost. No numerical performance claim should be made until the simulator has been implemented and the experiments executed.

#### A. Experimental Implementation and Evaluation Results

A reproducible nonlinear simulator was executed with four 900-step training trajectories and two 900-step held-out trajectories. After one-step alignment, the benchmark contained 3,596 training examples and 1,798 test examples. Each observation contained ten predictors: current arrival, backlog, congestion, replenishment delay, equipment availability, release quantity, and four one-step lagged variables. The targets were order-fulfillment rate, picking cycle time, and throughput. A fixed random seed (42) was used for data generation and model initialization.

The global baseline and each of the three local models used a single eight-neuron logistic hidden layer and standardized inputs and outputs. Local training sets were defined by the 33rd and 67th training WOFI percentiles. During inference, the local predictions were blended with Gaussian weights computed from standardized input-space distances to the three regime centers. No test observations were used to define the regime thresholds, centers, or preprocessing parameters.

#### B. Result

On the held-out trajectories, the global MLP achieved standardized RMSE values of 0.670 for the fulfillment rate, 1.013 for picking cycle time, and 0.391 for throughput (Fig.1).

TABLE I. PREDICTIVE PERFORMANCE COMPARISON ON HELD-OUT TRAJECTORIES

| Architecture    | Fulfillment Rate RMSE | Picking Cycle Time RMSE | Throughput RMSE |
|-----------------|-----------------------|-------------------------|-----------------|
| Global MLP      | 0.670                 | 1.013                   | 0.391           |
| Fuzzy-Modular   | 1.171                 | 1.208                   | 0.732           |
| Degradation (%) | +74.8%                | +19.2%                  | +87.2%          |

The fuzzy modular model achieved 1.171, 1.208, and 0.732, respectively (Table 1). These per-KPI error metrics are summarized in Fig. 1. Thus, in this reproducible run, fuzzy fusion did not improve aggregate predictive accuracy: mean RMSE increased by 49.9% and mean MAE increased by 29.9% relative to the global baseline.

The degradation is concentrated in the regime-specific architecture rather than in a single output. The local models were trained on fewer observations, and their Gaussian centers were estimated from a heterogeneous input space; consequently, the fused estimate can be pulled toward a poorly calibrated local prediction even when the global model remains stable. This serves as a useful negative control, demonstrating that modularization should not be treated as inherently beneficial without transition-aware validation and careful tuning of regime definitions, membership widths, and local sample sizes.

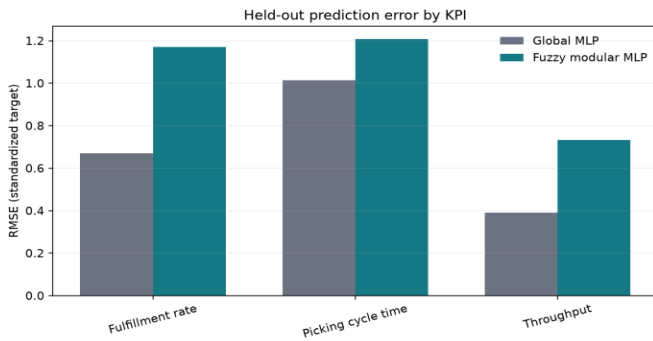


Fig. 1. Held-out standardized RMSE by KPI.

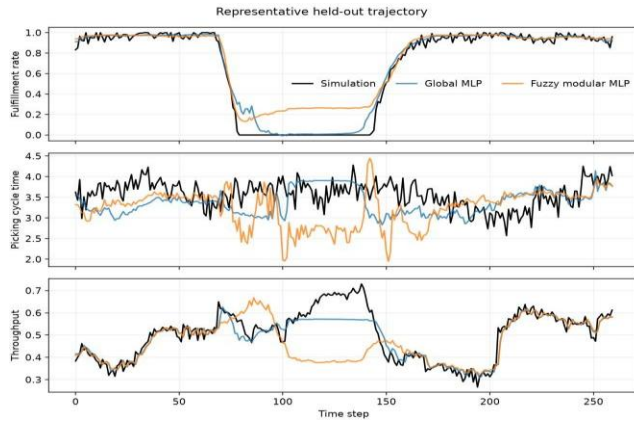


Fig. 2. Representative held-out trajectory.

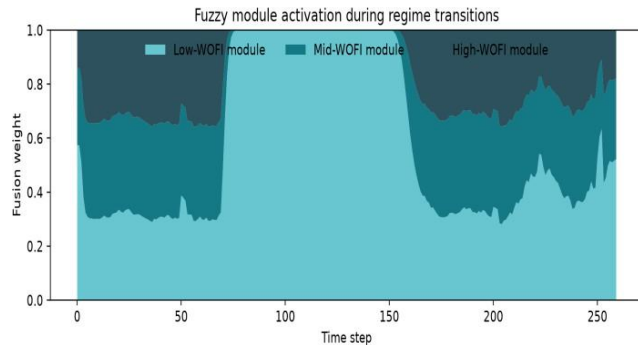


Fig. 3. Gaussian fusion weights across a held-out trajectory.

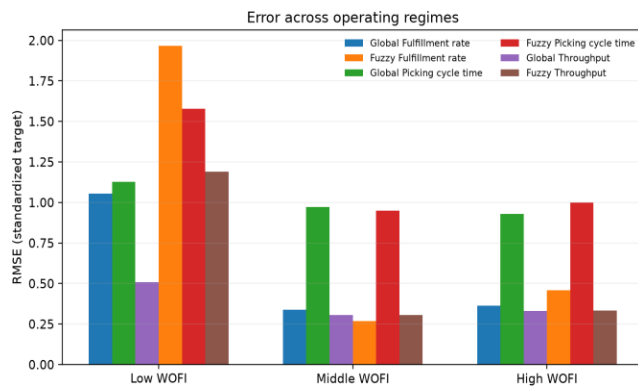


Fig. 4. Prediction error across low-, middle-, and high-WOFI test regimes.

### C. Evaluation Interpretation

The representative held-out trajectory (Fig. 2) shows that both predictors follow the broad simulated dynamics, but the fuzzy model exhibits larger deviations during and after workload shocks. Concurrently, the Gaussian fusion weights across the trajectory (Fig. 3) confirm that the model's activation changes smoothly rather than switching discontinuously. However, as the aggregate error metrics demonstrate, mathematical smoothness alone does not guarantee predictive accuracy. Finally, the prediction error stratified across low-, middle-, and high-WOFI test regimes (Fig. 4) provides a detailed view of where the modular model struggles, indicating whether future improvements should focus on better local specialization or a more reliable global baseline.

The current experiment supports a reproducible benchmark and establishes a clear evaluation protocol, rather than making a direct deployment claim. Before drawing operational conclusions, future studies should repeat the experiment over multiple random seeds, report confidence intervals across trajectories, tune membership widths using validation data only, and compare against simpler regime-aware baselines (such as a hard-switch model) and a mixture-of-experts model with learned gating. Furthermore, real warehouse-management-system (WMS) data are required to assess external validity.

### D. Limitations and Future Work

Although this study provides a rigorous and reproducible benchmark for predicting multiple warehouse-performance measures, it has several limitations. First, the evaluation was conducted entirely in a simulated environment. While the simulator represents important warehouse dynamics, such as congestion, replenishment delays, and capacity constraints, results generated from assumed system equations cannot, by themselves, demonstrate that the models will perform equally well in an industrial setting. Second, the experiments used a single random seed for both data generation and model initialization. This limits the extent to which the reported error metrics can be generalized. Finally, the weaker performance of the fuzzy-modular architecture points to a specific methodological issue. In a high-dimensional and heterogeneous input space, using simple distance-based Gaussian fusion can give too much influence to poorly calibrated local models, particularly when only a small number of training examples are available for a given operating regime.

Future research will address these limitations in three main ways. First, the experiments will be repeated using multiple random seeds, with mean performance and 95% confidence intervals reported across a broader range of operational scenarios. Second, the architecture will be improved by replacing fixed Gaussian validity functions with learned, data-driven gating mechanisms, such as those used in mixture-of-experts models. Membership-function widths will also be tuned using a dedicated validation set to reduce the risk of overfitting regime boundaries. Third, and most importantly, the framework will be tested using high-frequency data from real Warehouse Management Systems.

This step is essential for determining whether regime-aware modularization can improve predictive accuracy, robustness during operational transitions, and decision support in complex warehouse environments beyond the simulated setting.

## V. CONCLUSION

This study presented a simulation-based, multi-output predictive framework for warehouse performance, benchmarking a global multilayer perceptron (MLP) against a fuzzy-integrated modular neural architecture. Experimental evaluation revealed that the modular model underperformed the global baseline, with aggregate RMSE increasing by 49.9%. Rather than invalidating modular prediction, this outcome highlights the critical challenges of applying naive Gaussian fusion in high-dimensional, heterogeneous state spaces. It demonstrates that mathematical smoothness across regime boundaries does not guarantee predictive accuracy without rigorous local calibration and sufficient training data.

Despite this comparative performance outcome, the study establishes three foundational contributions: 1) a holistic multi-output prediction task capturing interdependent warehouse dynamics; 2) a fully specified, reproducible nonlinear simulator; and 3) a rigorous evaluation protocol that explicitly isolates within-regime accuracy from transition-phase robustness. Future work will address the identified limitations by replacing distance-based Gaussian validity functions with learned gating mechanisms and validating the framework against real-world warehouse management system (WMS) data to assess industrial deployment readiness.

## 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. 265120].

## REFERENCES

- [1] J. J. Bartholdi III and J. P. Hackman, Warehouse & Distribution Science, First draft: May 22, 1998; revised August 20, 2006.
- [2] Kabak, Ö., Güleç, N. Data-Driven Approach to Order Picking Time Prediction Using Fuzzy Clustering and ANN. In: Kahraman, C., Cebi, S., Cevik Onar, S., Oztaysi, B., Tolga, A.C., Sari, I.U. (eds) Intelligent and Fuzzy Techniques for Emerging Conditions and Digital Transformation. *INFUS. Lecture Notes in Networks and Systems*, vol 307. Springer, Cham. 2021. Doi: 10.1007/978-3-030-85626-7\_3
- [3] Frikha M.A. and Mraihi R. "Audit of Authorised Economic Operators (AEO): A new risk management approach based on Logistics Information System (LIS) analysis", 2020 IEEE, 13 th international Conference on Logistics (LOGISTIQUA), 2020. 10.1109/LOGISTIQUA49782.2020.9353944
- [4] Wenrong Lu, Duncan McFarlane, Vaggelis Giannikas, Quan Zhang, An algorithm for dynamic order-picking in warehouse operations. *European Journal of Operational Research*, vol 248, no. 1, pp. 107-122, ISSN 0377-2217, 2016. Doi: 10.1016/j.ejor.2015.06.074.
- [5] J. A. Riedel, R. Facilities planning – 4th edition by J.A. Tompkins, J.A. White, Y.A. Bozer and J.M.A. Tanchoco. *International Journal of Production Research*, vol. 49, no. 24, pp. 7519–7520. Doi : 10.1080/00207543.2011.563164.
- [6] Rodrigo Furlan de Assis, Alexandre Frias Faria, Vincent Thomasset-Laperrière, Luis Antonio Santa-Eulalia, Mustapha Ouhimmou, William de Paula Ferreira. *Machine Learning in Warehouse Management: A Survey, Procedia Computer Science*, Vol 23, pp 2790-2799, ISSN 1877-0509, 2024. Doi: 10.1016/j.procs.2024.02.096.
- [7] Md Redwanul Islam, Md Nahid Hossain, & Md. Zahid Hasan Tusar. predictive analytics in supply chain management a review of business analyst-led optimization tools. *Review of Applied Science and Technology*, vol. 6, no. 1, pp. 34-73. 2021. Doi: 10.63125/5aypx555
- [8] J. -S. R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, no. 3, pp. 665-685, May-June 1993. Doi: 10.1109/21.256541.
- [9] Robert A. Jacobs, Michael I. Jordan, Steven J. Nowlan, Geoffrey E. Hinton; Adaptive Mixtures of Local Experts. *Neural Comput*; vol. 3, no 1, pp 79-87, 1991. Doi: 10.1162/neco.1991.3.1.79
- [10] Noam Shazeer, Azalia Mirhoseini, Krzysztof Maziarsz, Andy Davis, Quoc Le, Geoffrey Hinton, Jeff Dean. Outrageously Large Neural Networks. *The Sparsely-Gated Mixture-of-Experts Layer*, vol. 1, 2017. DOI: 10.48550/arXiv.1701.06538.
- [11] Justus Ngunjiri, artificial intelligence-driven supply chain optimisation: integrating predictive analytics for resilient logistics systems. *Gulf Journal of Advance Business Research*. Vol. 1 No. 3, 2023.
- [12] Frikha, M. A., Mrad, M., Boujelbene, Y., & Soliman, Y. M. A. Hybrid State Estimation and Recurrent Neural Networks for Supply Chain Disruption Detection. *International Journal of Advanced Computer Science and Applications*, vol 17, no 7, 2026. Doi: 10.14569/IJACSA.2026.0170789.
- [13] De Assis, R. F., Faria, A. F., Thomasset-Laperrière, V., Santa-Eulalia, L. A., Ouhimmou, M., & de Paula Ferreira, W.. Machine learning in warehouse management: A survey. *Procedia Computer Science*, vol 232, pp 2790–2799, 2024. Doi: 10.1016/j.procs.2024.02.096
- [14] Kabak, Ö., & Güleç, N.. Data driven approach to order picking time prediction using fuzzy clustering and ANN. In C. Kahraman, S. Cebi, S. C. Onar, B. Oztaysi, A. C. Tolga, & I. U. Sari (Eds.), *Intelligent and fuzzy techniques for emerging conditions and digital transformation* (Lecture Notes in Networks and Systems, Springer, vol. 307. 2022. Doi: 10.1007/978-3-030-85626-7\_3
- [15] Lu, W., McFarlane, D., Giannikas, V., & Zhang, Q. An algorithm for dynamic order-picking in warehouse operations. *European Journal of Operational Research*, vol. 248 no. 1, pp 107–122, 2016. Doi: 10.1016/j.ejor.2015.06.074
- [16] J. -S. R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, no. 3, pp. 665-685, May-June 1993. Doi: 10.1109/21.256541.
- [17] Jacobs, R. A., Jordan, M. I., Nowlan, S. J., & Hinton, G. E.. Adaptive mixtures of local experts. *Neural Computation*, vol. 3, no. 1, pp. 79–87, 1991. Doi: 10.1162/neco.1991.3.1.79.
- [18] M. A. Frikha, M. Mrad, Y. Boujelbene, and S. Ben Othman, "Deep hybrid learning for sustainable industrial forecasting: Integrating CNN–LSTM models to enhance economic efficiency and carbon performance, *International Journal of Advanced Computer Science and Applications*, vol. 17, no. 2, 2026. Doi: 10.14569/IJACSA.2026.0170264.
- [19] Rad, F.F., Oghazi, P., Onur, İ. et al. Adoption of AI-based order picking in warehouse: benefits, challenges, and critical success factors. *Rev Manag Sci* 19, pp. 3495–3540, 2025. Doi: 10.1007/s11846-025-00858-1.
- [20] M. A. Frikha, "Sustainable and transparent global supply chains through blockchain-enabled logistics information systems," *Quality-Access to Success*, vol. 27, no. 213, pp. 46–54, 2026. Doi: 10.47750/QAS/27.213.05.
- [21] X. Wang, L. Zhang, L. Wang, E. Ruiz Zuñiga, X. V. Wang, and E. Flores-García, Dynamic multi-tour order picking in an automotive-partwarehouse based on attention-aware deep reinforcement learning, *Robotics and Computer-Integrated Manufacturing*, vol. 94, p. 102959, 2025. Doi: 10.1016/j.rcim.2025.102959.