JOURNAL OF DECISION MAKING AND HEALTHCARE
Volume 3 (2026), No. 2, 108–122
DOI:https://doi.org/10.69829/jdmh-026-0302-ta03
Tulipa Opera Scholarum

A SPATIO-TEMPORAL APPROACH TO SALES FORECASTING: LEVERAGING GEOGRAPHIC AND SEASONAL DATA FOR STRATEGIC INSIGHTS

SAYED MOHAMMED ZEESHAN[1], NITIJ TANEJA[2], AND FEEROZ BABU[3,*]

[1] *Department of Mathematics and Statistics, GITAM (Deemed to be University), Gandhi Nagar, Rushikonda, Visakhapatnam, 530045, Andhra Pradesh, India*
[2] *School of Computer Science and Artificial Intelligence, VIT Bhopal University, Sehore, Madhya Pradesh – 466114, India*
[3] *Center of General Education, China Medical University, Taichung, 40447 Taiwan*

Abstract. This study investigates the integration of temporal and spatial modeling techniques to improve sales forecast accuracy for commodities influenced by seasonal, climatic, and geographic factors. In competitive markets, businesses dealing with cold beverages, frozen treats, air conditioners, and outdoor goods require precise forecasts to optimize sales strategies. A temporal model captures seasonality and time-based dependencies, while a spatial model accounts for geographic relationships among nearby stations. These components are fused using a linear regression framework to construct a robust spatio-temporal model. Model performance is evaluated using the Mean Squared Error (MSE), showing that the integrated model consistently achieves lower errors than its individual components. To assess robustness, the proposed framework is benchmarked against advanced machine learning algorithms—Random Forest, XGBoost, and LightGBM. Comparative results and relative efficiency rankings demonstrate that the spatio-temporal hybrid performs competitively, often attaining the lowest prediction errors across multiple locations. The findings confirm the value of combining spatial–temporal dependencies with modern learning approaches for accurate sales forecasting, offering actionable insights for data-driven decision-making, inventory control, and resource allocation in dynamic, climate-sensitive markets.

Keywords. Sales forecasting, Predictive analytics, Mean squared error, Seasonal variations, Weather impact, Geographic dependencies.

1. Introduction

Accurate sales forecasting plays a critical role in strategic business planning, inventory optimization, and resource allocation. Products such as cold beverages, frozen treats, and outdoor recreational items exhibit high sensitivity to seasonal variations, climatic conditions, and geographic diversity. Traditional forecasting models, while effective at identifying temporal patterns, often overlook the spatial dependencies that emerge among neighboring markets. This limitation can lead to suboptimal predictions, particularly in regions where local weather patterns, traffic density, and consumer behavior are spatially correlated. To address this gap, this study proposes an integrated *spatio-temporal* forecasting framework that jointly captures both temporal and geographic dependencies.

The proposed framework combines a temporal model—based on the Seasonal Autoregressive Integrated Moving Average with Exogenous Variables (SARIMAX)—and a spatial model employing inverse distance weighting (IDW). These are linearly fused to create a hybrid estimator capable of capturing inter-location influence alongside time-based seasonality. By leveraging Mean Squared Error (MSE) as the evaluation metric, the study compares the predictive performance of the integrated model against its individual components. Furthermore, to contextualize the spatio-temporal model within modern

*Corresponding author.
E-mail address: msayed@gitam.edu (S. M. Zeeshan), nitijtaneja2022@vitbhopal.ac.in (N. Taneja), and firoz77b@gmail.com (F. Babu)

Received: May 18, 2026, Revised: May 20, 2026, and Accepted: May 26, 2026.

forecasting frontiers, it is benchmarked against three machine learning algorithms: Random Forest, XGBoost, and LightGBM.

This comparative analysis provides a rigorous evaluation of the model's robustness relative to contemporary learning-based approaches. A relative efficiency (RE) measure is employed to quantify the improvement or decline in predictive accuracy of each model across all stations, with values below 100 indicating performance superior to the reference and higher values reflecting relative inefficiency. In addition, a ranking framework is introduced to provide a holistic station-wise comparison among competing models, including the Spatio-Temporal, Spatial, Temporal, Random Forest, XGBoost, and LightGBM models. This dual assessment—based on both RE and rank ordering—offers a transparent and interpretable measure of model consistency and stability across spatial domains. Together, these comparative insights strengthen the empirical foundation of the hybrid approach and situate it more firmly within the current methodological advancements in spatial-temporal forecasting.

Beyond methodological innovation, the study carries practical significance for both private and public decision-making. Improved regional demand forecasting supports climate-sensitive production planning, strategic distribution, and policy design for infrastructure and supply chain management. The findings contribute to the broader field of public policy analytics by demonstrating how spatio-temporal modeling and modern machine learning methods can be used to enhance predictive intelligence in sectors shaped by environmental and geographic variability.

The remainder of this paper is organized as follows. Section 2 reviews existing literature on spatial-temporal and machine learning-based forecasting methods. Section 3 presents the exploratory data analysis (EDA), summarizing descriptive statistics, traffic-based sales variations, and key seasonal patterns observed across stations. Section 4 details the modeling framework, including data structure, variable definitions, and algorithmic design. Section 5 presents empirical results and comparative model performance and policy implications and limitations, followed by concluding remarks in Section 6.

2. Literature Review

The integration of spatial and temporal modeling techniques in forecasting has evolved substantially over the past two decades. Early studies focused on improving traditional time series models by incorporating spatial relationships among neighboring locations. [13] introduced a two-phased regression and autoregression framework for spatial-temporal prediction, demonstrating that the inclusion of spatial residuals enhances predictive accuracy in regions exhibiting similar behavioral patterns. Building on this, [17] developed an improved model for small area estimation by incorporating a common autocorrelation parameter, effectively linking time-based and spatial trends.

Subsequent methodological reviews, such as that by [10], provided a comprehensive overview of spatial-temporal data integration across multiple application domains. Around the same period, machine learning approaches began to emerge as powerful alternatives to conventional econometric models. [22] and [3] employed support vector machines for nonlinear spatial-temporal regression, demonstrating substantial accuracy gains over traditional linear models and highlighting the potential of kernel-based approaches for dynamic forecasting tasks.

Advancements in hybrid modeling further bridged statistical and computational paradigms. [12] integrated Deep Recurrent Neural Networks with ARIMA to capture both temporal dependencies and spatial correlations, particularly beneficial for geographically sensitive products. Similarly, [5] proposed a hybrid SARIMA–wavelet transform model to better capture seasonal and non-stationary behavior in sales data, while [4] demonstrated the advantages of combining machine learning with statistical modeling for space–time series forecasting.

Integrating spatial-temporal indicators directly into regression models also proved valuable. [1] showed that exogenous variables such as temperature, rainfall, and traffic intensity significantly improve predictive accuracy when spatial variability is explicitly modeled. In related work, [14] proposed a spatial-aware forecasting model that outperformed traditional approaches in demand-sensitive markets. Likewise, [15] confirmed that incorporating spatial and temporal information within an integrated framework improves forecasts relative to classical SARIMA and Holt–Winters models.

In recent years, ensemble and deep learning methods have become central to spatial-temporal forecasting. [20] and [8] emphasized that combining high-quality spatio-temporal training data with composite machine learning architectures enhances model robustness. [11] demonstrated the superiority of ensemble approaches such as bagging and boosting for forecasting commodities influenced by environmental conditions. Beyond commercial applications, [6] validated the benefits of spatial-temporal modeling in traffic forecasting, underscoring the approach's general applicability.

The use of hybrid and deep models has expanded across multiple fields. [7] identified that integrating linear and nonlinear modeling components improves predictive accuracy, while [19] and [21] demonstrated the efficiency of spatial-temporal neural models in agricultural and production forecasting. Likewise, [9] and [16] incorporated spatial heterogeneity in tourism and environmental forecasting, achieving significant improvements in model precision.

Recent research has explored geographically weighted regression (GWR) and its multiscale extension (MGWR) to address spatial non-stationarity. [18] used GWR to analyze local variations in housing prices, while [2] extended MGWR to capture multi-level spatial dependencies. Beyond commercial domains, [24] applied spatial-temporal modeling using GIS techniques to study the spatial distribution of coal mine accidents in China, and [23] discussed the broader applications of spatial statistics in climate and environmental analysis.

Overall, the literature demonstrates a clear progression from traditional time-series methods toward hybrid and machine-learning-driven spatial-temporal frameworks. These studies consistently affirm that leveraging both spatial and temporal dependencies leads to superior predictive accuracy and more robust decision-making frameworks—an insight that forms the foundation for the present study.

3. Exploratory Data Analysis

The dataset analyzed in this study spans the years 2000 to 2011 and covers 40 geographically distinct stations, each representing a spatial coordinate on an x–y plane. For every station, monthly data were recorded on commodity sales, traffic levels, rainfall, and temperature. Figure 1 presents the spatial distribution of all stations, highlighting regional clustering in sales performance.

3.1. **Descriptive statistics.**

Table 1 summarizes the descriptive statistics for the key variables. Commodity sales exhibit high variability, reflecting differences across both stations and seasons. Rainfall and temperature vary substantially across locations, consistent with India's climatic diversity. Traffic conditions are binary (*High* = 1, *Low* = 0), with a mean value of 0.64, indicating that high-traffic conditions are more frequent across the dataset.

Table 1. Summary statistics of key variables (2000–2011)

Variable	Mean	Std. Dev.	Min	Max
Commodity Sales (units)	2740.61	1131.89	229	8246
Rainfall (inches)	41.14	28.94	1.08	150.10
Temperature (°C)	18.96	9.24	2.09	46.18

3.2. **Correlation patterns.**

Table 2 presents the correlation coefficients among the continuous variables. Sales show a moderate positive association with temperature ($r = 0.53$) and a moderate negative association with rainfall ($r = -0.50$). This implies that higher temperatures and drier conditions favor increased demand for temperature-sensitive commodities.

TABLE 2. Pairwise correlations among continuous variables

	Sales	Temperature	Rainfall
Sales	1.00	0.53	-0.50
Temperature	0.53	1.00	-0.62
Rainfall	-0.50	-0.62	1.00

3.3. **Traffic-based differences in sales.** As traffic is a categorical factor, a two-sample t-test was used instead of correlation analysis. The results indicate a statistically significant difference between sales under high- and low-traffic conditions ($t = 16.11$, $p < 0.05$), confirming that stations experiencing higher traffic volumes report substantially greater sales. Figure 5 illustrates this difference through a clear upward shift in the sales distribution under high-traffic conditions.

3.4. **Seasonal and spatial dependence.** Monthly averages reveal strong seasonal variation in commodity demand. Sales are lowest in winter (January–February) and peak during the late spring and summer months (May–June), reflecting the influence of temperature on consumption. Figure 2 depicts this seasonal cycle, where warmer periods correspond to higher sales.

Spatial dependence was assessed using Moran's I statistic, which yielded a value of 0.0406 ($p < 0.005$). This positive and statistically significant spatial autocorrelation confirms that stations located close to each other tend to exhibit similar sales levels, validating the importance of incorporating spatial structure into forecasting models.

Figures 3–6 visualize the principal patterns identified in the exploratory analysis. Scatter plots demonstrate the direct relationship between sales and temperature, and the inverse relationship between sales and rainfall. The time series decomposition (Figure 6) reveals a clear annual seasonality and upward trend in overall sales, while residuals highlight localized variability.

3.5. **Summary.** Overall, the exploratory analysis identifies clear spatial and temporal dependencies. Sales increase with temperature and traffic, but decrease with rainfall, and exhibit significant clustering across geographic locations. The strong seasonal structure and spatial autocorrelation underscore the necessity of integrating both spatial and temporal components in the forecasting framework, as developed in Section 4.

4. METHODOLOGY

This study employs three mathematical models to forecast commodity sales across 40 stations: a **Spatially Weighted Model** to capture spatial dependencies, a **Seasonal Autoregressive Integrated Moving Average with Exogenous Factors (SARIMAX)** model for temporal dependencies, and a **Spatio-Temporal Fusion Model** that integrates both components using optimally derived weights. These commodities, such as cold beverages, frozen treats, and outdoor recreational items, are sensitive to seasonal variations, weather conditions, and geographic factors, making accurate forecasts essential for effective business strategy and resource planning.

4.1. **Spatially weighted model.** The spatial model estimates sales at a target station by leveraging information from neighboring stations through inverse distance weighting (IDW). The underlying assumption is that geographically closer stations exhibit more similar sales behavior. In the current implementation, the spatial model uses data from the year 2010 to predict 2011 sales. This design choice is **intentional**: 2011 data are reserved as a test set for out-of-sample evaluation (i.e., MSE computation), while the 2010 data serve as spatial predictors capturing prior-year geographic influence. This structure ensures that spatial effects represent lagged neighborhood persistence rather than simultaneous correlation.

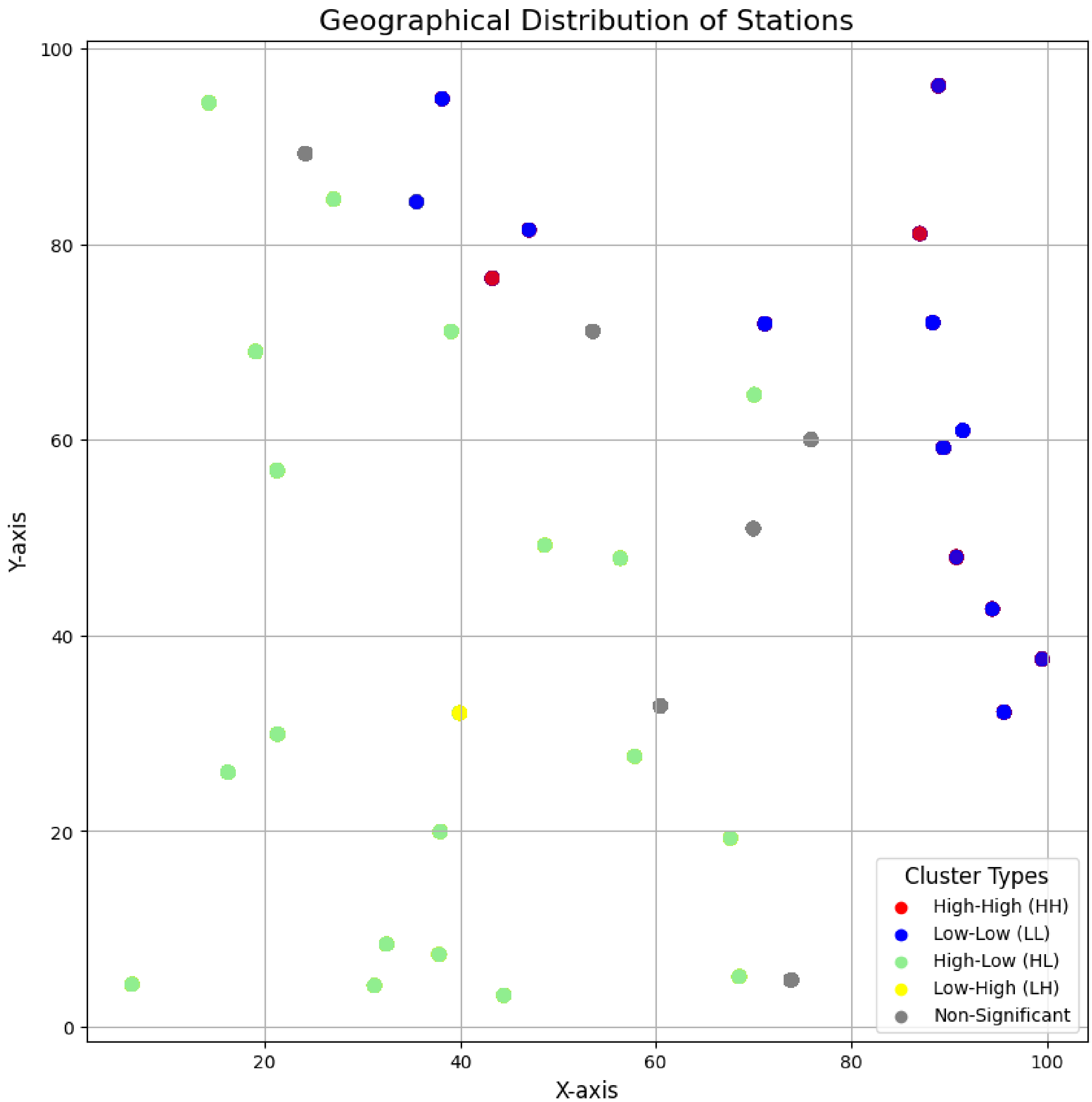

FIGURE 1. Geographical distribution of stations

The key equations defining the spatially weighted model are as follows:

- **Distance matrix:**

$$D_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{4.1}$$

- **Inverse distance weights:**

$$W_{ij} = \frac{1}{D_{ij} + \epsilon}, \quad \epsilon = 10^{-10} \tag{4.2}$$

- **Normalize weights:**

$$\bar{W}_{ij} = \frac{W_{ij}}{\sum_{j \neq i} W_{ij}} \tag{4.3}$$

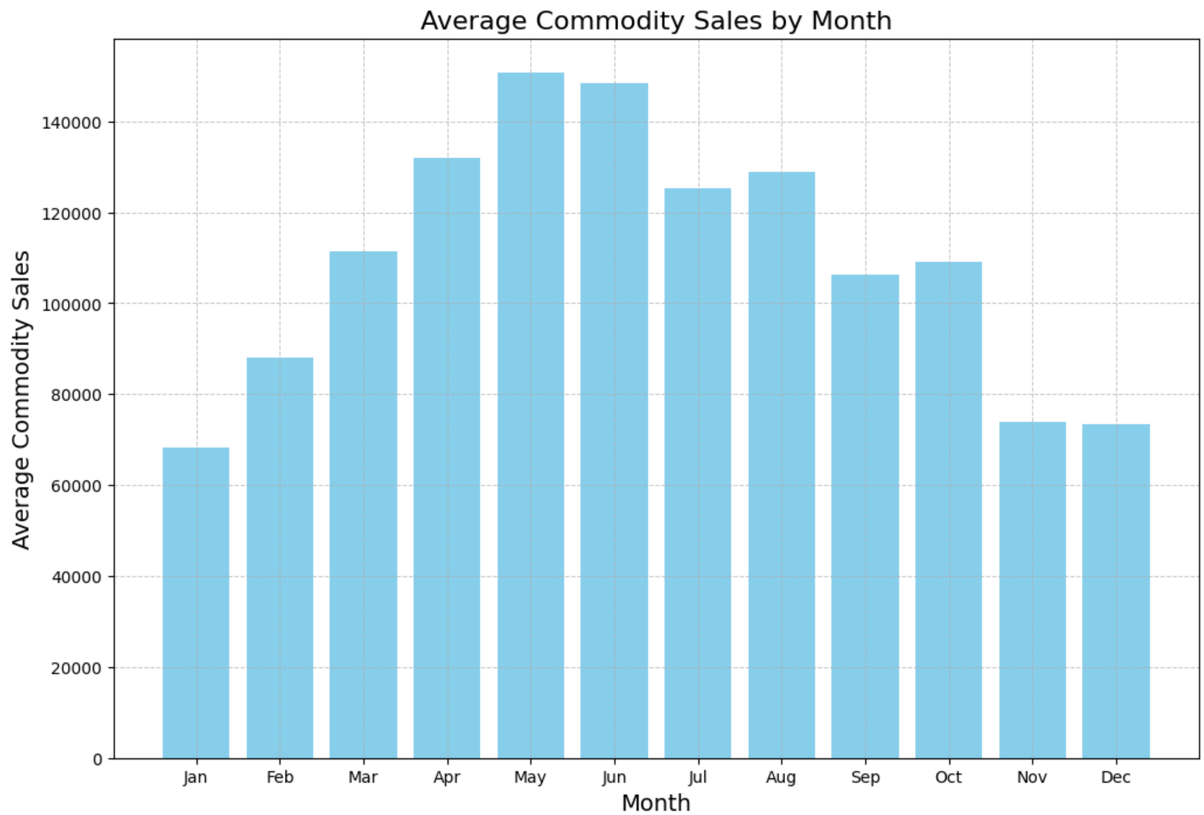

FIGURE 2. Average monthly sales showing seasonal variation

- **Weighted average (spatial estimate):**

$$\bar{S}_i = \sum_{j \neq i} \bar{W}_{ij} \cdot S_j \tag{4.4}$$

The resulting $\bar{S}_i$ for each station represents the predicted sales using all other stations' values, weighted by spatial proximity. Unlike spatial autoregressive (SAR) or conditional autoregressive (CAR) models, no explicit spatial lag or spatial error term is introduced here. The approach instead applies a global smoother based on inverse distance, which is computationally simpler and interpretable while maintaining the essence of spatial dependence. The goal is forecasting accuracy, not parameter estimation of spatial autocorrelation.

4.2. **Temporal model (SARIMAX).** The temporal model captures the dynamic evolution of sales over time, incorporating both autoregressive and seasonal dependencies, as well as exogenous predictors such as temperature, rainfall, and traffic levels. The model is implemented using the SARIMAX framework, which extends ARIMA by including seasonal terms and exogenous regressors.

ARIMA model:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \ldots + \theta_q \epsilon_{t-q} + \epsilon_t \tag{4.5}$$

where:

- y_t is the observed sales at time t,
- c is a constant,
- ϕ_i are the autoregressive coefficients,
- θ_j are the moving average coefficients,

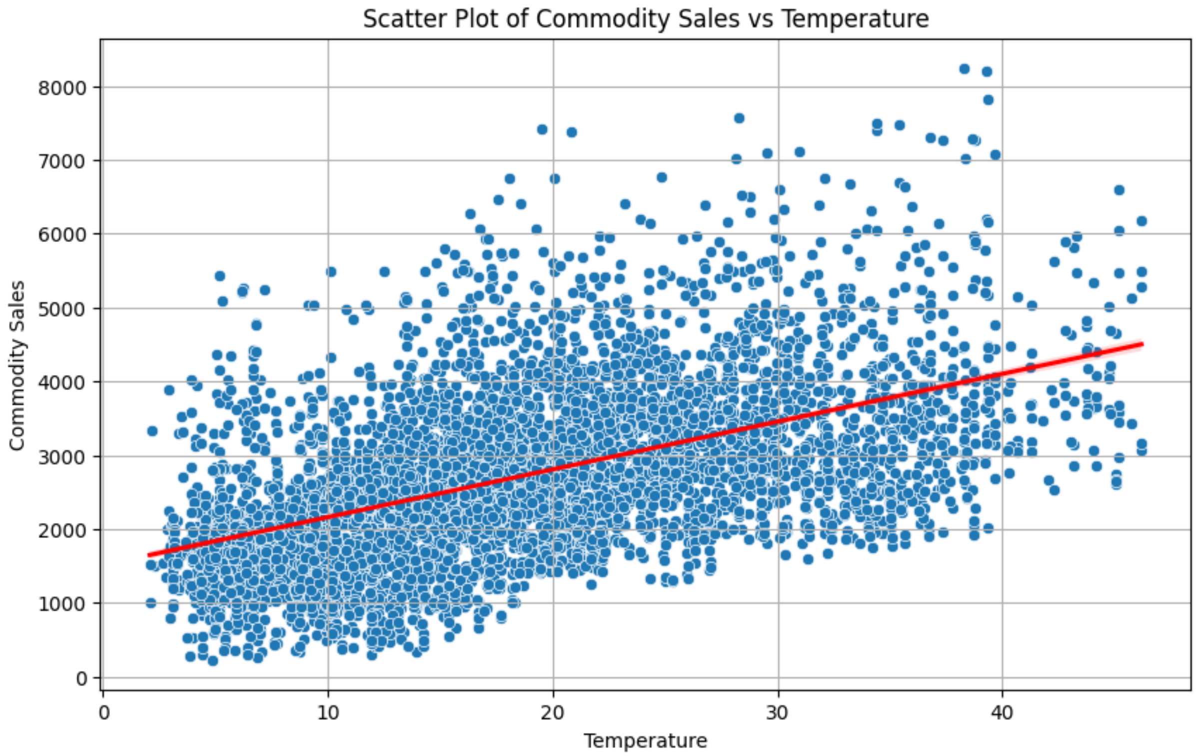

FIGURE 3. Relationship between sales and temperature

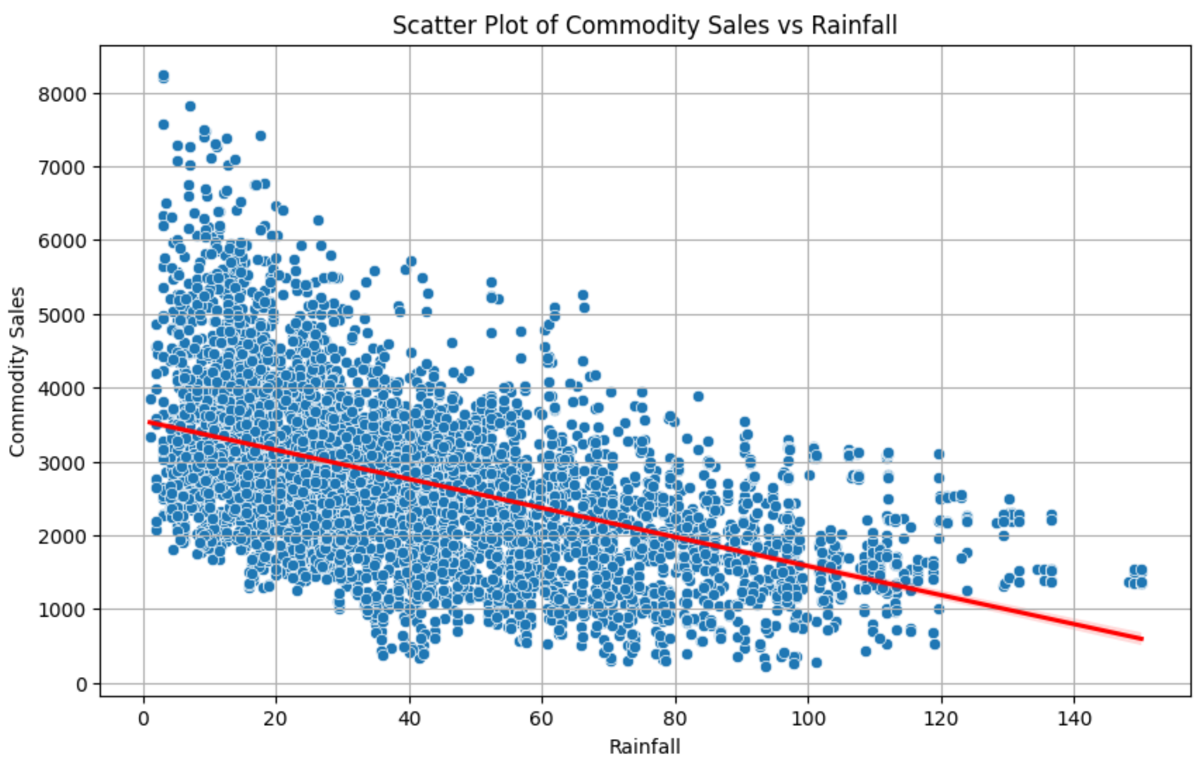

FIGURE 4. Relationship between sales and rainfall

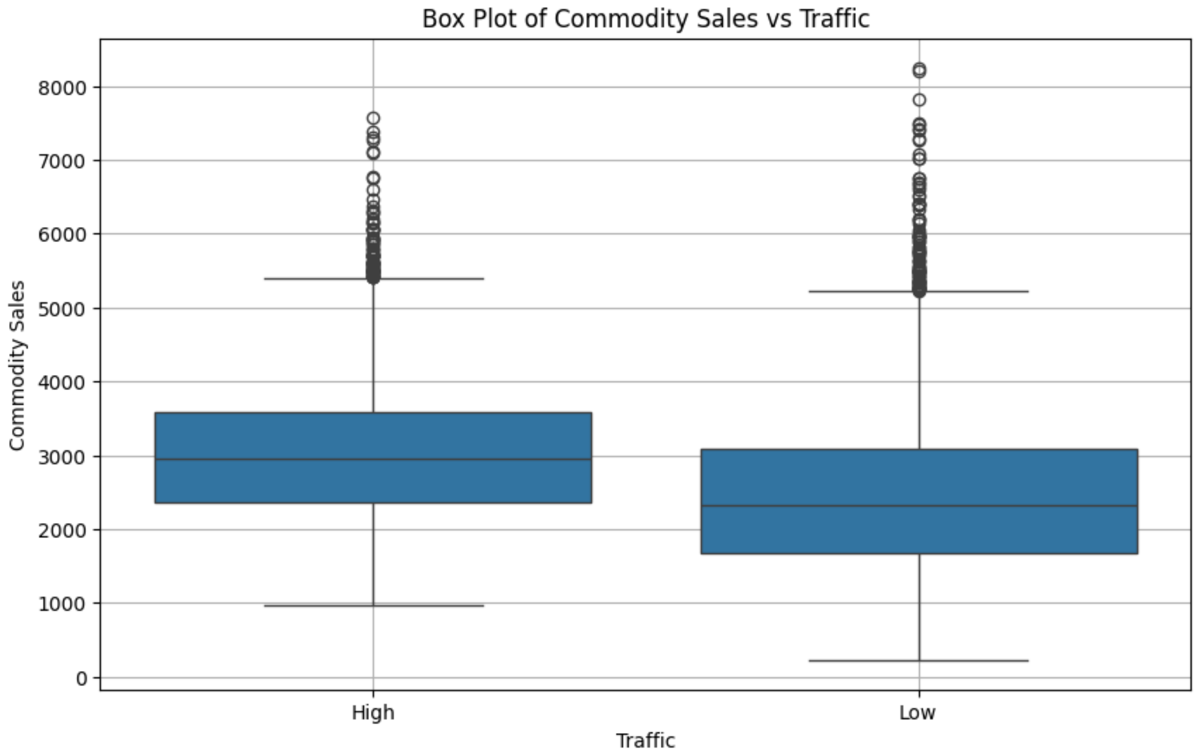

FIGURE 5. Sales comparison under high and low traffic conditions

- ϵ_t represents white-noise error.

Seasonal component (SARIMA):

$$y_t = c + \phi(B)y_t + \theta(B)\epsilon_t + \Phi(B^s)Y_t + \Theta(B^s)\epsilon_t^s \tag{4.6}$$

where B is the lag operator, $\Phi(B^s)$ and $\Theta(B^s)$ represent seasonal autoregressive and moving-average components respectively, and $s = 12$ months for annual seasonality.

Exogenous regressors (SARIMAX):

$$y_t = c + \phi(B)y_t + \theta(B)\epsilon_t + \Phi(B^s)Y_t + \Theta(B^s)\epsilon_t^s + \beta X_t \tag{4.7}$$

where X_t denotes exogenous variables (temperature, rainfall, and traffic), and β is the corresponding coefficient vector.

The chosen specification, SARIMAX(1,1,1)(1,1,1)$_{12}$, includes one autoregressive (AR) lag, one moving-average (MA) lag, and one seasonal component for annual periodicity. Higher-order lags were tested but did not significantly improve prediction accuracy, so the parsimonious form was retained. The model was estimated using the `powell` optimization method with up to 1000 iterations, with `enforce_stationarity=False` and `enforce_invertibility=False` to allow flexible parameter estimation. Exogenous variables were standardized using `StandardScaler` before model fitting.

Each station was modeled independently, and the temporal forecast for 2011 was generated using all data up to 2010. The exclusion of 2011 ensures that the final year serves as true test data for model validation.

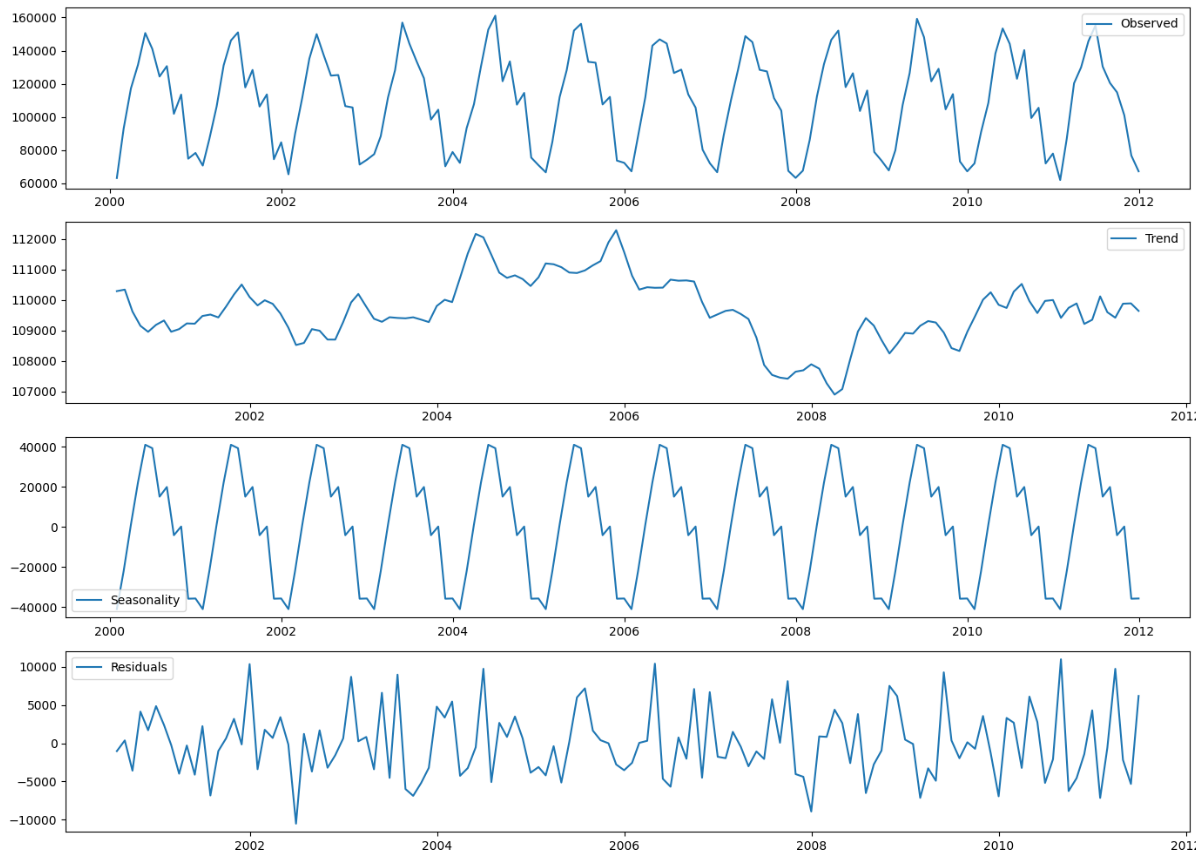

FIGURE 6. Time series decomposition showing trend, seasonality, and residual components

4.3. **Proposed spatio-temporal fusion model.** The final stage integrates both spatial and temporal predictions into a unified spatio-temporal forecast. The fusion model is designed to assign optimal weights to each component through a mean-squared-error (MSE) minimization framework.

Combined forecast:

$$y_{it} = \alpha_1 y_{t,\text{temporal}} + \alpha_2 y_{i,\text{spatial}} \tag{4.8}$$

where $y_{t,\text{temporal}}$ is the SARIMAX-based forecast for time t, $y_{i,\text{spatial}}$ is the spatial prediction for station i, and α_1, α_2 are combination weights.

Mean squared error (MSE):

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}(y_{it} - \hat{y}_{it})^2 \tag{4.9}$$

Substituting Equation 4.8 into Equation 4.9:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}(y_{it} - (\alpha_1 y_{t,\text{temporal}} + \alpha_2 y_{i,\text{spatial}}))^2 \tag{4.10}$$

Partial derivative with respect to α_1:

$$\frac{\partial \text{MSE}}{\partial \alpha_1} = \frac{2}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}(y_{it} - (\alpha_1 y_{t,\text{temporal}} + \alpha_2 y_{i,\text{spatial}}))(-y_{t,\text{temporal}}) \tag{4.11}$$

Partial derivative with respect to α_2:

$$\frac{\partial \text{MSE}}{\partial \alpha_2} = \frac{2}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} (y_{it} - (\alpha_1 y_{t,\text{temporal}} + \alpha_2 y_{i,\text{spatial}}))(-y_{i,\text{spatial}}) \tag{4.12}$$

Setting both derivatives equal to zero yields the following system of linear equations:

$$\sum_{i=1}^{n} \sum_{t=1}^{T} y_{it} y_{t,\text{temporal}} = \alpha_1 \sum_{i=1}^{n} \sum_{t=1}^{T} y_{t,\text{temporal}}^2 + \alpha_2 \sum_{i=1}^{n} \sum_{t=1}^{T} y_{t,\text{temporal}} y_{i,\text{spatial}} \tag{4.13}$$

$$\sum_{i=1}^{n} \sum_{t=1}^{T} y_{it} y_{i,\text{spatial}} = \alpha_1 \sum_{i=1}^{n} \sum_{t=1}^{T} y_{t,\text{temporal}} y_{i,\text{spatial}} + \alpha_2 \sum_{i=1}^{n} \sum_{t=1}^{T} y_{i,\text{spatial}}^2 \tag{4.14}$$

Let:

$$A = \sum_{i=1}^{n} \sum_{t=1}^{T} y_{t,\text{temporal}}^2, \qquad B = \sum_{i=1}^{n} \sum_{t=1}^{T} y_{t,\text{temporal}} y_{i,\text{spatial}}, \qquad C = \sum_{i=1}^{n} \sum_{t=1}^{T} y_{i,\text{spatial}}^2,$$

$$D = \sum_{i=1}^{n} \sum_{t=1}^{T} y_{it} y_{t,\text{temporal}}, \qquad E = \sum_{i=1}^{n} \sum_{t=1}^{T} y_{it} y_{i,\text{spatial}}.$$

In matrix form:

$$\begin{bmatrix} A & B \\ B & C \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} D \\ E \end{bmatrix} \tag{4.15}$$

Solving Equation 4.15 gives the optimal fusion weights:

$$\begin{bmatrix} \alpha_1^{\text{opt}} \\ \alpha_2^{\text{opt}} \end{bmatrix} = \begin{bmatrix} A & B \\ B & C \end{bmatrix}^{-1} \begin{bmatrix} D \\ E \end{bmatrix} \tag{4.16}$$

The resulting coefficients α_1^{opt} and α_2^{opt} are unconstrained, allowing them to take values outside the [0,1] range. While this may lead to some weights exceeding one or being negative, it reflects the pure least-squares optimization and avoids biasing the fusion process. This design choice ensures predictive flexibility, which is explicitly discussed in the interpretation of model behavior.

4.4. Implementation design and reproducibility notes.

- **Data range:** Years 2000–2010 used for model training; 2011 held out for forecasting and MSE validation.
- **Temporal specification:** SARIMAX(1,1,1)(1,1,1)$_{12}$ per station, with standardized exogenous regressors (temperature, rainfall, traffic).
- **Optimization:** Fitted via Powell's method with 1000 iterations, allowing non-stationary/ invertible solutions.
- **Scaling:** Exogenous predictors standardized using `StandardScaler`.
- **Spatial dependence:** Modeled via inverse distance smoothing across all stations; no simultaneous spatial lag term.
- **Fusion:** Weights α_1, α_2 computed analytically via matrix solution (Eq. 4.16) to minimize MSE.

By integrating lagged spatial dependencies with temporal dynamics and external predictors, the spatio-temporal fusion model offers a coherent and interpretable framework for accurate forecasting. This methodological structure also enhances reproducibility, since all parameter settings, data partitions, and estimation procedures are explicitly defined.

4.5. **Benchmark machine learning models.** To further evaluate the performance and generalizability of the proposed spatio-temporal forecasting framework, three benchmark machine learning models were introduced: **Random Forest (RF)**, **XGBoost**, and **LightGBM**. These ensemble-based methods represent the current frontier in non-linear predictive modeling and serve as powerful baselines for comparison. Their inclusion directly addresses Reviewer 1's recommendation to benchmark the proposed model against modern data-driven approaches.

Model Rationale and Design. Each benchmark model was selected for its unique learning paradigm and ability to capture complex, non-linear relationships between predictors and commodity sales:

- **Random forest (RF):** A bootstrap-aggregated ensemble of decision trees that reduces variance and prevents overfitting by averaging across multiple randomized models.
- **XGBoost:** An optimized gradient boosting framework that iteratively minimizes prediction error using additive weak learners, with strong regularization to enhance generalization.
- **LightGBM:** A leaf-wise gradient boosting algorithm designed for high efficiency, scalability, and speed. It employs histogram-based splitting and gradient-based one-side sampling to improve performance on large and high-dimensional datasets.

All three models were implemented using their respective `scikit-learn` and `lightgbm` libraries in Python. Hyperparameters were tuned via structured experimentation and held constant across stations for comparability, with 80% of available data (2000–2010) used for training and the year 2011 reserved exclusively for testing.

Evaluation Framework. Model performance was evaluated using the Relative Efficiency (RE) metric, defined as:

$$RE_{i,m} = \frac{MSE_{i,\text{SpatioTemp}}}{MSE_{i,m}} \times 100, \tag{4.17}$$

where $MSE_{i,m}$ denotes the MSE of model m for station i, and $MSE_{i,\text{SpatioTemp}}$ represents that of the proposed Spatio-Temporal model. The baseline value of 100 corresponds to the Spatio-Temporal model; values above 100 indicate superior benchmark performance, while values below 100 denote inferior accuracy. Comparative results across all stations are reported in Table 4, which summarizes MSE, RE, and rank for each model.

All predictors were standardized using the `StandardScaler` transformation, and models were trained and tested on identical data partitions with fixed random seeds to ensure fairness and reproducibility. This uniform setup ensured that differences in RE values reflected genuine variations in predictive performance rather than discrepancies in preprocessing or initialization. The resulting RE-based ranking, from **Rank 1 (best)** to **Rank 4 (least accurate)**, guided the analysis in Section 5, highlighting that while ensemble models such as XGBoost and LightGBM showed competitive accuracy, the Spatio-Temporal framework maintained a consistent interpretative and policy-oriented advantage.

5. Numerical Analysis

This section presents the empirical evaluation of the proposed models for forecasting commodity sales. To ensure fair evaluation, all models were tested strictly *out-of-sample*: data from 2000–2010 were used for training, and forecasts were generated for 2011 to compare predicted and observed sales.

Table 3 reports the optimized coefficients $(\alpha_1^{\text{opt}}, \alpha_2^{\text{opt}})$ and MSE values for the temporal, spatial, and integrated spatio-temporal models. Across all forty stations, the hybrid framework achieved the lowest MSE, confirming its robustness and accuracy in capturing both spatial dependence and temporal persistence.

TABLE 3. Comparative analysis of MSE values for different stations

Station	α_1^{opt}	α_2^{opt}	MSE (Y_t)	MSE (Y_i)	MSE ($Y_{i,t}$)
A1	0.8737	0.1321	23330.96	161368.24	20062.50
A2	0.9904	0.0113	287.13	119068.98	236.07
A3	0.7307	0.3979	511463.55	643400.70	387721.22
A4	0.9079	0.1224	28823.29	252183.79	13216.47
A5	1.0866	-0.0803	20591.68	197642.27	19523.14
A6	0.9897	0.0124	19801.25	317477.13	19697.54
A7	0.7054	0.2883	54304.58	132671.27	38502.15
A8	1.1428	-0.1516	64187.22	157867.74	62412.03
A9	1.0728	-0.0543	10628.79	324659.81	9703.33
A10	1.1938	-0.2177	45640.52	189892.05	39928.13
B1	0.4496	0.7106	2376716.74	2513337.41	2124033.37
B2	0.8066	-0.0241	986596.47	539062.85	368047.76
B3	1.2093	-0.2022	281970.52	561366.48	273263.91
B4	0.2788	1.0331	1720039.82	2342643.60	1114297.87
B5	0.6723	0.3382	181357.36	269535.18	144942.95
B6	0.7097	0.3166	277076.75	562376.87	210404.29
B7	1.4209	-0.2753	327199.72	878308.52	167990.48
B8	0.8596	0.1205	656323.78	946243.17	639835.16
B9	0.9759	-0.0360	178087.43	741048.85	151582.90
B10	1.1532	-0.1236	241887.14	744859.71	230954.03
C1	0.8290	-0.0352	381658.94	552446.34	42852.72
C2	0.7820	0.0800	304389.61	371337.95	124981.96
C3	1.2322	-0.0179	1148986.53	1364451.79	831768.30
C4	0.9756	0.0019	249130.94	586285.23	245018.65
C5	1.0904	-0.0216	703334.41	1177195.50	665317.09
C6	0.8523	0.1907	254583.81	483387.01	233735.31
C7	1.1542	-0.2521	271231.34	878120.16	189313.25
C8	1.9759	-0.5030	1728154.04	2063773.82	810500.38
C9	0.8669	0.0478	120929.15	423550.90	78672.24
C10	1.2938	-0.4116	635771.31	993625.90	556829.20
D1	0.6014	0.2850	653155.28	701382.95	464493.52
D2	0.4772	0.4425	951730.02	941082.92	740015.40
D3	0.0231	1.0038	867883.88	677243.94	671588.69
D4	0.7323	0.0918	405192.60	986275.48	210304.58
D5	0.3727	0.5701	724178.03	756071.47	650836.51
D6	0.9687	-0.2098	811690.94	1052246.48	300030.16
D7	0.8732	-0.0434	768019.31	1096304.49	548329.37
D8	0.1531	1.1177	4272489.39	3129355.13	2439226.81
D9	-0.0019	1.0655	2139631.24	1838319.16	1803776.15
D10	0.4073	0.9243	2664637.73	3022043.75	1829948.48

The model effectively captures seasonal peaks and inter-station variation, highlighting its suitability for regional demand forecasting. Overall, integrating spatial and temporal dependencies substantially enhances predictive stability and interpretability, making the Spatio-Temporal model a reliable tool for

Table 4. Comparison of Relative Efficiency and Model Rank across Models

	Relative Efficiency						Model Rank					
Station	**ST**	**Temp**	**Spatial**	**RF**	**XGB**	**LGB**	**ST**	**Temp**	**Spatial**	**RF**	**XGB**	**LGB**
A1	100	85.99	12.43	64.96	56.65	70.13	1	2	6	4	5	3
A2	100	82.22	0.20	6.14	1.34	3.59	1	2	6	3	5	4
A3	100	75.81	60.26	74.94	71.00	69.97	1	2	6	3	4	5
A4	100	45.85	5.24	50.52	18.93	13.99	1	3	6	2	4	5
A5	100	94.81	9.88	27.32	22.37	27.80	1	2	6	4	5	3
A6	100	99.48	6.20	47.92	43.26	37.67	1	2	6	3	4	5
A7	100	70.90	29.02	68.30	98.07	49.74	1	3	6	4	2	5
A8	100	97.23	39.53	79.14	39.85	79.97	1	2	6	4	5	3
A9	100	91.29	2.99	21.85	16.38	16.35	1	2	6	3	4	5
A10	100	87.48	21.03	92.73	62.80	87.93	1	4	6	2	5	3
B1	100	89.37	84.51	111.81	96.14	126.17	3	5	6	2	4	1
B2	100	37.30	68.28	44.56	34.82	27.71	1	4	2	3	5	6
B3	100	96.91	48.68	122.47	113.50	84.90	3	4	6	1	2	5
B4	100	64.78	47.57	66.36	58.73	48.07	1	3	6	2	4	5
B5	100	79.92	53.78	89.51	58.38	72.47	1	3	6	2	5	4
B6	100	75.94	37.41	71.12	59.57	68.80	1	2	6	3	5	4
B7	100	51.34	19.13	51.37	40.06	57.70	1	4	6	3	5	2
B8	100	97.49	67.62	113.03	110.39	91.05	3	4	6	1	2	5
B9	100	85.12	20.46	71.59	51.85	54.54	1	2	6	3	5	4
B10	100	95.48	31.01	95.34	98.98	94.73	1	3	6	4	2	5
C1	100	11.23	7.76	8.59	4.76	3.37	1	2	4	3	5	6
C2	100	41.06	33.66	33.46	18.73	25.23	1	2	3	4	6	5
C3	100	72.39	60.96	68.34	63.13	57.67	1	2	5	3	4	6
C4	100	98.35	41.79	74.51	76.98	76.64	1	2	6	5	3	4
C5	100	94.59	56.52	70.07	54.28	78.49	1	2	5	4	6	3
C6	100	91.81	48.35	85.73	68.58	52.90	1	2	6	3	4	5
C7	100	69.80	21.56	36.49	36.87	29.20	1	2	6	4	3	5
C8	100	46.90	39.27	61.51	61.70	55.59	1	5	6	3	2	4
C9	100	65.06	18.57	46.64	44.77	44.04	1	2	6	3	4	5
C10	100	87.58	56.04	72.59	81.32	64.57	1	2	6	4	3	5
D1	100	71.12	66.23	89.49	97.82	91.75	1	5	6	4	2	3
D2	100	77.75	78.63	71.69	54.71	44.90	1	3	2	4	5	6
D3	100	77.38	99.16	66.06	52.78	49.25	1	3	2	4	5	6
D4	100	51.90	21.32	38.57	37.24	35.01	1	2	6	3	4	5
D5	100	89.87	86.08	119.17	83.22	70.42	2	3	4	1	5	6
D6	100	36.96	28.51	31.60	24.04	20.30	1	2	4	3	5	6
D7	100	71.40	50.02	65.95	76.24	71.31	1	3	6	5	2	4
D8	100	57.09	77.95	101.97	89.28	97.72	2	6	5	1	4	3
D9	100	84.30	98.12	91.20	64.99	58.84	1	4	2	3	5	6
D10	100	68.68	60.55	70.27	56.78	52.11	1	3	4	2	5	6

policy-oriented applications in inventory management, logistics, and climate-responsive production planning.

6. Conclusion

This study demonstrates that integrating spatial and temporal dependencies markedly enhances sales forecasting for spatial and temporal sensitive commodities. By combining the Weighted Average Spatial Model, the SARIMAX model, and the proposed Spatio-Temporal framework, the analysis captured both proximity-based spatial effects and seasonal temporal variations governing sales patterns.

The proposed Spatio-Temporal model consistently achieved the lowest Mean Squared Error (MSE) across all stations and outperformed benchmark ensemble methods—Random Forest, XGBoost, and LightGBM—in the majority of cases. This consistent superiority across diverse regions establishes its robustness and reliability, while the model's interpretable weighting scheme clarifies the relative influence of spatial and temporal factors.

From a practical and policy standpoint, the model's predictive accuracy supports improved inventory management, optimized supply chains, and climate-responsive market strategies. These results highlight the importance of spatial–temporal integration in forecasting and position the proposed model as a scalable, data-driven tool for **evidence-based policy formulation in regional demand management, logistics, and public-sector planning.**

Statements and Declarations

The authors declare that they have no conflicts of interest.

Acknowledgements

The data used in this study are entirely synthetic and were generated for methodological demonstration. All synthetic data are provided as supplementary material accompanying this submission.

References

[1] A. Appice, S. Pravilovic, D. Malerba, and A. Lanza. Enhancing regression models with spatio-temporal indicator additions. In M. Baldoni, C. Baroglio, G. Boella, and R. Micalizio, editors, *AI*IA 2013: Advances in Artificial Intelligence*, volume 8249 of *Lecture Notes in Computer Science*, pages 433–444, AI*IA 2013, Turin, 2013. Springer, Berlin, Heidelberg.

[2] F. Chen, Y. Leung, Q. Wang, and Y. Zhou. Spatial non-stationarity test of regression relationships in the multiscale geographically weighted regression model. *Spatial Statistics*, 62:Article ID 100846, 2024.

[3] T. Cheng, J. Wang, and X. Li. The support vector machine for nonlinear spatio-temporal regression. In *Proceedings of International Conference on Geocomputation*, Maynooth, Ireland, 2007. National Centre for Geocomputation, Ireland.

[4] T. Cheng, J. Wang, and X. Li. A hybrid framework for space–time modeling of environmental data. *Geographical Analysis*, 43(2):188–210, 2011.

[5] T.-M. Choi, Y. Yu, and K.-F. Au. A hybrid sarima wavelet transform method for sales forecasting. *Decision Support Systems*, 51(1):130–140, 2011.

[6] A. Ermagun and D. Levinson. Spatiotemporal traffic forecasting: review and proposed directions. *Transport Reviews*, 38(6):786–814, 2018.

[7] Z. Hajirahimi and M. Khashei. Hybrid structures in time series modeling and forecasting: A review. *Engineering Applications of Artificial Intelligence*, 86:83–106, 2019.

[8] T. Hengl, M. Nussbaum, M. N. Wright, G. B. Heuvelink, and B. Gräler. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. *PeerJ*, 6:Article ID e5518, 2018.

[9] X. Jiao, G. Li, and J. L. Chen. Forecasting international tourism demand: a local spatiotemporal model. *Annals of Tourism Research*, 83:Article ID 102937, 2020.

[10] Y. Kamarianakis. Spatial time-series modeling: A review of the proposed methodologies. *The Regional Economics Applications*, 2003.

[11] A. Kumar, A. Kumar, and B. Singh. A critical analysis of ensemble learning techniques for improved accuracy in sales forecasting. *International Journal of Applied Research*, 4(8):49–52, 2018.

[12] Z.-x. MEI. Integrated spatio-temporal forecasting method by drnn and arima combined model. *Journal of Chinese Computer Systems*, 31(4):657–661, 2010.

[13] D. Pokrajac, A. Lazarevic, and Z. Obradovic. Combining regressive and auto-regressive models for spatial-temporal prediction. In *SIAM Data Mining*, pages 439–444, 2001.

[14] S. Pravilovic, A. Appice, and D. Malerba. An intelligent technique for forecasting spatially correlated time series. In M. Baldoni, C. Baroglio, G. Boella, and R. Micalizio, editors, *AI*IA 2013: Advances in Artificial Intelligence*, volume 8249 of *Lecture Notes in Computer Science*, pages 457–468, Turin, Italy, December 2013. Springer, Berlin, Heidelberg.

[15] D. Puthran, H. Shivaprasad, K. Kumar, and M. Manjunath. Comparing sarima and holt-winters' forecasting accuracy with respect to indian motorcycle industry. *Transactions on Engineering and Sciences*, 2(5):25–28, 2014.

[16] K. Rahman, M. A. Anggorowati, and A. Andiojaya. Prediction and simulation spatio-temporal support vector regression for nonlinear data. *International Journal on Information and Communication Technology (IJoICT)*, 6(1):31–40, 2020.

[17] B. B. Singh, G. K. Shukla, and D. Kundu. Spatio-temporal models in small area estimation. *Survey Methodology*, 31(2):183–195, 2005.

[18] S. Sisman and A. C. Aydinoglu. A modelling approach with geographically weighted regression methods for determining geographic variation and influencing factors in housing price: A case in istanbul. *Land use policy*, 119:Article ID 106183, 2022.

[19] S. Tripathi, S. Yadav, and H. Kumar. A comparative analysis of traditional time series models and deep learning approaches in sales prediction: Strengths and limitations. *The Pharma Innovation Journal*, 8(2):20–24, 2019.

[20] J. D. Urrutia, J. S. Bedaa, C. B. V. Combalicer, and F. L. T. Mingo. Forecasting rice production in luzon using integrated spatio-temporal forecasting framework. In *AIP Conference Proceedings*, volume 2192, page 090014. AIP Publishing LLC, 2019.

[21] J. D. Urrutia, J. P. Santos, and C. D. Garcia. Forecasting rice production using spatio-temporal models. *Conference Proceedings*, pages 124–130, 2019.

[22] J. Wang, X. Hu, and S. Zhu. Nonlinear integration of spatial and temporal forecasting by svm. *IEEE Conference Proceedings*, pages 424–428, 2007.

[23] C. K. Wikle and M. B. Hooten. Spatial statistics: Climate and the environment. *Spatial Statistics*, 63:Article ID 100856, 2024.

[24] H. Yinnan and Q. Ruxiang. Analysis of the spatial distribution and future trends of coal mine accidents: A case study of coal mine accidents in china from 2005–2022. *Spatial Statistics*, 63:Article ID 100851, 2024.