Journal of Computer Science and Electrical Engineering
Print ISSN: 2663-1938
Online ISSN: 2663-1946
DOI: https://doi.org/10.61784/jcsee3163

DEEP LEARNING BASED DETECTION OF ANOMALOUS INVENTORY FLUCTUATIONS IN INDUSTRIAL SUPPLY CHAINS

JiaoJian Pang
Warwick Manufacturing Group (WMG), The University of Warwick, Coventry CV4 7AL, United Kingdom.

Abstract: Fluctuations in inventories across industrial supply chains are caused by interplay between demand, ordering, replenishment, and production which makes it challenging to detect anomalies with just one threshold or simple distributional rules. In this paper we introduce an anomaly detection approach based on four synchronized time series (inventory, demand, ordering, and replenishment). Our model captures normal operating patterns through sliding windows and computes the difference between actual and estimated inventory level as an anomaly score. We tested our model over a 240-step simulated sequence including sudden inventory drops and persistent inventory surpluses; for this sequence the multivariate model obtained an F1 score of 0.897 and an AUC of 0.951, surpassing Isolation Forest, LSTM, GRU, and baseline Transformer models. Detected anomalies matched the injected anomaly events. Results indicate potential use of multivariate time-series modeling for inventory anomaly screening, however, further validation is needed on real-world enterprise data.
Keywords: Industrial supply chain; Inventory anomaly detection; Deep learning; Time series

1 INTRODUCTION

The choice of inventories affects production continuity, working capital, and responsiveness to demand. A lack of inventory results in production interruptions/delivery delays; excess inventory uses capital and incurs warehousing cost. Transactional data that drives inventory fluctuations are recorded by modern digital supply chains, providing a chance at detecting anomalies from operational data. With uncertainties in demand and replenishment, forecasting techniques and data-driven inventory management methods have significantly enhanced planning capabilities [1-3], but anomaly detection is harder than simply forecasting because inventory fluctuations might be normal responses to demand, orders, inbound shipments, or production schedule. A good detection system needs to incorporate all such variables with full consideration on its time series characteristic.

Rule-based systems continue to be popular for inventory control because they are inexpensive and easy to understand. Statistical detection methods and traditional machine learning techniques provide more flexibility but tend to rely on preset thresholds or certain data assumptions which might miss anomalies which become evident upon reviewing previous sequential data. LSTM (Long Short-Term Memory) models able to handle sequential dependencies have been used for demand and spare parts forecasting [4-6], models incorporating attention mechanisms and Transformer architectures can capture longer time horizons and interactions amongst multiple variables and models like TranAD and Anomaly Transformer show how those features can detect anomalies deviating from normal multivariate behavior [7-8], domestic research has also considered the use of LSTMs in supply chain forecasting [9-10]. However very little work is done analyzing inventory anomalies by correlating inventory levels with the operational variables driving their fluctuations.

In this paper we address this problem through a prediction error based detection method. It makes use of time aligned historical data on inventory, demand, orders, and replenishment as input. The model predicts expected inventory levels under normal operating conditions given this historical data and large deviations between predicted and actual values indicate potential anomalies. In this study we aim to quantify the improvement of our multivariate model compared to standard benchmarks and determine if it's scoring mechanism can simultaneously identify both short term inventory fluctuations (shocks) and long term inventory accumulation. We evaluated the method using simulated data because enterprise datasets were unavailable.

2 DATA AND METHODOLOGY

2.1 Data Description and Preprocessing

We use four synchronized sequences including inventory, demand, order quantity and replenishment quantity which describe an inventory state and its main flow changes. Fluctuations on inventory that are anomalous by themselves can be explained after large orders/receipts so we evaluate all four sequences together for our detector. In this paper, as a demonstration, we created a dataset with 240 time steps containing periodic components, random fluctuation and some artificial injection of inventory anomaly. The definitions of those variables are listed in Table 1.

© By the Author(s) 2026, under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0).

Table 1 Main Variables Used for Inventory Anomaly Detection

Variable	Description
Inventory	Available inventory level at each time step
Demand	Actual demand or outbound quantity
Order	Order quantity generated during the corresponding period
Replenishment	Quantity replenished or received into inventory

Prior to training, this paper sorted the observational data chronologically and screened the records for missing or invalid values. Given the varying magnitudes of the four variables, this study standardized each one to prevent sequences with larger numerical values x from dominating the model. For variable x, the standardized observation is calculated as follows:

$$x_t^{'} = \frac{x_t - \mu_x}{\sigma_x} \tag{1}$$

Here, μ_x and σ_x denote the mean and standard deviation of x. This study then converted the continuous records into samples with a moving window. For a window of length w, the sequence available at time t is:

$$X_t = [x_{t-w+1}, x_{t-w+2}, \dots, x_t] \tag{2}$$

Samples were assigned to training, validation, and test sets in chronological order. The training set represents normal inventory behavior; the validation set is used to choose the model and calibrate the threshold; and the test set provides the final assessment. This paper did not shuffle the observations because doing so would expose the model to future records and break the temporal structure of the monitoring problem.

2.2 Deep Learning-Based Anomaly Detection

Our method relies on an error based anomaly detection. We feed historical data for inventory, demand, orders, and replenishment in a multivariate temporal model that predicts future inventory levels. Our main detector employs a Transformer-based temporal representation with all four variables while our vanilla Transformer baseline applies this forecasting principle but without the complete multivariate anomaly-detection setup. Prediction of inventory can be formulated as:

$$\hat{I}_{t+1} = f_\theta(X_t) \tag{3}$$

where X_t represents the historical supply chain state, f_θ denotes the temporal model parameterized by θ, and $\hat{I}_{t+1}$ is the estimated inventory level. Figure 1 summarizes the overall analytical workflow, from multivariate supply chain records to anomaly-score generation.



Figure 1 Overall Framework of the Deep Learning-Based Inventory Anomaly Detection Method

Under normal operating conditions, the discrepancy between observed and predicted inventory levels is expected to remain small. When the learned relationship among demand, orders, replenishment, and inventory is disrupted, the prediction error increases. Accordingly, the anomaly score is defined by the absolute deviation between the observed inventory I_t and the predicted value $\hat{I}_t$:

$$AS_t = \left| I_t - \hat{I}_t \right| \tag{4}$$

We consider an observation to be anomalous if it has a higher anomaly score than our decision threshold τ:

$$AS_t > \tau \tag{5}$$

This score does not follow any percentage change rules. The large movement for stocks could be anticipated if there are simultaneous changes in demand and replenishment; while a small movement might be suspicious against the recent relationship between the four variables. Thus our decisions depend on the prediction error in this particular context instead of the size of inventory fluctuation.

3 RESULTS AND ANALYSIS

3.1 Inventory Temporal Characteristics

In Figure 2 we show the fluctuation of inventory and demand for 240 simulated time steps. Most of this period showed that the level of inventory stayed fairly constant while the demand had smooth cyclic fluctuations. The variation of both these metrics was correlated but not synchronized; which is expected since there exists a time interval due to order processing and replenishment before changes in demand have an effect on inventory.

Two patterns emerge from Figure 2: (i) inventory levels drop sharply at steps 68 and 205 without a corresponding surge in demand; and (ii) inventory remains unusually high from approximately step 145 to step 153 while demand stays within its normal range.

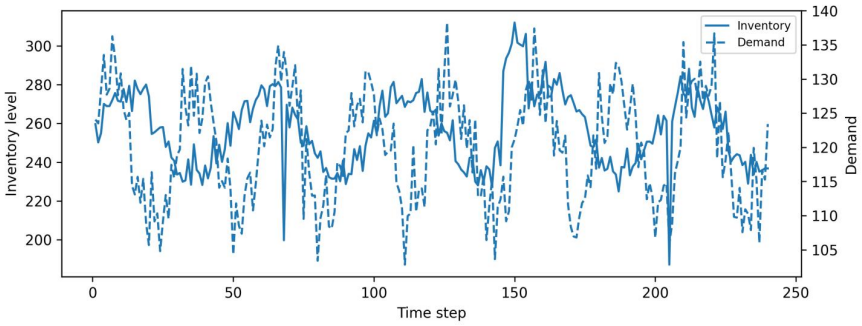


Figure 2 Temporal Variations in Inventory and Demand

3.2 Model Performance Comparison

We compared the proposed detector with Isolation Forest, LSTM, GRU, and a baseline Transformer. Isolation Forest provides a non-temporal baseline, LSTM and GRU provide recurrent sequence baselines, and the baseline Transformer provides an attention-based benchmark without the proposed multivariate detection configuration. Performance was evaluated using Precision, Recall, F1-score, and the Area Under the Receiver Operating Characteristic curve (AUC). Table 2 reports the results.

Table 2 Performance Comparison of Different Anomaly Detection Models

Model	Precision	Recall	F1-score	AUC
Isolation Forest	0.781	0.714	0.746	0.821
LSTM	0.845	0.812	0.828	0.901
GRU	0.861	0.829	0.845	0.914
Vanilla Transformer	0.887	0.851	0.869	0.932
Proposed multivariate model	0.913	0.882	0.897	0.951

Isolation Forest had the worst F1 score (0.746). Both recurrent models outperformed Isolation Forest: LSTM scored 0.828 while GRU reached 0.845. Baseline Transformer increased the F1 score up to 0.869 and the AUC up to 0.932. In short, on this simulated sequence, approaches using temporal order found inserted anomalies more accurately compared to baseline approaches disregarding temporal factors.

On this simulated sequence, the multivariate detector achieved a precision of 0.913, a recall of 0.882, an F1 score of 0.897, and an AUC of 0.951—all of them representing the best values in Table 2. Its F1 score is 0.028 higher and its AUC is 0.019 higher when compared against the baseline Transformer. Even though differences between both were minor, all trends regarding those four metrics were similar. This result confirmed our premise that demand, order, and replenishment data help the model determine if inventory responses are normal.

3.3 Inventory Anomaly Detection Results

In Figure 3 we show how it looks on an event basis: a plot of observed inventory versus the model-estimated normal inventory. In normal times they are similar but if certain events happen then there is deviation in the observed sequence compared with the estimated ones that will be flagged by our detector.

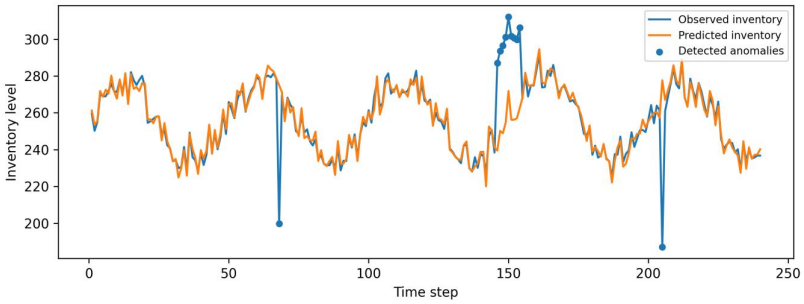


Figure 3 Detection Results of Anomalous Inventory Fluctuations

The drops at steps 68 and 205 produced isolated anomaly detections, whereas the high-inventory period from steps 145 to 153 produced consecutive detections.

Anomaly scores shown in Figure 4 also reflect this phenomenon. For most data points, scores remained under the decision threshold. Two isolated sudden drops gave rise to narrow spikes while scores during the high-inventory period remained over the threshold across several consecutive time steps. Consistency between the trajectory plot and the score plot enables us to precisely localize every alert within the original sequence.

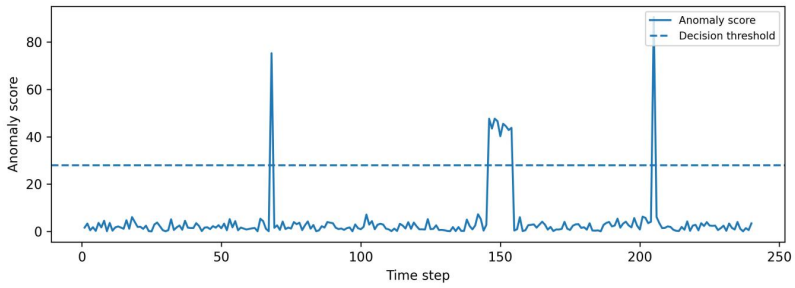


Figure 4 Anomaly Score and Decision Threshold

Table 3 links the three detected events to issues that an inventory analyst might need to investigate. The items listed represent only possible operational explanations; while the scoring identifies anomalous patterns, it cannot determine their root causes.

Table 3 Representative Cases of Detected Inventory Anomalies

Case	Time step	Observed pattern	Demand condition	Possible interpretation
A1	68	Sudden inventory drop	Normal range	Unexpected shortage, recording error, or emergency outbound activity
A2	145-153	Sustained high inventory	Mild fluctuation	Excess replenishment, delayed dispatch, or temporary stock accumulation
A3	205	Abrupt abnormal decline	Slight increase	Process disruption, stock loss, or an execution fault

4 CONCLUSION

This paper evaluates a prediction-based multivariate method for detecting anomalous inventory fluctuations. The method predicts the next inventory level from recent inventory, demand, order, and replenishment records, then uses the absolute prediction error as the anomaly score. In the simulation experiments, the sequence-based models outperformed Isolation Forest, and the proposed configuration achieved the highest values across all four evaluation metrics. It detected both isolated inventory drops and prolonged periods of excess inventory.

We rely exclusively on synthetic generation of data sequences. In future work, we plan to repeat these experiments using actual business data, which might affect our results due to data quality or operational changes, and the scarcity of confirmed anomaly events. Furthermore, extending our study to multi-echelon inventory networks and incorporating other variables like supplier reliability, lead times, and external demand signals are possible directions. Graph-based models, online update mechanism, and adaptive threshold setting are interesting areas worth exploring although they need to be evaluated against simpler methods using the same business data.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Cui L, Chen Y, Deng J, et al. A novel attLSTM framework combining the attention mechanism and bidirectional LSTM for demand forecasting. Expert Systems with Applications, 2024, 254: 124409.
[2] Douaioui K, Oucheikh R, Benmoussa O, et al. Machine learning and deep learning models for demand forecasting in supply chain management: a critical review. Applied System Innovation, 2024, 7(5): 93.
[3] Guo W, Ren Z, Du W, et al. TFT-MPIR: an end-to-end multi-period inventory replenishment strategy based on temporal fusion transformer. Expert Systems with Applications, 2025, 261: 125464.
[4] Kang H, Kang P. Transformer-based multivariate time series anomaly detection using inter-variable attention mechanism. Knowledge-Based Systems, 2024, 290: 111507.
[5] Lim B, Arik S O, Loeff N, et al. Temporal Fusion Transformers for interpretable multi-horizon time series forecasting. International Journal of Forecasting, 2021, 37(4): 1748-1764.
[6] Seyedan M, Mafakheri F, Wang C. Order-up-to-level inventory optimization model using time-series demand forecasting with ensemble deep learning. Supply Chain Analytics, 2023, 3: 100024.
[7] Tuli S, Casale G, Jennings N R. TranAD: deep Transformer networks for anomaly detection in multivariate time series data. Proceedings of the VLDB Endowment, 2022, 15(6): 1201-1214.
[8] Xu J, Wu H, Wang J, et al. Anomaly Transformer: time series anomaly detection with association discrepancy. International Conference on Learning Representations, 2022.

[9] Yu C, She C, Li J, et al. Research on measuring and predicting the carbon footprint of supply chain throughout its life cycle based on LSTM. Industrial Engineering Journal, 2024, 27(5): 161-171.
[10] Zhou Z, Jiang J. Research on demand forecasting of rarely-used spare parts based on deep artificial neural network. Journal of Hebei University of Water Resources and Electric Engineering, 2021, 31(4): 60-65.