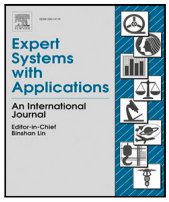
Expert Systems With Applications 261 (2025) 125464

Contents lists available at ScienceDirect

Expert Systems With Applications

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



TFT-MPIR: An end-to-end multi-period inventory replenishment strategy based on temporal fusion transformer

Guo Weixing [a], Ren Zhuoming [a,*], Du Wenli [b], Weng Tongfeng [a]

[a] Alibaba Business School, Hangzhou Normal University, Hangzhou, 311121, China
[b] Modeling Engineering Risk and Complexity, Scuola Superiore Meridionale, Naples, 80138, Italy

ARTICLE INFO

Dataset link: https://www.coap.online/competitions/1

Keywords:
Inventory replenishment
End-to-end
Temporal fusion transformer
Neural network

ABSTRACT

The inherent uncertainty associated with demand and vendor lead time significantly complicates replenishment strategies, which is challenge in the dynamic realm of e-commerce platforms. An end-to-end multi-period inventory replenishment strategy based on temporal fusion transformer (TFT-MPIR) is tailored for an integrated inventory replenishment decision-making process, and takes into account stochastic demand, vendor lead time, as well as linear transportation costs. TFT-MPIR which is fundamentally trained on an extensive amount of historical data, utilizes deep learning to directly calculate replenishment orders based on contextual and historical insights, deviating from the conventional two-step Predict-Then-Optimize (PTO) approach. The TFT-MPIR neural network framework designed with the concept of modularity enables an in-depth understanding and optimization of its structure and parameters. Specifically, the demand forecasting module utilizes temporal fusion transformer for advanced multi-quantile forecasting, generating comprehensive demand projections that significantly improve the accuracy of subsequent replenishment decisions. Numerical experiments incorporate authentic historical data from a prominent beverage supplier. Compared to the optimal solution (OPT) for inventory costs, TFT-MPIR exhibits a variance of 15.8%, markedly surpassing other integrated inventory strategies namely (t, R, S), PTO, and E2E-Multi-layer perceptron(E2E-MLP), which demonstrate divergences of 34.8%, 24.1%, and 22.3% respectively from OPT. Furthermore, TFT-MPIR framework achieves a cost reduction of 8.3% relative to the conventional PTO, and 19% in comparison to the (t, R, S). The robustness and scalability of the TFT-MPIR are substantiated through the adjustment of the ratio between unit stockout cost and unit transportation cost, coupled with sensitivity analysis.

1. Introduction

In the swiftly evolving landscape of e-commerce retail, the industry is witnessing a surge in customer diversity, an expanding range of products, and significant fluctuations in order volumes. Such evolution has resulted in the accumulation of extensive data, including characteristics of items (Stock Keeping Units or SKUs), historical sales figures, and historical vendor lead time (VLT) data. The challenge of integrating data-driven methodologies with large-scale data to boost inventory replenishment efficiency is an intriguing area of study (Snyder & Shen, 2019). Demand and VLT stand as pivotal factors in shaping inventory replenishment strategies. Traditional research often relied on probabilistic distribution assumptions for demand and VLT to devise replenishment policies (Levi, Janakiraman & Nagarajan, 2008; Levi, Pál, Roundy, & Shmoys, 2007; Levi, Roundy, Shmoys and Truong, 2008; Levi & Shi, 2013). However, in the real-world e-commerce setting, the pronounced unpredictability and variability of both demand and VLT render such assumptions impractical. Factors such as demand uncertainty, seasonal fluctuations, and promotional activities complicate inventory management, leading to the risk of high inventory costs for retailers. Therefore, precise inventory management is crucial for maintaining efficient operations and improving market competitiveness.

As the volume of datasets increases, it is likely to improve the effectiveness of order decision-making in practical applications (Maheshwari, Gautam, & Jaggi, 2021). The datasets contain numerous demand-related features, which aid in making better decisions based on real information (Babai, Boylan, & Rostami-Tabar, 2022; Kamble & Gunasekaran, 2020; Kuo & Kusiak, 2019). Machine learning (ML) techniques, leveraging historical and contextual data, offer highly accurate predictive capabilities (Mandi, Stuckey, Guns, et al., 2020). For example, Cohen et al. employed ML models to forecast future

* Corresponding author.
E-mail address: zhuomng.ren@hznu.edu.cn (Z. Ren).
URL: http://www.researcherid.com/rid/M-9158-2014 (Z. Ren).

https://doi.org/10.1016/j.eswa.2024.125464
Received 14 January 2024; Received in revised form 2 July 2024; Accepted 25 September 2024
Available online 1 October 2024

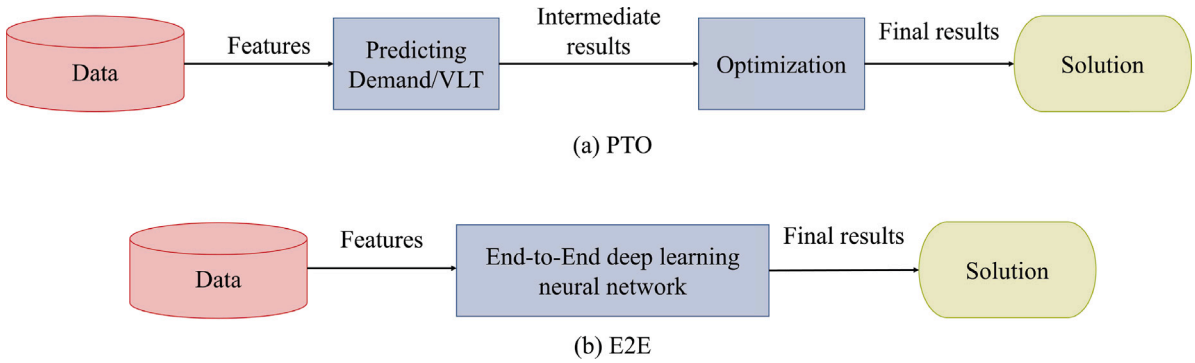


Fig. 1. PTO & E2E.

product demands, subsequently using predictions to determine optimal promotional pricing through non-linear integer programming (Cohen, Leung, Panchamgam, Perakis, & Smith, 2017). The methodology exemplifies the Predict-Then-Optimize (PTO) strategy (Elmachtoub & Grigas, 2022), a two-pronged approach that initially forecasts uncertain input parameters using ML models trained on historical data, followed by the resolution of the related optimization problem using forecasts above. While PTO strategy is prevalent, it fundamentally focuses on minimizing predictive errors without considering the ramifications of these predictions on subsequent optimization stages (Elmachtoub, Liang, & McNellis, 2020). Besides, even with demand and VLT predictions, addressing inventory issues remains challenging (Halman, Orlin, & Simchi-Levi, 2012). In the context of retail industry professionals, the principal emphasis lies on the efficacy of decision-making processes, as opposed to the mere minimization of predictive inaccuracies (Mukhopadhyay, Vorobeychik, Dubey, & Biswas, 2017; Wang et al., 2006). In addition, there is a trend towards simultaneously performing prediction and optimization. For the feature-based newsvendor problem, Ban and Rudin derived generalization bounds for the out-of-sample performance of the cost and the finite-sample bias from the true optimal decision (Ban & Rudin, 2019). Oroojlooyjadid et al. proposed a Multilayer Perceptron (MLP) model to optimize order quantities (Oroojlooyjadid, Snyder, & Takáč, 2020). However, their primary work focuses on the classical newsvendor problem, without considering multi-period settings and VLT.

Fig. 1(a) presents the decision-making mechanism within PTO framework. Paradoxically, demand and VLT forecasts that diverge further from actual values may lead to more suitable decisions, ultimately reducing inventory costs. This depiction underscores that a predictive model remains valuable even when significant errors are evident, provided that the forecasted cost vectors effectively lead to accurate decision-making.

There is often a discrepancy between the criteria for training algorithms and the final evaluation standards in many ML applications. To bridge this gap, Donti et al. introduced an End to End (E2E) methodology, and aimed at minimizing expected losses (Donti, Amos, & Kolter, 2017). Their approach has shown its strengths in addressing classic inventory challenges, managing real-world electrical grids, and optimizing energy storage arbitrage tasks. Additionally, Tian et al. proposed an E2E approach to address the data-driven newsvendor problem, which effectively reduces the costs associated with order decision-making (Tian & Zhang, 2023). In contrast to PTO framework, the E2E solution directly linking data input to replenishment decisions—proves to be more effective in inventory management, particularly when backed by a substantial amount of raw data. The E2E holistic approach to inventory decision-making as depicted in Fig. 1(b), embraces the entire process from input to output. The E2E method transcends traditional, segmented approaches by avoiding the need for manual feature engineering or intermediary steps, instead directly mapping inputs to outputs for a more streamlined and automated problem-solving process. This not only lightens the load of manual feature engineering but also minimizes reliance on intermediate outcomes, potentially revealing insights that conventional methods might miss. Consequently, E2E learning views the problem in its entirety, rendering the system more cohesive and efficient.

The E2E methodology presents an effective resolution for single-period inventory management challenges with uncertain demand, akin to the well-known newsvendor problem (Arrow, Harris, & Marschak, 1951), while the data-driven newsvendor problem has emerged as a research hotspot in the field of management science in recent years (Huber, Müller, Fleischmann, & Stuckenschmidt, 2019; Lin, Chen, Li, & Shen, 2022; Neghab, Khayyati, & Karaesmen, 2022). However, inventory management in e-commerce retail transcends this, encompassing complex, multi-period replenishment issues. The intricacy of multi-period replenishment stems from the dependency of each period's ordering quantity on both the current demand and VLT, as well as influenced by preceding and subsequent inventory levels. Crafting an ordering strategy, therefore, necessitates adapting to each period's evolving information, feedback, and potential future impacts. The examination of such problem began with the works of Ehrhardt (1984) and Kaplan (1970), which highlighted the benefits of (s, S) policies and myopic base stock policies, and base stock's optimality has been demonstrated under various settings (Gallego & Özer, 2001; Iida & Zipkin, 2006; Muharremoglu & Tsitsiklis, 2008). However, transportation costs should be considered. Moreover, with transportation costs being a significant factor, the larger the order quantity, the higher the transportation expense. Hence, minimizing the total inventory cost involves a delicate balance: preventing exorbitant stockout costs due to inadequate ordering while controlling the excess transportation costs associated with overordering.

Recognizing the pivotal role of demand and VLT in replenishment decision-making, we develop a deep learning framework which adopts a modular structure, comprising modules for demand prediction, VLT forecasting, and replenishment quantity estimation. The demand module harnesses historical demand data to project future demands, and the VLT prediction module forecasts future VLT. During the modules' forecasting processes, contextual elements like holidays, weekends, and SKU specifics are considered. To capitalize on such characteristics and discern the periodic patterns within historical demand data for multi-quantile forecasting, we integrate the Temporal Fusion Transformer (TFT) model (Lim, Arık, Loeff, & Pfister, 2021). The interpretable time series forecasting algorithm, pioneered by Google Cloud AI, adeptly processes varying input features and incorporates interpretability aspects. It adeptly handles complex time series structures, necessitating the extraction of diverse data features, including static covariates, known future inputs, and other historically observed exogenous time series. The TFT model skillfully utilizes these extensive features to amalgamate various inputs, thereby boosting the forecasting efficacy across multiple tasks. Following the design of two modules, corresponding predictive outputs, along with other elements like replenishment status, inventory levels, and historical arrival quantities,

Table 1
Contribution of this article compared to previous studies.

Method	Articles	Period	Demand	VLT	Trans cost	Neural network
Base stock and (s,S)	Ehrhardt (1984) and Kaplan (1970)	Multi	Assume certain knowledge	Assume certain knowledge	No	–
Approximation algorithms	Levi, Janakiraman et al. (2008), Levi et al. (2007) and Levi, Roundy et al. (2008)	Multi	Assume certain knowledge	No	No	–
Approximation algorithms	Levi and Shi (2013)	Multi	Assume certain knowledge	Assume as a known constant	No	–
Joint estimation optimization	Ban and Rudin (2019)	Single	Uncertain	No	No	–
Joint estimation optimization	Oroojlooyjadid et al. (2020)	Single	Uncertain	No	No	MLP
E2E	Donti et al. (2017)	Single	Randomly generated	No	No	Fully connected network
E2E	Tian and Zhang (2023)	Single	Uncertain	No	No	Modular designed and integrated textual review features
E2E	Qi et al. (2023)	Multi	Uncertain	Uncertain	No	Modular designed and based on MQRNN (Wen, Torkkola, Narayanaswamy, & Madeka, 2017)
E2E	This paper	Multi	Uncertain	Uncertain	Linear	Modular designed and based on TFT

are inputted into the replenishment quantity prediction module to determine the final replenishment decision. The proposed method integrates tasks that were previously separated in PTO (such as demand and VLT prediction and replenishment decision-making) into a single deep learning framework. Here, predictions of demand and VLT serve as auxiliary tasks, facilitating rapid model training, enabling performance monitoring, and enhancing generalizability to mitigate overfitting. In contrast to neural networks with direct full-feature interconnectivity, the modular approach delivers superior interpretability and efficiency, ensuring robust performance (Qi et al., 2023).

From the contributions outlined in Table 1, the differences between the study and previous approaches can be clearly seen. The proposed method focuses on using an E2E deep learning approach to address the multi-period replenishment problem, which involves uncertainties in both demand and VLT, while also considering linear transportation costs. To better handle historical demand and VLT as time-series data, we have designed a modular neural network model based on the TFT time-series forecasting model. The contributions are encapsulated as follows. Firstly, we innovatively applied an E2E deep learning framework to address the challenges of multi-period inventory replenishment under the uncertainty of demand and VLT, including linear transportation costs. Our framework adeptly outputs replenishment decisions for order replenishment days, guided by the input features. Secondly, utilizing both simulated and real-world data from a beverage supplier, our experiments revealed that our end-to-end learning framework surpasses several benchmark models, particularly in achieving reduced inventory costs. lastly, proposed approach integrates a modular neural network design, incorporating TFT and LSTM models for predicting unknown demand and VLT, subsequently integrated with a replenishment quantity decision module.

The remainder of the paper is organized as follows: Section 2 delves into the multi-period inventory replenishment issue with linear transportation costs, outlining a post-hoc marking process using the gurobi solver to derive the target vector. Section 3 introduces the TFT model and explicates our deep learning neural network architecture. In Section 4, we engage in offline numerical experiments using both simulated and real-world data to validate the efficacy of our proposed model in reducing total inventory costs compared to other methods. Finally, Section 5 concludes with a summary and suggestions for potential future research directions.

2. Considering linear transportation costs and post-hoc optimal solutions in multi-period inventory replenishment challenges

2.1. Overview of the inventory replenishment problem

In the study, we tackle a multi-period inventory replenishment challenge characterized by unpredictable demand and VLT, coupled with transportation costs that vary linearly with order volume. Here's the scenario: for an individual item stored in a single distribution center (DC), we consider a planning horizon denoted as T. The uncertain demand on day t is represented by d_t. If a day concludes with surplus inventory, the inventory level is positive, which noted as the end-of-day stock quantity. Conversely, if a shortage occurs at day's end, leading to stockouts, which noted as stockout costs are immediately incurred, and the inventory is reset to zero. Hence, at the end of each cycle, we incur either a holding cost h per excess unit or a stockout cost b per unit short. Moreover, if a replenishment order is placed on any given day, associated transportation costs based on the order volume are also applicable.

The proposed model incorporates a periodic replenishment policy, where the replenishment cycle is predetermined by specific dates, allowing for orders to be placed at set intervals (such as on Mondays and Fridays). Concurrently, the model considers the stochastic nature of VLT, labeled as l_t, denoting that an order made on day t is expected to be delivered on day $t + l_t$, with l_t being a random positive integer. The delivery time l_t for a given day t remains uncertain until the order is received on day $t + l_t$. Inspired by the graphical representations of inventory level changes, Bhavani, Mahapatra and Kumar (2023) and Bhavani, Mishra and Mahapatra (2023) considered a sustainable two-echelon green supply chain model and a two green-warehouse inventory systems with green investments, we drew Fig. 2 which depicts the inventory replenishment scenario, derived from the real data provided in this paper. The blue line illustrates the fluctuations in inventory levels, which change in response to daily sales and arrivals. The timeline begins at t_0, with certain dates marked as replenishment days, suggesting possible order placements. The horizontal purple dashed line shows the quantity of orders in transit. For instance, an order placed on t_0 will arrive four days later on t_4, since $l_{t_0} = 4$, hence $t_0 + l_{t_0} = t_4$, thereby replenishing the inventory.

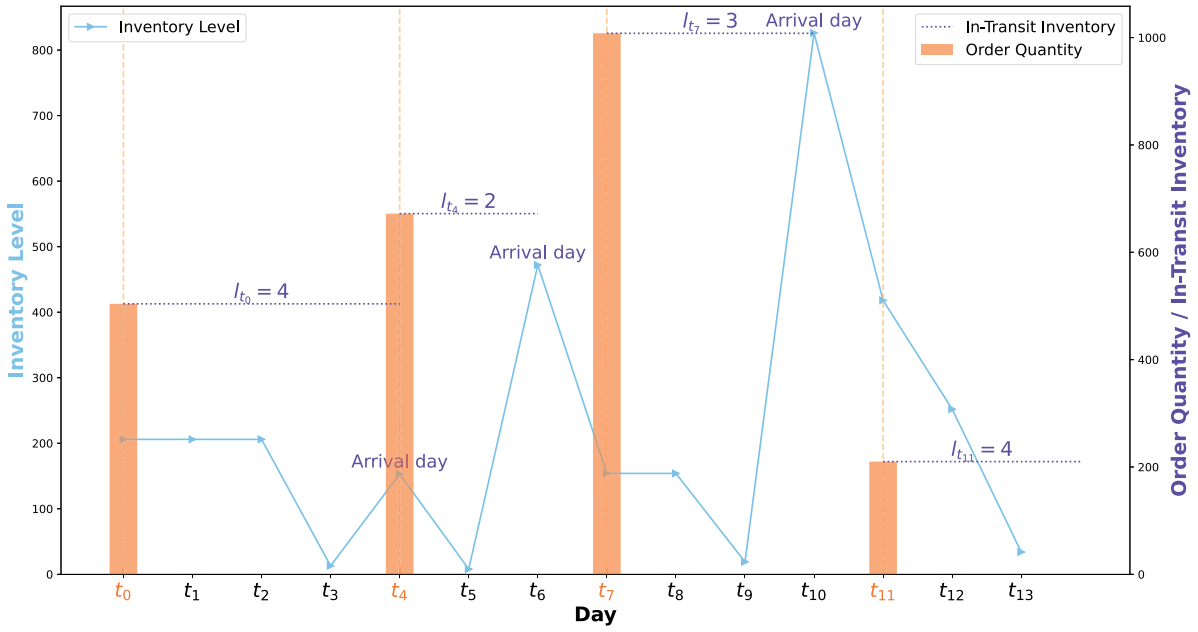


Fig. 2. Periodic replenishment.

Table 2
Parameter symbols and explanatory notes.

Parameter	Description
$T, \tilde{T}, T_H$	Planning horizon, length of input contextual data, historical data
d_t	Demand on day t
l_t	VLT on day t
a_t	Whether replenishment occurs on day t
q_t, q_t^*	Replenishment quantity on day t, OPT on day t
p	Number of boxes per pallet
K_f	Transportation cost per box from the current DC to factory f
h, b	Holding cost per box, stockout cost per box
m_t, n_t	Inventory level on day t, stockout quantity on day t
λ_i	Loss weight for each output i

Inventory costs are calculated in the following manner: The overall inventory cost is comprised of transportation, holding, and stockout costs. At the beginning of each cycle, it is imperative to verify whether any replenishment orders are arriving on that day, acknowledging that several orders might arrive simultaneously. Subsequently, demand levels are considered, and the inventory level at the end of the period is updated. Depending on final inventory level and any shortages, either holding costs or stockout costs are incurred. Moreover, on days designated for replenishment, when orders are placed, corresponding transportation costs are also incurred. Our goal is to minimize the total inventory cost within the specified planning period by accurately determining the necessary replenishment quantities.

2.2. Symbol explanation and model representation

Below are the parameters pertinent to the problem and model, as outlined in Table 2.

The problem of the multi-period inventory replenishment is encapsulated in a mathematical framework, using the function$(\cdot)^+ = max(\cdot, 0)$ to denote the maximum value in comparison with 0. In the context of problem (P1), q_t symbolizes the replenishment quantity on day t, quantified in pallets. The term p refers to the conversion ratio between pallets and boxes pertinent to the current SKU. Although boxes are typically employed for storage and sales in DCs, pallets are more efficient for picking and transportation processes. This setting necessitates conversion between different measurement units for diverse products. It is crucial to note that factories require DC replenishment in minimum units of pallets, with each SKU having a distinct number of boxes per pallet. Consequently, DCs must undertake replenishment in whole pallets. The binary variable a_t indicates whether an order is placed on day t. K_f denotes the per-box transportation cost for replenishment from two factories. It is worth noting that the scenario considers linear transportation costs, where costs are directly proportional to the transport quantity, implying that each unit of cargo transported incurs a specific cost. Thus, when placing orders for stock replenishment, it is essential to consider the increase in transportation costs corresponding to the order volume. Larger order sizes, while potentially reducing stockout rates and costs, can also lead to escalated holding and transportation expenses. The variables m_t and n_t represent the inventory level and the quantity of stockouts on day t, respectively. The initial inventory level is assumed to be m_{-1}. M is a number large enough to be significant. The goal of problem (P1) is to minimize the aggregate costs of transportation, holding, and stockouts.

Constraint (1) pertains to inventory equilibrium, allowing for the possibility of multiple orders arriving concurrently on the same day. Constraint (2) establishes the definition of daily stockout quantity. According to constraint (3), if day t is not a designated ordering day, then the order quantity on day t should be zero. Constraint (5) applies a binary condition to a_t, constraint (6) stipulates that inventory levels must be non-negative integers, and constraint (7) serves as a binary indicator for selecting between the two factories for procurement.

$$\min_{q_t, y_t, z_t} \sum_{t=0}^{T} \left(a_t q_t p K_f + h m_t + b n_t \right) \tag{P1}$$

$$\text{s.t. } m_t = \left(m_{t-1} + \sum_{i \in \{j : j + l_j = t\}} q_i p - d_t \right)^+ \quad \forall t = 0, \ldots, T \tag{1}$$

$$n_t = \left(-m_{t-1} - \sum_{i \in \{j : j + l_j = t\}} q_i p + d_t \right)^+ \quad \forall t = 0, \ldots, T \tag{2}$$

$$q_t \le M a_t \quad \forall t = 0, \ldots, T \tag{3}$$

$$m_t, n_t, q_t \ge 0 \quad \forall t = 0, \ldots, T \tag{4}$$

$$a_t \in \{0, 1\} \quad \forall t = 0, \ldots, T \tag{5}$$

$$m_t, n_t, q_t \in \mathbb{Z}^+ \quad \forall t = 0, \ldots, T \tag{6}$$

$$f \in \{0, 1\} \tag{7}$$

2.3. Approach to post-hoc optimal solutions and labeling process

In deep neural network training, mastering the transformation from input characteristics to output labels is vital. For our replenishment issue at hand, the objective revolves around ascertaining the ideal replenishment amount for each specified replenishment day after scrutinizing all encompassing context data X, which encompasses historical demand, historical VLT, and chronological data (like days, months, weekends, holidays, etc.). Nevertheless, given that the primary dataset only encompasses observational data of demand and lead times for replenishment without including the optimal replenishment quantities (the principal target vector), it is unsuitable for direct use as a training dataset. For each historical replenishment juncture, using the observed contextual data X, one needs to compute q^*, signifying the optimal daily replenishment order quantity, to act as the target vector in supervised learning. The aforementioned phase is termed as "labeling" within supervised learning techniques (James, Witten, Hastie, Tibshirani, et al., 2013). Once labeled data is accessible, the neural network training association can be exemplified as follows.

$$\min_{f:X\to R}\sum_{i=1}^{N} L(f(x_i);q_i^*) \tag{1}$$

N represents the aggregate count of training data, while L is the loss function outlined by the discrepancy between the model's forecast $f(x_i)$ and the optimal order volume q_i^*. Notably, we contemplate the neural network model for the function f, with the intricacies of the neural network architecture to be delineated in Section 3.2.

Here, we detail a methodology for calculating the optimal order labels based on historical data. Qi et al. originally developed a labeling process for multi-period inventory replenishment issues without accounting for transportation costs (Qi et al., 2023). However, incorporating linear transportation costs requires an alternate approach to compute the labels. Moreover, as both demand and VLT are accessible from historical data, problem (P1) can be redefined as a deterministic bulk problem, incorporating stockout costs, termed as problem (P2).

$$\min_{q_t,y_t,z_t}\sum_{t=-T_H}^{-1}\left(a_t q_t p K_f + h m_t + b n_t\right) \tag{P2}$$

$$\text{s.t. } m_t=\left(m_{t-1}+\sum_{i\in\{j:j+l_j=t\}} q_i p - d_t\right)^+ \quad \forall t=-T_H,\dots,-1 \tag{1}$$

$$n_t=\left(-m_{t-1}-\sum_{i\in\{j:j+l_j=t\}} q_i p + d_t\right)^+ \quad \forall t=-T_H,\dots,-1 \tag{2}$$

$$q_t \le M a_t \quad \forall t=-T_H,\dots,-1 \tag{3}$$

$$m_t,n_t,q_t \ge 0 \quad \forall t=-T_H,\dots,-1 \tag{4}$$

$$a_t \in \{0,1\} \quad \forall t=-T_H,\dots,-1 \tag{5}$$

$$m_t,n_t,q_t \in \mathbb{Z}^+ \quad \forall t=-T_H,\dots,-1 \tag{6}$$

$$f \in \{0,1\} \tag{7}$$

In the context of problem (P2), the historical data on demand and VLT extends from $-T_H$ day to -1 day, T_H is usually the larger value. Retailers amass a wealth of historical information, including past sales and VLT, for training data. Consequently, in this scenario, d_t and l_t are replaced with their actual historical counterparts. Addressing such a deterministic bulk problem, including stockout costs, can be addressed through dynamic programming or efficiently solved using commercial solvers. Solving this problem enables us to ascertain the optimal replenishment quantity, q_t^*, for each designated ordering day, specifically on days t where $a_t = 1$, culminating in the generation of labeled data.

3. An end-to-end multi-period inventory replenishment strategy based on temporal fusion transformer

This section presents the development of the E2E deep learning model namely Multi-Period Inventory Replenishment by Temporal Fusion Transformer (TFT-MPIR), which tailored for addressing multi-period inventory replenishment challenges with linear transportation costs.

3.1. Framework for resolution

The proposed framework for resolving multi-period replenishment issues is depicted in Fig. 3, comprising two distinct segments. The first segment involves the labeling. We commence by defining the multi-period inventory replenishment issue inclusive of transportation costs, preprocess the initial data, and utilize the gurobi solver to ascertain the optimal quantities for replenishment. Optimal quantities are then employed as the target vectors (labels) for supervised learning, resulting in an updated labeled dataset. The second segment focuses on the model training and prediction. We begin by segmenting the dataset, then proceed to train the training set within the TFT-MPIR neural network model. The model is further assessed on a validation set, with the aim of determining the most cost-effective model during the training phase as the ultimate predictive model. Finally, the trained model is utilized to generate the predicted quantities for replenishment.

3.2. TFT-MPIR neural network

With the post-hoc optimal problem resolved and the training data suitably labeled in Section 2, we have now developed training data incorporating the target vector. Our aim is to train the neural network to accurately map the relationship between contextual information and optimal decision-making. The detailed TFT-MPIR neural network is shown in Fig. 4. The inputs to the TFT-MPIR model are segmented into six distinct parts $(\{D_i\}_{i=-\tilde{T}+t}^{-1+t}, \{X_i\}_{i=-\tilde{T}+t}^{-1+t}, \{L_i\}_{i=-\tilde{T}+t}^{-1+t}, \{G_i\}_{i=-\tilde{T}+t}^{-1+t}, \{I_i\}_{i=-\tilde{T}+t}^{-1+t}, \{A_i\}_{i=t}^{t+h})$, each representing a different context sequence. D_t denotes features related to the historical demand noted prior to day t, while X_t encompasses standard features acquired before day t, including aspects like product category, warehouse location, and time-sensitive factors such as weekends or holidays. L_t relates to features associated with VLT observed before day t. G_t and I_t correspond to the optimal quantities for order arrivals and inventory levels respectively, calculated prior to day t. A_t is a determined Boolean value indicating whether replenishment is planned from day t to $t+h$. $\tilde{T}$ signifies the duration of the input contextual data. The TFT-MPIR model produces three distinct outputs (Out1, Out2, Out3), where Out1 is the forecasted demand, Out2 is the forecasted VLT and the principal output, Out3 symbolizes the ultimate replenishment decision, namely the replenishment quantity. Hence, the training goal is to fine-tune the function $f(\{D_i\}_{i=-\tilde{T}+t}^{-1+t}, \{X_i\}_{i=-\tilde{T}+t}^{-1+t}, \{L_i\}_{i=-\tilde{T}+t}^{-1+t}, \{G_i\}_{i=-\tilde{T}+t}^{-1+t}, \{I_i\}_{i=-\tilde{T}+t}^{-1+t}, \{A_i\}_{i=t}^{t+h})$ to closely replicate Out3. Considering that the input for demand is a time series format and future demand significantly influences replenishment decisions, the demand forecasting module is crafted used TFT model which has garnered remarkable results in time series forecasting, particularly in sectors like transportation and energy (Nazir, Shaikh, Shah, & Khalil, 2023; Wu, Wang, & Zeng, 2022, 2023; Zhang, Zou, Yang, & Yang, 2022).

TFT model is a multi-horizon time series forecasting deep learning model that possesses natural interpretability, offering stronger explanatory power compared to standard black-box models. Fig. 5 outlines the architecture of the TFT model. The TFT adeptly distinguishes and leverages various input types, encompassing static variables, known dynamic variables, and unknown dynamic variables. Within the model, its time-dependency processing component is capable of deriving both long and short-term temporal associations from observed or known

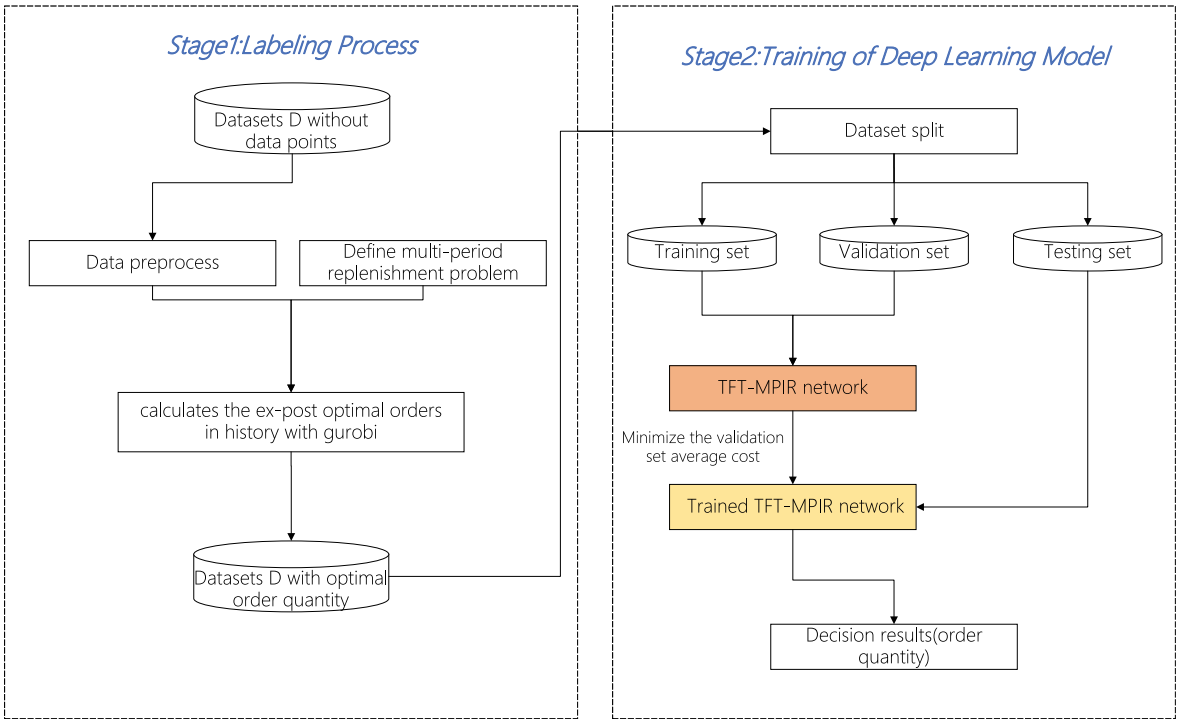


Fig. 3. E2E prediction framework.

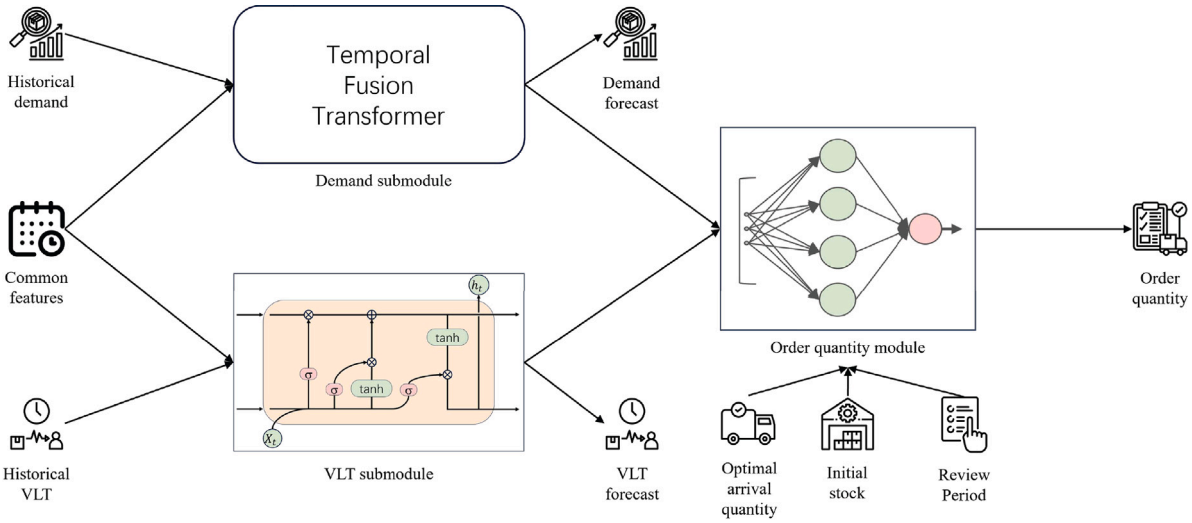


Fig. 4. TFT-MPIR neural network framework.

time-variable inputs. It effectively creates feature representations for each input type, thereby enhancing the predictive performance across various forecasting tasks. TFT model has five fundamental components: the gating mechanism, the variable selection network, the static covariate encoder, the temporal dependency processing module, and multi-level prediction interval prediction. (a) Gating Mechanism: Its purpose is to disregard unnecessary components within the architecture, adapting to various datasets and scenarios. (b) Variable Selection Network: Select the most significant input variables at each time step while ignoring less relevant variables simultaneously. (c) Static Covariate Encoder: It integrates static features into the network, which enhances the accuracy of predictions. (d) Temporal Dependency Processing Module: Capture long-term or short-term temporal dependencies from observed or known time-varying inputs. The Seq2Seq layer captures short-term dependencies, while long-term dependencies rely on the multi-head attention mechanism. (e) Multi-level Prediction Interval Prediction: Use quantile forecasting to determine the possible range of target values within each prediction interval. It generates multi-quantile daily demand forecasts as output, leading to more accurate replenishment decision-making. Complete details of the TFT can be found in the referenced literature (Lim et al., 2021).

In current demand forecasting task for the DC_SKU combination under consideration, there is a static covariate S(e.g. DC info or SKU info). y_t represents the target variable on day t (e.g. the sales value on the target day). x_t represents other time-dependent inputs, which include two parts, expressed as $x_t = [o_t, z_t]$. o_t denotes observed inputs that are unknown in advance and can only be measured when time step t arrives (e.g. sales or inventory level on day t). The z_t values can be predetermined (e.g. knowing the day of the week, holidays, or replenishment days). This study uses the observed inputs and targets until day

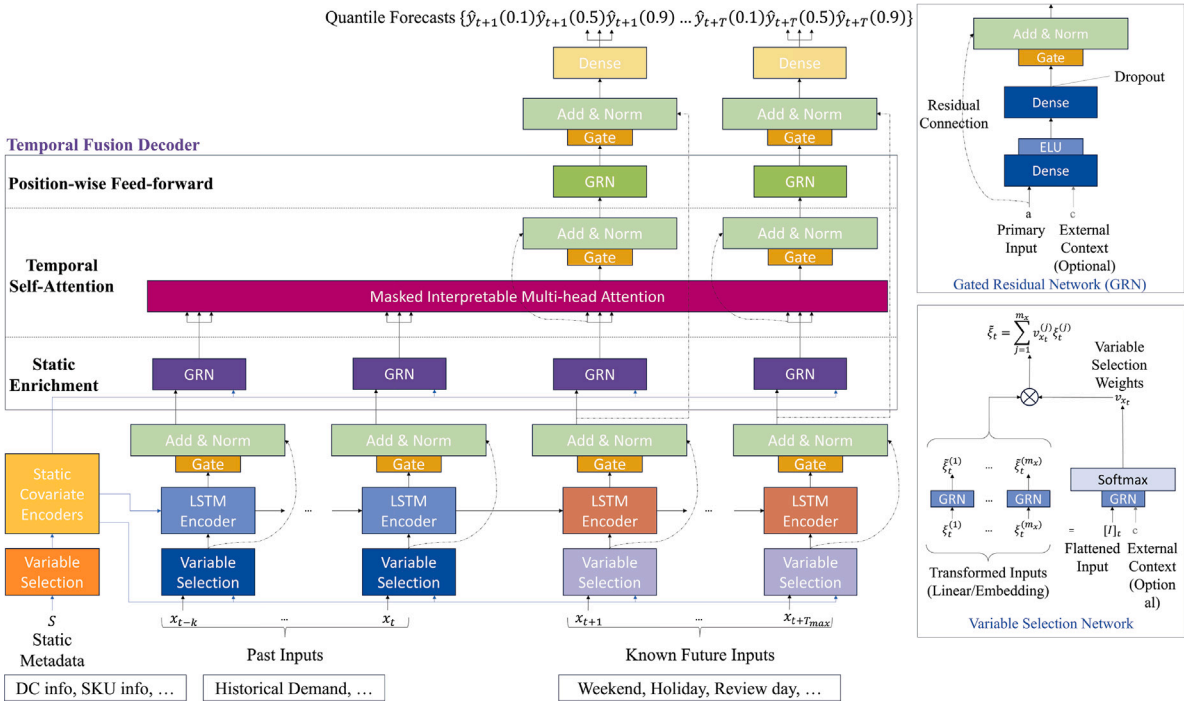


Fig. 5. The model architecture of TFT.

t (i.e. $y_{t-k:t} = \{y_{t-k}, \ldots, y_t\}$, $o_{t-k:t} = \{o_{t-k}, \ldots, o_t\}$) and known inputs across the entire range $z_{t-k:t+\tau}$ (i.e. $z_{t-k:t+\tau} = \{z_{t-k}, \ldots, z_t, \ldots, z_{t+\tau}\}$), where k is a finite look-back window, τ is the prediction point, $\tau \in (1, \ldots, \tau_{\max})$. Lastly, each quantile prediction is expressed as follows:

$$\hat{y}_i(q,t,\tau) = f_q\left(\tau, y_{t-k:t}, o_{t-k:t}, z_{t-k:t+\tau}, S\right) \tag{2}$$

In TFT-MPIR model, we integrate the Long Short-Term Memory (LSTM) network for VLT forecasting, a feedback neural network optimally designed for processing time series data (Graves & Graves, 2012). LSTM has demonstrated its efficacy in various domains, including finance and tourism (Cao, Li, & Li, 2019; Chimmula & Zhang, 2020; He, Ji, Wu, & Tso, 2021; Livieris, Pintelas, & Pintelas, 2020). The final module for replenishment decision-making is constructed as a fully connected layer. ReLU is employed as the activation function within our model, functioning as a segmented linear function that outputs positive inputs as they are, while nullifying negative ones. To combat overfitting, a common challenge in deep learning, we incorporate a Dropout layer which randomly disables a selection of neurons and their connections during the training process, effectively preventing overfitting and enhancing the integration efficiency across diverse network architectures (Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014). The training phase is governed by an objective function, defined as:

$$\min_{\theta} \sum_{i=1}^{N} \{\lambda_1 \hat{L}_1(Out1_i, d_t) + \lambda_2 \hat{L}_2(Out2_i, l_t) + \lambda_3 \hat{L}_3(Out3_i, q_t^*)\} \tag{3}$$

In this function, θ represents the collection of neural network parameters slated for optimization, and N denotes the total count of training data. The terms λ_1 and λ_2 are smaller numerical values penalizing inaccuracies in demand and VLT predictions, instrumental in regulating the gradient descent's iteration step size for streamlined model training. The first two components of the function calculate the loss between the predictions of demand and VLT and their actual values. The third component assesses the loss between the predicted replenishment quantities and the optimal quantities identified earlier. This design, anchored in the principle that replenishment decisions heavily depend on accurate demand and VLT information, is crucial for effective decision-making. Notably, when deploying the trained model on the test dataset, the output requires further adjustment, rounding it off to non-negative integers.

To sum up, the designed TFT-MPIR model can simultaneously predict demand, VLT, and replenishment quantities in each iteration. Additionally, the intermediate variables, demand and VLT, can be used to monitor model performance. By observing anomalies in demand and VLT outputs during real-time operation of the TFT-MPIR model, decision-makers can identify abnormal replenishment outputs and analyze their underlying reasons. Moreover, the model can continuously identify and adapt to long-term market trends based on historical data, providing ongoing inventory management support to retailers and helping them make more accurate inventory decisions.

4. TFT-MPIR evaluation

To appraise the efficacy of TFT-MPIR model and compare it with benchmark strategies, we created a series of simulated datasets and incorporated historical sales data from an actual beverage supplier, which effectively mirrors real-life replenishment situations. The ReLU function was employed as the activation function, with a Dropout rate set at 0.2. The learning rate during training was maintained at 0.01. For the loss function's objective, the weighting parameters were established as $\lambda_1 = \lambda_2 = 0.1$. We utilized PyTorch for neural network backpropagation, implementing an early stopping policy with a patience threshold of 20. The computational environment comprised a Linux operating system, powered by an Intel(R) Xeon(R) Gold 6330 CPU @ 2.00 GHz, with 256 GB RAM and an NVIDIA A40 GPU.

4.1. Data

For assessing the TFT-MPIR model, we engaged both artificially simulated and real-world datasets. The simulated dataset was crafted with a framework comprising a single factory and a single distribution warehouse, covering a spectrum of 1000 distinct SKUs, equating to 1000 different DC_SKUs combinations. This setup was informed by the "80/20 rule" of logistics research, which typically observes that approximately 20% of a company's products generate 80% of its sales

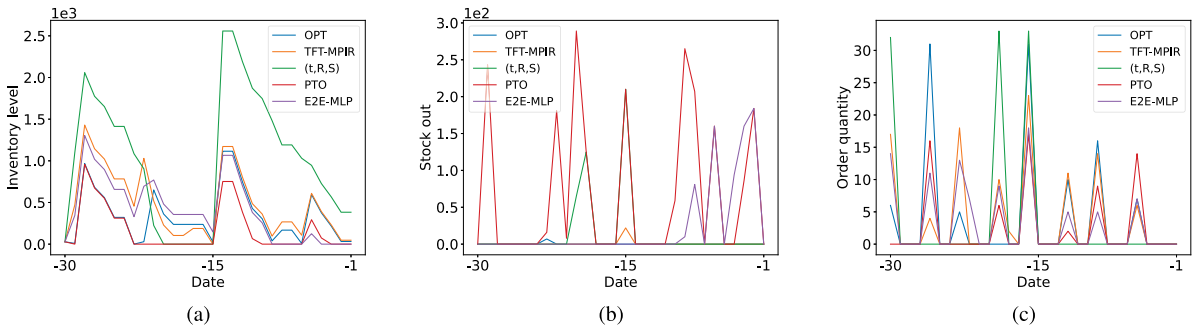


Fig. 6. Inventory change chart of a certain DC_SKU: (a) inventory level, (b) stockout quantity, (c) order quantity.

revenue (Coelho & Mendes, 2019). Accordingly, the average sales volumes in our simulated dataset of 1000 combinations adhered to a Pareto distribution. Each DC_SKU combination began recording sales and VLT from January 1, 2018, accumulating 900 records, where the sales data adhered to a normal distribution, and VLT values ranged from 1 to 10, following a gamma distribution.

The real dataset was derived from a leading beverage supplier, and included detailed information on the supply warehouse network and sales of a specific product category within a certain region. The company operated 2 production factories and 18 DC warehouses in that area, distributing 77 different SKUs. The original dataset encompassed 1080 DC and SKU combinations, which were narrowed down to 708 after a preliminary data cleansing process. Included in this dataset were historical sales figures, VLT records, initial inventory holdings, and various standard constraints like the unit transportation cost and box-to-pallet conversion ratios. The descriptive statistics of data is seen in Supplementary Material Tab. S1–S7.

4.2. Experiments

In the study, the experimental setup for both the simulated and actual datasets involved allocating the final thirty days as the test set for each DC_SKU pairing. The rest of the data was divided into 80% for training purposes and 20% for validation. The validation set played a crucial role in evaluating the neural network model's response to various combinations of hyperparameters, which was instrumental in choosing the most effective parameters and averting overfitting. The context for each day encompassed data from the previous 30 days, denoted as T being 30. For instance, the context for May 1st entailed historical data ranging from April 1st to April 30th, inclusive of target vectors. Considering the real dataset, each DC's transportation costs varied between the two supplying factories. Assuming no storage capacity limits at the factories and DCs, the DCs would naturally opt for the factory offering lower unit transportation costs. In the base scenario, the unit transportation cost for each DC_SKU combination was predetermined. It was assumed that the daily unit stockout cost for each combination was set at 150% of the unit transportation cost, with the unit holding cost being 1% of the unit stockout cost.

In experimental design, we utilized three benchmark strategies for comparison. The first is the (t, R, S) policy, an integrated inventory strategy that combines the (t, S) and (R, S) policies. (t, R, S) policy involves a fixed review period t, a maximum inventory level S, and a predetermined reorder point R. If inventory falls below the reorder point after the specified review period, a replenishment order is issued up to level S; otherwise, no order is placed, with the order quantity being the difference between the maximum inventory level and the inventory level at the time of review. The second benchmark employs PTO method, where a deep neural network is trained to precisely forecast future demand and VLT. This yields daily point forecasts for future demand and VLT, followed by solving a deterministic batch problem to ascertain the daily replenishment quantity, which can be efficiently solved using a commercial solver. The third benchmark involves an E2E model comprised entirely of MLP for predicting replenishment quantities, In the implementation of E2E-MLP methodology, the neural network architecture employs a MLP exclusively to map the input feature data onto optimal decision-making outcomes (Liu, Letchford, & Svetunkov, 2022; Oroojlooyjadid et al., 2020; Tian & Zhang, 2023).

To evaluate the performance, we use six metrics: total inventory cost, holding cost, stockout cost, transportation cost, stockout rate, and turnover rate. The total inventory cost is the aggregate of the holding, stockout, and transportation costs. The stockout rate is defined as the percentage of days with stockouts in the test set, indicating the frequency of stockouts. It is calculated as the ratio of days with stockouts (days where the stockout cost is above zero) to the total number of days. The inventory turnover rate is derived by dividing the average daily inventory level by the average demand.

4.3. Results

In the case of the simulated dataset as shown in Table 3, the proposed TFT-MPIR method outperformed the other four methodologies, securing the minimum total inventory cost. It also recorded the lowest figures in both stockout cost and stockout rate. We can see that an E2E deep learning model has two key advantages. Firstly, by comparing the costs of TFT-MPIR against PTO, and E2E-MLP against PTO, we can observe that using an E2E framework rather than a two-step PTO framework offers benefits. Both TFT-MPIR and PTO use the same deep learning models to predict demand and VLT, but TFT-MPIR provides replenishment decision results directly from the input, whereas PTO follows a two-step process. We suspect that prediction errors in demand and VLT can accumulate during the optimization phase, misleading the final replenishment decisions. The main optimization target of the end-to-end framework is the final replenishment quantity, which shortens the decision-making process and leads to more optimal results. Comparing the costs of TFT-MPIR with E2E-MLP, we believe that the use of TFT and LSTM models in our sub-modules gives our model a greater advantage in accurately predicting demand and VLT, resulting in more reasonable replenishment decisions and subsequently lower stockout costs.

Regarding the real dataset, Fig. 6 illustrates the inventory levels and ordering dynamics for a specific DC_SKU combination over the test period. In this scenario, it is assumed that for the thirty days leading up to the test period, the ordering was in line with the optimal replenishment quantity (OPT), and any in-transit orders, which were placed but had not arrived yet, were accurately included in the inventory calculations for these thirty days. In Fig. 6(a), the variations in inventory levels as per different methods are depicted. Fig. 6(b) showcases the stockout occurrences on a daily basis for each method. Finally, Fig. 6(c) demonstrates the quantities ordered on replenishment days by each method, quantified in terms of pallets.

Furthermore, we computed a cumulative cost change graph for this case study. As illustrated in Fig. 7, the (t, R, S) method is prone

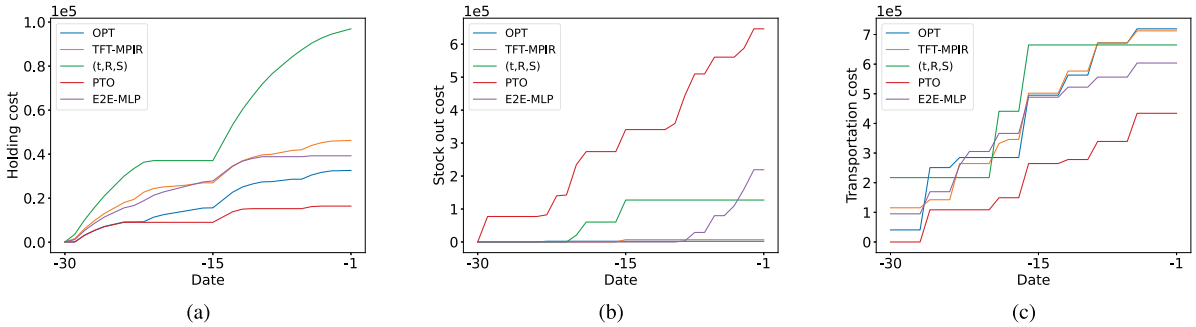


Fig. 7. Cumulative cost graph of a DC_SKU — (a) cumulative holding cost, (b) cumulative stockout cost, (c) cumulative transportation cost.

Table 3
Comparison results of different methods on simulated data. (The percentage values signify the positive (+) or negative (−) growth rates, indicating how the values of the current method compare to those of the OPT benchmark).

Method	Total cost	Holding cost	Stockout cost	Transportation cost	Stockout ratio	Turnover rate
OPT	286 242	10 361	5698	270 183	14.12	3.51
TFT-MPIR	**+15.7%**	+25.5%	**+270.9%**	+9.9%	**+20.0%**	+12.0%
(t, R, S)	+40.4%	+50.9%	+1148.5%	+16.6%	+188.5%	**−33.1%**
PTO	+25.2%	**−21.6%**	+1988.6%	**−14.4%**	+148.7%	−43.9%
E2E-MLP	+21.8%	+3.4%	+1131.5%	−0.9%	+101.3%	+1.9%

Table 4
Comparison results of different methods. (The percentage values signify the positive (+) or negative (−) growth rates, indicating how the values of the current method compare to those of the OPT benchmark).

Method	Total cost	Holding cost	Stockout cost	Transportation cost	Stockout ratio	Turnover rate
OPT	2 668 932	77 248	3078	2 588 606	3.61	2.89
TFT-MPIR	**+15.8%**	+7.1%	+28 185.0%	−17.4%	+354.0%	−8.4%
(t, R, S)	+34.8%	+137.0%	**+16 042.7%**	+12.7%	**+248.5%**	+46.4%
PTO	+24.1%	**−29.9%**	+44 051.3%	**−26.7%**	+729.4%	**−55.8%**
E2E-MLP	+22.3%	+78.1%	+31 847.8%	−17.2%	+377.6%	+44.9%

to ordering up to higher inventory levels, and witnesses the swiftest escalation in holding costs, culminating in the highest overall holding costs. Conversely, the PTO method, is characterized by the smallest order quantities, saws a slower increment in total holding costs, ending up with the lowest final total holding costs. However, the PTO method leads to a steep increase in total stockout costs, ultimately the highest among the methods. Notably, the changes across various aspects in TFT-MPIR method shows a closer alignment with the OPT method.

Table 4 presents a comprehensive performance assessment for 708 DC_SKU combinations from the actual dataset. According to the table, TFT-MPIR model reports lower total costs than all three benchmark strategies. Notably, it outperforms the PTO and E2E-MLP methods in achieving lower stockout rates and costs, achieving lower holding and transportation costs compared to (t, R, S) and E2E-MLP models. Besides, as a basic method, (t, R, S) performs the worst in both simulated and real data scenarios, resulting in the highest inventory costs.

We evaluate the performance of each strategy under varying ratios between unit stockout cost and unit transportation cost. As depicted in Fig. 8(a), TFT-MPIR method consistently achieves the lowest total inventory costs at ratios of 150%, 200%, and 250%. According to the detailed cost breakdown shown in Figs. 8(b), 8(c) and 8(d), it is evident that the escalation in stockout costs is the main driver of the overall increase in total inventory costs as the ratio increases, while the transportation costs remain relatively unchanged. Notably, even though the (t, R, S) strategy incurs the lowest stockout costs across varying ratios, it is responsible for the highest associated holding and transportation costs.

4.4. Sensitivity analysis

In this section, we conduct extensive sensitivity analyses, extending prior experiments to assess TFT-MPIR robustness and adaptability

Table 5
Network hyper-parameters.

Hyperparameter	Range
Learning rate	{0.0001, 0.001,0.005,0.01}
Dropout rate	{0.1,0.2,0.3}
TFT hidden size	{32,64,128}
TFT attention head size	{2,4,8}
LSTM hidden size	{32,64,100,128}
LSTM layer	{1,2}
Integration module	{50,50},{100,100},{100,100,50}

across a variety of hyperparameter choices and model covariates. Hyperparameters are pivotal in machine learning models as they essentially shape the model's architecture and its overall efficacy. The critical hyperparameters for the TFT-MPIR model are outlined in Table 5. During the model's training phase, a grid search method is utilized to explore various hyperparameter combinations within a predefined scope, and the average total inventory cost is computed on the validation set to identify the most optimal hyperparameter values. The standard setting for the ratio between unit stockout cost and unit transportation cost is set as 150%.

(1) Analysis of Learning Rate Sensitivity: The learning rate stands as a key hyperparameter in the training of neural networks (LeCun, Bengio, & Hinton, 2015). Table 6 delineates the impact of varying learning rates on the TFT-MPIR model's performance in the test set. In our analysis, an increase in the learning rate from 0.0001 to 0.01 corresponding with a reduction in both total and stockout costs, signifying the development of a more proficient model.

(2) Sensitivity Analysis of TFT Hidden Layer Size: The demand prediction module employs the TFT model, with the size of the TFT hidden layer being a primary hyperparameter. Table 7 indicates how different sizes of the TFT hidden layer impact the performance of

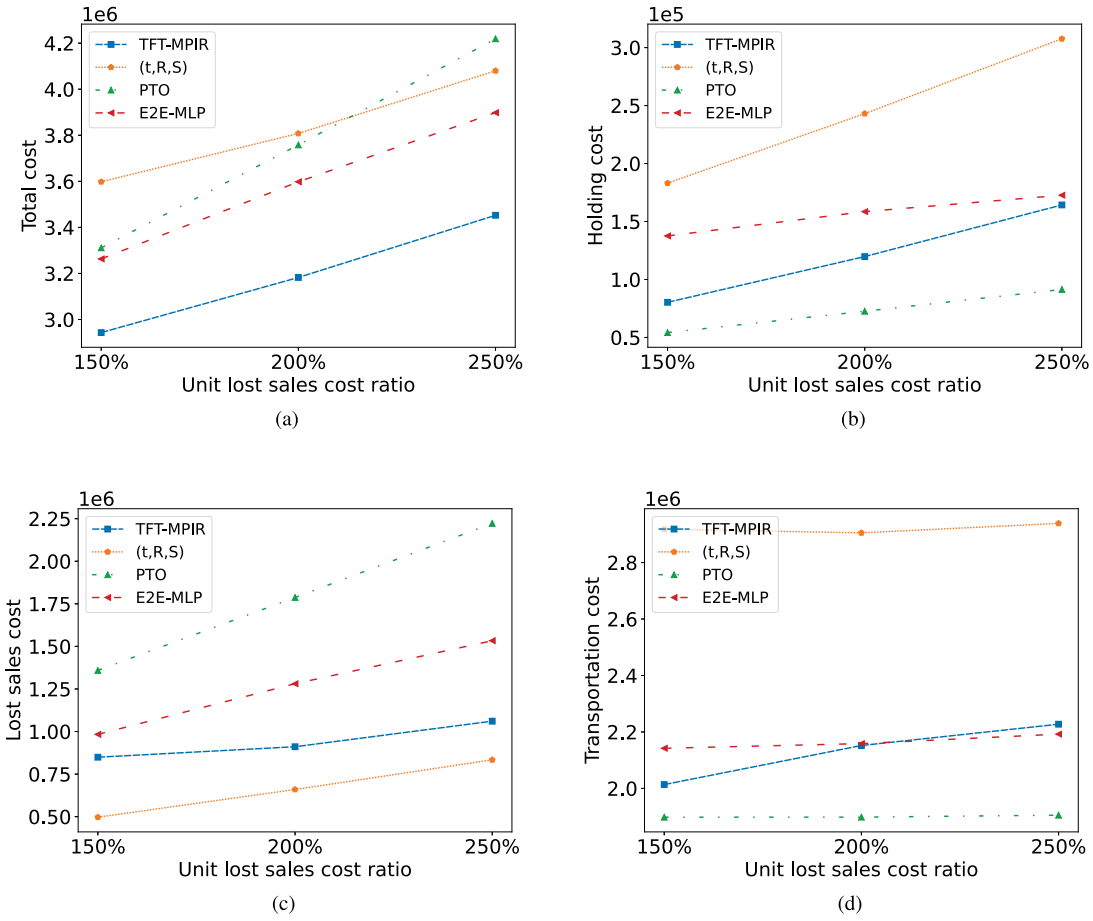


Fig. 8. Performance of each method under different ratios of unit stockout cost to unit transportation cost: (a) Comparison chart of total inventory cost for each method, (b) holding cost, (c) stockout cost, (d) transportation cost.

Table 6
Sensitivity analysis of TFT-MPIR to learning rate. (The percentage values signify the positive (+) or negative (−) growth rates, indicating how the values of the current parameter compare to those of the OPT benchmark).

	OPT	lr = 0.0001	lr = 0.001	lr = 0.005	**lr = 0.01**
Total cost	3 900 093	+18.1%	+11.1%	+9.8%	**+7.7%**
Holding cost	103 899	+13.6%	−1.5%	**−4.7%**	+14.2%
Stockout cost	2469	+62 189.9%	+53 125.1%	+53 258.1%	**+44 590.8%**
Transportation cost	3 793 725	−22.3%	−23.2%	**−24.4%**	−21.5%

Table 7
Sensitivity analysis of TFT-MPIR to TFT hidden size. (The percentage values signify the positive (+) or negative (−) growth rates, indicating how the values of the current parameter compare to those of the OPT benchmark).

	OPT	32	64	**128**
Total cost	3 900 093	+13.5%	+13.2%	**+10.4%**
Holding cost	103 899	10.7%	**+3.6%**	+3.7%
Stockout cost	2469	+58 062.5%	+53 756.1%	**+49 086.1%**
Transportation cost	3 793 725	**−24.3%**	−21.6%	21.4%

Table 8
Sensitivity analysis of TFT−MPIR to TFT attention head size. (The percentage values signify the positive (+) or negative (−) growth rates, indicating how the values of the current parameter compare to those of the OPT benchmark).

	OPT	2	**4**	8
Total cost	3 900 093	+11.1%	**+10.4%**	+11.6%
Holding cost	103 899	**−0.4%**	+3.7%	+11.7%
Stockout cost	2469	+53 125.1%	**+49 086.1%**	+51 034.8%
Transportation cost	3 793 725	**−23.2%**	−21.4%	−21.7%

the TFT-MPIR model in the test dataset. Overall, as the size of the TFT hidden layer increases from 32 to 128, there is an improvement in the model's performance. With a too-small TFT hidden layer size, the demand prediction module fails to extract sufficient features from historical sales data, leading to insufficient information transfer to subsequent neural network layers and suboptimal decision-making.

(3) Evaluating the Impact of TFT Attention Heads: In our study, we assess how the number of attention heads within the TFT−MPIR model's TFT affects its efficiency. Table 8 details the influence of different counts of attention heads on the TFT−MPIR model's performance in the test set, offering insights into how this key aspect of the model's structure impacts its overall functionality.

(4) Assessing the Impact of LSTM Hidden Layer Size: The quantity of hidden layers within the LSTM, particularly in the VLT prediction segment, plays a pivotal role. Table 9 elaborates on how variations in the size of LSTM hidden layers influence the efficiency of the

Table 9
Sensitivity analysis of TFT−MPIR to LSTM hidden size. (The percentage values signify the positive (+) or negative (−) growth rates, indicating how the values of the current parameter compare to those of the OPT benchmark).

	OPT	32	64	100	**128**
Total cost	3 900 093	+12.6%	+11.1%	+12.0%	**+10.4%**
Holding cost	103 899	+14.8%	+4.9%	**−0.8%**	+3.7%
Stockout cost	2469	+53 859.1%	+51 984.4%	+60 362.9%	**+49 086.1%**
Transportation cost	3 793 725	−22.6%	−22.6%	**−26.9%**	−21.4%

TFT−MPIR model in the test environment, highlighting the correlation between hidden layer dimensions and overall model performance.

5. Conclusion and discussion

Inventory costs form an integral part of a company's expenses. Effective inventory replenishment strategies are crucial for businesses, as they not only reduce inventory costs but also optimize capital utilization, lower risks, and bolster competitiveness and profitability. In today's context of escalating market competition and complex supply chain dynamics, the significance of inventory management, particularly in terms of replenishment decisions, cannot be overstated. It is essential for companies to implement scientifically-backed inventory replenishment methods to maintain optimal inventory levels, thereby circumventing the risks and costs linked to disproportionate inventory levels. In our study, we introduced an E2E deep learning framework for addressing multi-period inventory replenishment challenges that include transportation costs, eliminating the need for presumptions about future demand and VLT distributions. Our training objective focused on identifying replenishment order quantities that minimize the total inventory cost, encompassing transportation, holding, and stockout costs. The efficacy of our model was validated through numerical implementations using both simulated and real datasets, establishing its superiority compared to benchmark models. The key findings of our research can be summarized as follows: (1) In the simulated dataset analysis, our TFT-MPIR model successfully achieved the lowest total inventory, stockout costs, and stockout rates in the test set. Compared to the optimal post-hoc solution, the total inventory cost with TFT-MPIR was higher by 15.67%. The traditional (t, R, S) method performed poorly, showing a 40.36% higher total inventory cost and generating the most transportation costs. The E2E-MLP method resulted in a 21.78% higher total inventory cost than the optimal solution. The findings unequivocally demonstrated the exceptional efficacy of TFT-MPIR in the realm of inventory management. (2) Within the realm of actual data, in contrast to the total inventory costs ascertained by the OPT, our TFT-MPIR model in the test dataset consistently exhibits the narrowest margin of difference from OPT, at 15.8%. This performance is markedly superior to the disparities observed in the (t, R, S) method, PTO method, and E2E-MLP method when compared with OPT, which register at 34.8%, 24.1%, and 22.3%, respectively. Furthermore, a rigorous sensitivity analysis conducted on the dataset corroborated the robustness and versatility of the TFT-MPIR model. A pivotal observation emerged when altering the ratio between unit stockout and transportation costs: the primary driver of total inventory cost escalation was the increment in stockout costs. Notably, the TFT-MPIR model demonstrated a relatively gradual increase in stockout costs, thus preserving its capability to consistently attain the lowest total inventory costs across all benchmarks. This insight underscores the method's superiority, attributed to its reduction of intermediate steps, thereby minimizing information loss, averting errors in intermediary processes, and ensuring that each component of the model significantly influences decision-making, effectively safeguarding against further adverse effects on decision performance. (3) Our approach, which integrates the prediction and optimization of crucial decision variables in a modular framework, proved to be a multi-faceted model. It not only outputs recommended replenishment decisions from historical contextual data but also generates demand and VLT forecasts. These features guide the decision-making process and enable the early detection of training and prediction issues, enhancing the model's applicability and stability. Facing scenarios with diverse product categories and uncertain demand and VLT, adopting an end-to-end concept in inventory replenishment strategy represents an innovative and efficient method, minimizing manual intervention and thereby facilitating more precise inventory decisions at reduced operational costs. Future research directions outlined in this work include: (1) Expanding the scope of the E2E model by factoring in additional practical elements, such as the storage capacity limits of factories and distribution centers, and integrating varying transportation costs among multiple warehouses and distribution hubs. This will facilitate the development of a more comprehensive E2E inventory replenishment framework capable of determining optimal replenishment sources for different distribution centers. (2) While the study focused on replenishment on predetermined order days, it acknowledges that certain products may require more frequent restocking. Future efforts will aim to also determine the optimal scheduling of subsequent replenishment cycles, leading to the creation of a more holistic solution.

CRediT authorship contribution statement

Guo Weixing: Data curation, Formal analysis, Methodology, Validation, Writing – original draft, Writing – review & editing. **Ren Zhuoming:** Conceptualization, Formal analysis, Methodology, Writing – original draft, Writing – review & editing, Supervision. **Du Wenli:** Formal analysis, Validation, Writing – original draft, Writing – review & editing. **Weng Tongfeng:** Formal analysis, Methodology, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The datasets used to support the findings of this study are available from the website (https://www.coap.online/competitions/1).

Acknowledgments

The work is partially supported by National Natural Science Foundation of China (Grant Nos. 61803137 and 11805128), Scientific Research Fund of Zhejiang Provincial Education Department, China (Y202353590).

Ethical approval statements

Ethical approval was not required as the study did not involve human participants.

Informed consent statements

This article does not contain any studies with human participants performed by any of the authors.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.eswa.2024.125464.

References

Arrow, K. J., Harris, T., & Marschak, J. (1951). Optimal inventory policy. *Econometrica*, 250–272.

Babai, M. Z., Boylan, J. E., & Rostami-Tabar, B. (2022). Demand forecasting in supply chains: a review of aggregation and hierarchical approaches. *International Journal of Production Research*, *60*(1), 324–348.

Ban, G.-Y., & Rudin, C. (2019). The big data newsvendor: Practical insights from machine learning. *Operations Research*, *67*(1), 90–108.

Bhavani, G. D., Mahapatra, G. S., & Kumar, A. (2023). A sustainable two-echelon green supply chain coordination model under fuzziness incorporating carbon pricing policies. *Environmental Science and Pollution Research*, *30*(38), 89197–89237.

Bhavani, G. D., Mishra, U., & Mahapatra, G. (2023). A case study on the impact of green investment with a pentagonal fuzzy storage capacity of two green-warehouse inventory systems under two dispatching policies. *Environment, Development and Sustainability*, 1–35.

Cao, J., Li, Z., & Li, J. (2019). Financial time series forecasting model based on CEEMDAN and LSTM. *Physica A: Statistical Mechanics and its Applications*, *519*, 127–139.

Chimmula, V. K. R., & Zhang, L. (2020). Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos, Solitons & Fractals*, *135*, Article 109864.

Coelho, M. P., & Mendes, J. Z. (2019). Digital music and the "death of the long tail". *Journal of Business Research*, *101*, 454–460.

Cohen, M. C., Leung, N.-H. Z., Panchamgam, K., Perakis, G., & Smith, A. (2017). The impact of linear optimization on promotion planning. *Operations Research*, *65*(2), 446–468.

Donti, P., Amos, B., & Kolter, J. Z. (2017). Task-based end-to-end model learning in stochastic optimization. *Advances in Neural Information Processing Systems*, *30*.

Ehrhardt, R. (1984). (S, S) policies for a dynamic inventory model with stochastic lead times. *Operations Research*, *32*(1), 121–132.

Elmachtoub, A. N., & Grigas, P. (2022). Smart "predict, then optimize". *Management Science*, *68*(1), 9–26.

Elmachtoub, A. N., Liang, J. C. N., & McNellis, R. (2020). Decision trees for decision-making under the predict-then-optimize framework. In *International conference on machine learning* (pp. 2858–2867). PMLR.

Gallego, G., & Özer, Ö. (2001). Integrating replenishment decisions with advance demand information. *Management Science*, *47*(10), 1344–1360.

Graves, A., & Graves, A. (2012). Long short-term memory. In *Supervised sequence labelling with recurrent neural networks* (pp. 37–45). Springer.

Halman, N., Orlin, J. B., & Simchi-Levi, D. (2012). Approximating the nonlinear newsvendor and single-item stochastic lot-sizing problems when data is given by an oracle. *Operations Research*, *60*(2), 429–446.

He, K., Ji, L., Wu, C. W. D., & Tso, K. F. G. (2021). Using SARIMA–CNN–LSTM approach to forecast daily tourism demand. *Journal of Hospitality and Tourism Management*, *49*, 25–33.

Huber, J., Müller, S., Fleischmann, M., & Stuckenschmidt, H. (2019). A data-driven newsvendor problem: From data to decision. *European Journal of Operational Research*, *278*(3), 904–915.

Iida, T., & Zipkin, P. H. (2006). Approximate solutions of a dynamic forecast-inventory model. *Manufacturing & Service Operations Management*, *8*(4), 407–425.

James, G., Witten, D., Hastie, T., Tibshirani, R., et al. (2013). *An introduction to statistical learning*: *vol. 112*, Springer.

Kamble, S. S., & Gunasekaran, A. (2020). Big data-driven supply chain performance measurement system: a review and framework for implementation. *International Journal of Production Research*, *58*(1), 65–86.

Kaplan, R. S. (1970). A dynamic inventory model with stochastic lead times. *Management Science*, *16*(7), 491–507.

Kuo, Y.-H., & Kusiak, A. (2019). From data to big data in production research: the past and future trends. *International Journal of Production Research*, *57*(15–16), 4828–4853.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

Levi, R., Janakiraman, G., & Nagarajan, M. (2008). A 2-approximation algorithm for stochastic inventory control models with lost sales. *Mathematics of Operations Research*, *33*(2), 351–374.

Levi, R., Pál, M., Roundy, R. O., & Shmoys, D. B. (2007). Approximation algorithms for stochastic inventory control models. *Mathematics of Operations Research*, *32*(2), 284–302.

Levi, R., Roundy, R. O., Shmoys, D. B., & Truong, V. A. (2008). Approximation algorithms for capacitated stochastic inventory control models. *Operations Research*, *56*(5), 1184–1199.

Levi, R., & Shi, C. (2013). Approximation algorithms for the stochastic lot-sizing problem with order lead times. *Operations Research*, *61*(3), 593–602.

Lim, B., Arık, S. Ö., Loeff, N., & Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, *37*(4), 1748–1764.

Lin, S., Chen, Y., Li, Y., & Shen, Z.-J. M. (2022). Data-driven newsvendor problems regularized by a profit risk constraint. *Production and Operations Management*, *31*(4), 1630–1644.

Liu, C., Letchford, A. N., & Svetunkov, I. (2022). Newsvendor problems: An integrated method for estimation and optimisation. *European Journal of Operational Research*, *300*(2), 590–601.

Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). A CNN–LSTM model for gold price time-series forecasting. *Neural Computing and Applications*, *32*, 17351–17360.

Maheshwari, S., Gautam, P., & Jaggi, C. K. (2021). Role of big data analytics in supply chain management: current trends and future perspectives. *International Journal of Production Research*, *59*(6), 1875–1900.

Mandi, J., Stuckey, P. J., Guns, T., et al. (2020). Smart predict-and-optimize for hard combinatorial optimization problems. *Vol. 34*, In *Proceedings of the AAAI conference on artificial intelligence* (pp. 1603–1610).

Muharremoglu, A., & Tsitsiklis, J. N. (2008). A single-unit decomposition approach to multiechelon inventory systems. *Operations Research*, *56*(5), 1089–1103.

Mukhopadhyay, A., Vorobeychik, Y., Dubey, A., & Biswas, G. (2017). Prioritized allocation of emergency responders based on a continuous-time incident prediction model. In *Proceedings of the 16th conference on autonomous agents and multiAgent systems* (pp. 168–177). Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems.

Nazir, A., Shaikh, A. K., Shah, A. S., & Khalil, A. (2023). Forecasting energy consumption demand of customers in smart grid using temporal fusion transformer (TFT). *Results in Engineering*, *17*, Article 100888.

Neghab, D. P., Khayyati, S., & Karaesmen, F. (2022). An integrated data-driven method using deep learning for a newsvendor problem with unobservable features. *European Journal of Operational Research*, *302*(2), 482–496.

Oroojlooyjadid, A., Snyder, L. V., & Takáč, M. (2020). Applying deep learning to the newsvendor problem. *IISE Transactions*, *52*(4), 444–463.

Qi, M., Shi, Y., Qi, Y., Ma, C., Yuan, R., Wu, D., et al. (2023). A practical end-to-end inventory management model with deep learning. *Management Science*, *69*(2), 759–773.

Snyder, L. V., & Shen, Z.-J. M. (2019). *Fundamentals of supply chain theory*. John Wiley & Sons.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, *15*(1), 1929–1958.

Tian, Y.-X., & Zhang, C. (2023). An end-to-end deep learning model for solving data-driven newsvendor problem with accessibility to textual review data. *International Journal of Production Economics*, *265*, Article 109016.

Wang, H., Xie, H., Qiu, L., Yang, Y. R., Zhang, Y., & Greenberg, A. (2006). COPE: Traffic engineering in dynamic networks. In *Proceedings of the 2006 conference on applications, technologies, architectures, and protocols for computer communications* (pp. 99–110).

Wen, R., Torkkola, K., Narayanaswamy, B., & Madeka, D. (2017). A multi-horizon quantile recurrent forecaster. arXiv preprint arXiv:1711.11053.

Wu, B., Wang, L., & Zeng, Y.-R. (2022). Interpretable wind speed prediction with multivariate time series and temporal fusion transformers. *Energy*, *252*, Article 123990.

Wu, B., Wang, L., & Zeng, Y.-R. (2023). Interpretable tourism demand forecasting with temporal fusion transformers amid COVID-19. *Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies*, *53*(11), 14493–14514.

Zhang, H., Zou, Y., Yang, X., & Yang, H. (2022). A temporal fusion transformer for short-term freeway traffic speed multistep prediction. *Neurocomputing*, *500*, 329–340.