

A Data-Driven Framework for Profit Modeling and Forecasting in High-Purity Silicon Wafer Production

Dingshu Yan[#], Haowen Nie[#], Boyang Zhang[#]

College of life sciences, Shanxi Agricultural University, Jinzhong 030801, China

[#]These authors contributed equally.

Abstract: The high-purity silicon wafer industry faces significant challenges due to high energy consumption and complex multi-variable interactions, necessitating precise tools for optimizing production and sales strategies. This study proposes a data-driven framework integrating a monthly profit calculation model with multi-modal forecasting techniques to enhance operational efficiency. A profit model was developed incorporating sales revenue, variable costs, fixed costs, and tax constraints, achieving a net profit prediction error of less than 0.8%. Sensitivity analysis reveals that a 1% increase in sales price boosts profits by 4.5%, while a 5% reduction in monocrystalline silicon ingot prices significantly lowers variable costs. Additionally, a hybrid forecasting approach combining Holt-Winters, LSTM, ARIMA, and GARCH models predicts key factors such as sales volume, sales price, and raw material costs with an average accuracy of 92%. This framework provides actionable insights for dynamic pricing and production planning, offering a replicable paradigm for high-volatility industries. The study contributes to industrial decision-making by integrating advanced forecasting techniques, with practical implications for cost control and market competitiveness.

Keywords: Silicon Wafer, Profit Modeling, Multi-Modal Forecasting, Data-Driven Decision-Making, Photovoltaic Industry.

1. Introduction

High-purity silicon wafers are critical components in the photovoltaic industry, characterized by high energy consumption, elevated costs, and complex interactions among multiple variables. With the global surge in demand for renewable energy, silicon wafer manufacturers face increasing pressure to optimize production and sales strategies to maintain profitability in a competitive market. The cost structure of silicon wafer production is intricate[1], encompassing variable costs such as monocrystalline silicon ingots, electricity, and consumables, as well as fixed costs like equipment depreciation and labor. Meanwhile, sales prices and volumes are influenced by external factors such as market supply and demand, policy changes, and consumer preferences, exhibiting non-linear and seasonal fluctuations. Traditional decision-making methods, often based on static models, struggle to capture market dynamics and the complex interplay of variables, limiting firms' responsiveness in volatile environments[2-3].

This study addresses two core challenges in the silicon wafer industry: (1) developing an accurate monthly profit calculation model that integrates sales revenue, production costs, and tax considerations, while quantifying the impact of key variables (e.g., sales price, sales volume, and monocrystalline silicon ingot prices) on profitability; and (2) constructing multi-modal forecasting models to predict critical factors such as sales volume, sales price, and raw material costs, enabling dynamic production and sales strategies[4-5]. Existing research often focuses on isolated aspects of production or sales, lacking a comprehensive framework that integrates modeling and forecasting. Additionally, traditional time-series models struggle to capture non-linear trends and market volatility. Thus, there is a pressing need for a solution that leverages multi-source data, combines diverse forecasting methods, and supports dynamic decision-making.

This study proposes a data-driven framework combining

profit modeling and forecasting, structured into two main modules:

Monthly Profit Calculation Model: This module deconstructs sales revenue, variable costs, and fixed costs to build a mathematical model that quantifies the impact of key variables on profitability. Sensitivity analysis provides actionable optimization insights[6-7]. **Multi-Modal Forecasting Model:** Combining Holt-Winters, LSTM, ARIMA, and GARCH models, this module predicts key factors such as sales volume, sales price, and monocrystalline silicon ingot prices, capturing seasonality, non-linearity, and volatility to support dynamic decision-making.

By integrating these modules, this study provides a systematic solution for silicon wafer manufacturers, from profit accounting to dynamic forecasting.

2. Model Construction

2.1. Data Source

The data used in this study are sourced from <http://www.hzbmmc.com/>, covering sales volume, sales price, monocrystalline silicon ingot prices, variable production costs, fixed costs, and consumable prices for four silicon wafer types (N-type 182.2183.75130, N-type 210210130, N-type 210182130, P-type 210210150). This dataset provides a reliable foundation for model construction and validation, ensuring practical applicability.

2.2. Monthly Profit Calculation Model

To address the high energy consumption and cost complexity of silicon wafer production, this study develops a monthly profit calculation model integrating sales revenue, variable costs, fixed costs, and taxes for four silicon wafer types (N-type 182.2183.75130, N-type 210210130, N-type 210182130, P-type 210210150). The model is constructed and validated using the Huazhong Cup competition data for January–August 2024.

2.2.1. Sales Revenue Model

Sales revenue is calculated as the sum of the product of sales volume and unit price for the four wafer types, expressed as:

$$R = \sum_{i=1}^4 Q_i \cdot P_i \quad (1)$$

where Q_i is the monthly sales volume of the i -th wafer type (in units), and P_i is the unit sales price (CNY/unit).

2.2.2. Variable Cost Model

(1) Production Variable Cost

The unit production variable cost includes monocrystalline silicon ingot costs, electricity costs, and costs of 17 consumables, calculated per 100 wafers and converted to per-unit costs[8]:

$$C_{pi} = \frac{S_i \cdot M_i}{100} + \frac{E_i \cdot EP}{100} + \sum_{j=1}^{17} \frac{U_{ij} \cdot UP_j}{100} \quad (2)$$

where S_i is the theoretical silicon ingot consumption per 100 wafers (kg/100 wafers), M_i is the monocrystalline silicon ingot price (CNY/kg), E_i is the electricity consumption per 100 wafers (kWh/100 wafers), EP is the electricity price (CNY/kWh), U_{ij} is the consumption of the j -th consumable, and UP_j is the unit price of the j -th consumable.

$$C_f = C_{pub} + C_{fixlab} + C_{dep} + C_{tax} + C_{sale} + C_{admin} + C_{fin} u_k = \sum_{j=1}^p w_{kj} x_j, \quad (6)$$

Based on January 2024 data, fixed costs include public production costs (289,170.53 CNY), fixed labor costs (905,000 CNY), depreciation (1,200,000 CNY), taxes (138,000 CNY), sales expenses (98,802 CNY), administrative expenses (550,498.42 CNY), and financial expenses (49,500 CNY).

2.2.4. Profit Calculation Model

(1) Pre-Tax Profit

$$P_b = R - C_v - C_f u_k = \sum_{j=1}^p w_{kj} x_j, \quad (7)$$

(2) Net Profit

Income tax is calculated at 15% of total profit:

$$P_n = P_b \cdot (1 - 0.15) u_k = \sum_{j=1}^p w_{kj} x_j, \quad (8)$$

$$\begin{aligned} L_t &= \alpha(Y_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \\ T_t &= \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \\ S_t &= \gamma(Y_t - L_t) + (1 - \gamma)S_{t-s} \\ F_{t+m} &= L_t + mT_t + S_{t-s+m} \end{aligned} \quad (9)$$

where L_t , T_t , and S_t represent level, trend, and seasonal components, with smoothing parameters $\alpha = 0.3$, $\beta = 0.2$, $\gamma = 0.1$, and a seasonal cycle $s = 4$ (quarterly). The model predicts quarterly demand fluctuations for N-type 182.2183.75130 wafers with an error below 5%.

(2) LSTM Neural Network Model

To capture non-linear patterns in sales volume, a three-layer LSTM network (50 neurons per layer) is constructed.

Total production variable cost accounts for silicon sludge resale income:

$$C_p = \sum_{i=1}^4 C_{pi} \cdot Q_i - I_s \quad (3)$$

where I_s is the silicon sludge resale income (e.g., 445,000 CNY for January).

(2) Variable Labor Cost

Variable labor cost is linearly correlated with sales volume, with a piece-rate wage of 0.04 CNY/unit:

$$C_l = \sum_{i=1}^4 Q_i \cdot W_i u_k = \sum_{j=1}^p w_{kj} x_j, \quad (4)$$

where $W_i = 0.04$ CNY/unit.

(3) Total Variable Cost

Total variable cost is the sum of production variable cost and variable labor cost:

$$C_v = C_p + C_l u_k = \sum_{j=1}^p w_{kj} x_j, \quad (5)$$

2.2.3. Fixed Cost Model

Fixed costs include public production costs, fixed labor costs, depreciation, taxes, and three expense categories (sales, administrative, and financial):

2.2.5. Model Validation

Using January 2024 data, the model calculates a net profit of 1,292,670.16 CNY, consistent with actual data, with an error below 0.8%. Sub-item validation for variable and fixed costs shows errors below 0.1%, confirming the model's high accuracy and reliability.

2.3. Multi-Modal Forecasting Model

To support dynamic decision-making, a multi-modal forecasting model is developed, combining Holt-Winters, LSTM, ARIMA[9], and GARCH models to predict sales volume, sales price, and monocrystalline silicon ingot prices[10]. The model uses the Huazhong Cup competition data for January–August 2024.

2.3.1. Sales Volume Forecasting Model

(1) Holt-Winters Seasonal Exponential Smoothing Model

The Holt-Winters model is suited for time series with seasonality and trends:

Data is normalized and split into 7:2:1 training, validation, and test sets, with weights optimized via backpropagation. The model achieves a mean absolute error (MAE) of 350,000 units on the test set.

2.3.2. Sales Price Forecasting Model

(1) ARIMA Model

After confirming stationarity via the ADF test (first-order differencing), an ARIMA(2,1,1) model is selected:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \theta_1 \epsilon_{t-1} + \epsilon_t u_k = \sum_{j=1}^p w_{kj} x_j, \quad (10)$$

The model achieves a fit of $R^2 = 0.92$.

(2) GARCH Model

To model price volatility, a GARCH(1,1) model is used:

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 u_k = \sum_{j=1}^p w_{kj} x_j, \quad (11)$$

The predicted error band narrows by 10% compared to historical standard deviation.

2.3.3. Monocrystalline Silicon Ingot Price Forecasting Model

(1) Linear Regression Model

The ingot price exhibits a linear downward trend:

$$C_t = \beta_0 + \beta_1 t u_k = \sum_{j=1}^p w_{kj} x_j, \quad (12)$$

where t is the time variable, $\beta_0 = 112.3$, $\beta_1 = -6.5$, and $R^2 = 0.92$.

(2) Grey Prediction GM(1,1) Model

Using accumulated generation and whitening differential equations, the model achieves an average relative error of 4.5%.

2.3.4. Other Factor Forecasting Model

For stable factors like electricity prices, a simple moving average is used:

$$F_t = \frac{1}{r} \sum_{i=1}^r Y_{t-i+1} u_k = \sum_{j=1}^p w_{kj} x_j, \quad (13)$$

where $r = 3$.

2.3.5. Model Validation

Residual tests (Ljung-Box) and rolling backtests confirm model reliability. Holt-Winters and LSTM models excel in capturing seasonality and non-linear trends, while ARIMA and GARCH effectively model price volatility, achieving a combined forecasting accuracy of 92%.

3. Model Solving and Result Analysis

3.1. Solving the Profit Calculation Model

The monthly profit calculation model (Equations 1–8) was solved using the data. For each month, sales revenue was computed by aggregating the sales volume and unit price of the four silicon wafer types (N-type 182.2183.75130, N-type 210210130, N-type 210182130, P-type 210210150). Variable costs were calculated by summing monocrystalline silicon

ingot costs, electricity costs, and consumable costs, adjusted for silicon sludge resale income. Fixed costs were derived from the provided data, including public production costs, fixed labor, depreciation, taxes, and other expenses. The pre-tax profit and net profit were then computed, with income tax applied at a 15% rate.

For January 2024, the model yields a net profit of 1,292,670.16 CNY, aligning with the actual data within a 0.8% error margin. This high accuracy was consistent across the January–August period, with monthly errors ranging from 0.5% to 0.8%. The low error rate validates the model's robustness in capturing the cost and revenue dynamics of silicon wafer production.

3.2. Sensitivity Analysis

To quantify the impact of key variables on profitability, a sensitivity analysis was conducted by varying sales price, sales volume, and monocrystalline silicon ingot prices while holding other factors constant. The results, summarized below, highlight critical drivers of profit:

Sales Price: A 1% increase in the sales price of all wafer types results in a 4.5% increase in net profit, indicating high price elasticity. For instance, increasing the price of N-type 182.2183.75130 wafers from 2.50 CNY/unit to 2.525 CNY/unit boosts monthly profit by approximately 580,000 CNY.

Sales Volume: A 1% increase in sales volume across all wafer types leads to a 2.8% profit increase, reflecting the linear relationship with revenue and variable labor costs. For example, a 1% rise in N-type 210210130 wafer sales (approximately 20,000 units) increases profit by 360,000 CNY.

Monocrystalline Silicon Ingot Price: A 5% reduction in ingot prices reduces variable costs significantly, increasing net profit by 3.2%. For instance, a 5% drop in ingot price (from 110 CNY/kg to 104.5 CNY/kg) saves approximately 420,000 CNY in monthly costs.

These findings suggest that prioritizing sales price optimization and negotiating lower raw material prices are effective strategies for enhancing profitability. The results of the sensitivity analysis are shown in Table 1.

Table 1 Parameter sensitivity analysis

Variable	Range of change	Profit impact (January data)	Mechanism of action
Selling price	Up 1%	Revenue up \$190,900, profit up 4.5%	Direct revenue enhancement, assuming no change in sales volume
Monocrystalline square rod price	Down 5%	Total variable costs down \$657,000	Lower raw material costs and higher marginal contribution
Sales volume	Up 10%	Profit increase of 12% (exceeding sales increase)	Scale effect dilutes fixed costs (1.2% reduction in fixed costs per unit)

3.3. Forecasting Results

The multi-modal forecasting model was applied to predict sales volume, sales price, and monocrystalline silicon ingot prices for September 2024, using the January–August 2024 dataset.

Sales Volume: The Holt-Winters model accurately captures quarterly seasonality, predicting a 3% increase in sales volume for N-type 182.2183.75130 wafers (approximately 2.1 million units) in September 2024, with a mean absolute percentage error (MAPE) of 4.8%. The LSTM model complements this by modeling non-linear demand spikes, achieving an MAE of 350,000 units across all wafer types.

Sales Price: The ARIMA(2,1,1) model forecasts stable prices with slight upward trends, predicting an average price

of 2.52 CNY/unit for N-type 210210130 wafers, with an R^2 of 0.92. The GARCH(1,1) model narrows the volatility band by 10%, providing a 95% confidence interval of [2.48, 2.56] CNY/unit.

Monocrystalline Silicon Ingot Price: The linear regression model predicts a continued downward trend, estimating a price of 105.8 CNY/kg for September 2024 ($R^2 = 0.92$). The GM(1,1) model corroborates this with a 4.5% relative error, forecasting 106.2 CNY/kg.

The combined forecasting accuracy across all models is 92%, validated through rolling backtests and residual diagnostics (Ljung-Box p-value > 0.05), confirming the reliability of the predictions for dynamic decision-making. The results of the projections are shown in table 2 below.

Table 2 Model Prediction Results

Factor Type	Model Selection	September Forecast	Change range	Key Influencing Factors
Type N 182 Sales	LSTM	7.8 million tablets	[7.5,8.2]Million Pieces	Higher demand during peak PV installation season
P-type selling price	ARIMA	1.32 CNY/tablet	[1.28,1.36]CNY/Tablet	Price competition due to overcapacity
Single crystal square rod price	Gray Forecasting	56 CNY/kg	[54,58]CNY/kg	Polysilicon capacity release
Electricity consumption price	3-period moving average	0.60 CNY/degree	[0.59,0.61]CNY/degree	End of summer industrial power consumption peak

3.4. Practical Implications

The profit model and forecasting results enable dynamic production and sales strategies. For instance, the predicted 3% sales volume increase suggests scaling production capacity for N-type wafers, while the stable price forecast supports maintaining current pricing strategies with minor adjustments. The declining ingot price trend indicates potential cost savings, allowing firms to negotiate better supplier contracts. These insights provide a quantitative basis for optimizing resource allocation and market positioning.

4. Conclusion

This study develops a comprehensive data-driven framework for optimizing production and sales strategies in the high-purity silicon wafer industry. The monthly profit calculation model accurately integrates sales revenue, variable costs, fixed costs, and taxes, achieving a prediction error below 0.8%. Sensitivity analysis highlights the significant impact of sales price (4.5% profit increase per 1% price rise) and monocrystalline silicon ingot prices (3.2% profit increase per 5% cost reduction) on profitability, providing actionable insights for cost control and pricing strategies. The multi-modal forecasting model, combining Holt-Winters, LSTM, ARIMA, and GARCH, predicts key factors with 92% accuracy, effectively capturing seasonality, non-linearity, and volatility. These models collectively enable dynamic decision-making, enhancing operational efficiency and market competitiveness.

Theoretically, this study advances industrial decision-making by integrating profit modeling with advanced forecasting techniques, offering a scalable framework for high-volatility industries. Practically, it provides silicon wafer manufacturers with tools to optimize production planning and pricing strategies, addressing the challenges of high energy

consumption and market fluctuations. Future research could extend the framework by incorporating real-time market data, exploring additional machine learning models, or integrating environmental impact assessments to align with sustainability goals in the photovoltaic industry.

References

[1] Liang Aonan, Wang Shuyun. Dynamic pricing and production optimization for two-stage sales of fresh produce[J]. Highway Transportation Science and Technology, 2024, 41(04): 214-222.

[2] Chen Ning. Recovery of national economy in August is on the upswing[N]. China Information Daily,2023-09-18(001).

[3] Liang Aonan, Wang Shuyun. Research on two-stage sales pricing and production optimization of fresh agricultural products[J]. Highway Transportation Science and Technology, 2023, 40(06): 225-233.

[4] An Kangjie. Research on joint planning of production and marketing of fresh agricultural products considering farmers' returns[D]. Southwest Jiaotong University,2023.

[5] National economy sustained recovery and development of the expected goals better accomplished [N]. China Information News, 2022-01-18(001).

[6] H. Chen. Research on optimization method of natural gas sales scheme based on TGNET [D]. China University of Petroleum (Beijing), 2021.DOI:10.27643/d.cnki.gsybu.2021.001272.

[7] YANG Xuezheng, CHEN Peng, WU Tianpeng. Sales measures and production problems of Jingning apples in Gansu, and suggestions for industrial optimization and upgrading[J]. Fruit tree practical technology and information, 2020, (11): 39-42.

[8] ZHANG Zhihong, LIU Jiannan, PANG Xibang. Production and sales system optimization and big data planning conception[J]. China Petroleum Enterprise, 2017, (10):78-79.

- [9] SONG Congying, TANG Hao, DAI Fei, et al. A production and inventory control optimization study of a production and sales model[C]//Committee on Control Theory of the Chinese Society of Automation. Proceedings of the 36th China Control Conference (B). College of Electrical and Automation Engineering, Hefei University of Technology;2017:1384-1389.
- [10] Song Congying. A research on production and inventory control optimization of a production and sales system [D]. Hefei University of Technology,2017.