Climate Effects and Driving Mechanisms of Primary Energy Consumption in Sichuan Province: An Integrated Framework Combining Time Series Analysis with Causal Inference

Liping Wang [1, *, #], Hao Wang [2, #]

[1] School of Mathematical Sciences, Chengdu University of Technology, Yibin, China, 644000

[2] College of Computer Science and Cyber Security (Pilot Software College), Chengdu University of Technology, Yibin, China, 644000

* Corresponding Author Email: 202219010214@stu.cdut.edu.cn

[#]These authors are contributed equally.

Abstract. This study quantifies the impact of primary energy consumption on climate change in Sichuan Province by analyzing 2012–2022 datasets through integrated time series analysis and causal inference methods. Key methodologies include anomaly detection via boxplot analysis, data correction using Lagrange interpolation and moving averages, and modeling via polynomial fitting, multiple linear regression, Pearson correlation, and Granger causality tests. Results demonstrate: (1) Significant positive correlations between fossil fuel consumption (coal, petroleum, natural gas) and CO_2 emissions (correlation coefficient: 0.9654), with coal exhibiting the strongest driving effect (regression coefficient: 0.9055); (2) CO_2 emissions act as a Granger cause of temperature variation (lag order: 3, p-value <0.05), while no causal link with precipitation is observed; (3) Weak nonlinear climate interactions, evidenced by a marginal positive correlation between CO_2 and temperature (Pearson: 0.2106) and a weak negative correlation with precipitation (-0.3616). As a critical energy hub and ecological barrier in western China, Sichuan's fossil fuel dependency underscores the urgency of transitioning to renewable energy and implementing climate-resilient water management. The proposed "time series-causal inference" framework offers scalable solutions for regional energy-climate governance, aligning with global sustainability goals and providing actionable insights for policymakers to prioritize coal reduction and enhance adaptive strategies in ecologically vulnerable regions.

Keywords: Primary Energy Consumption, Climate Change, Time Series Analysis, Pearson Correlation, Granger Causality Test.

1. Introduction

1.1. Research Background

Global climate change has emerged as a critical global challenge. Energy consumption, as a primary source of carbon emissions, is closely linked to climatic variations. Sichuan Province, a major energy-consuming region in China, has experienced continuous growth in primary energy consumption amid rapid economic development, potentially exerting significant impacts on regional and global climate systems. Investigating the relationship between primary energy consumption and climate change in Sichuan Province is essential for formulating rational energy policies and climate change mitigation strategies.

1.2. Literature Review

Global climate change has emerged as a critical challenge for humanity in the 21st century. Energy consumption, as a primary source of carbon emissions, and its interaction with climate systems have garnered significant academic attention. Scholars worldwide have extensively explored the linkages between energy consumption and climate change, providing critical insights for regional climate governance and energy policy formulation.

At the international level, numerous studies have confirmed the strong correlation between fossil energy consumption and greenhouse gas emissions. For instance, energy system analyses based on the LEAP model (Dai Zhipeng, 2021) [1] demonstrate that overreliance on coal and petroleum remains a core driver of carbon emission growth, while low-carbon energy transitions can markedly mitigate global warming trends. Furthermore, time series analysis methods, such as the Granger causality test, have been widely applied to investigate energy-climate dynamics. These methods reveal causal relationships between carbon emissions and temperature variations by analyzing lag effects (Tang Xiaoyan, 2024) [2]. However, existing research predominantly focuses on developed countries or global scales, with limited exploration of regional disparities in developing nations.

In the Chinese context, studies on energy-climate linkages have progressively shifted from macro-level analyses to region-specific investigations. Wang Yanan et al. (2022) [3] identified the positive role of clean energy adoption in climate improvement through CO_2 mediation but overlooked the differential impacts of distinct energy types. For southwestern China, Xu Haining et al. (2019) [4] analyzed spatiotemporal patterns of precipitation changes, highlighting climate-induced threats to regional hydrological cycles without integrating energy consumption drivers. Yang Mingxin et al. (2022) [5] simulated and projected summer climate changes in southwestern China using the CMIP6 model, revealing spatial heterogeneity in regional climate responses, which provides methodological insights for analyzing Sichuan's climate dynamics. Sichuan Province, a major energy consumer in western China, exhibits unique climatic influences due to its coal dominated energy structure and abundant hydropower resources. Xiang Bo et al. (2016) [6] quantified the potential of renewable energy utilization in Sichuan through carbon pinch analysis, indicating that synergistic development of hydropower and photovoltaic resources is a critical pathway for low-carbon energy transition in the region Li Xia (2020) [7] explored Sichuan's energy-saving potential from an energy-economy coupling perspective but failed to systematically quantify direct causal effects of energy consumption on climate indicators such as temperature and precipitation. The China Climate Change Blue Book (2023) [8] systematically summarizes observational facts of national climate change, providing authoritative data support for analyzing long-term trends in temperature and precipitation in Sichuan Province.

Current literature exhibits three key limitations: (1) Regional studies predominantly emphasize the economic impacts of energy consumption, with inadequate causal mechanism analysis of climatic effects;(2) Methodologies overly rely on correlation analysis, lacking integrated applications of time series and causal inference approaches;(3) Empirical research on the relationship between Sichuan's energy transition and climate responses remains absent. Long Tenggang (2024) [9] proposed an environmental impact assessment framework for climate change, offering theoretical references for future multi-factor coupling analysis in this field.

To address these gaps, this study employs 2012–2022 data from Sichuan Province, integrating time series analysis and causal inference methods to systematically investigate dynamic relationships among coal, petroleum, natural gas consumption, and climate variables (temperature, precipitation, CO_2 emissions). By combining polynomial fitting, multiple linear regression, Pearson correlation, and Granger causality tests, we aim to uncover nonlinear impacts and causal pathways of energy consumption on climate change. The findings seek to provide scientific support for optimizing energy policies and enhancing climate adaptation management in Sichuan, thereby filling critical gaps in regional research.

2. Materials and Methods

2.1. Data Acquisition and Preprocessing

This study collected data on petroleum, natural gas, coal consumption, temperature, precipitation, and CO_2 emissions in Sichuan Province from the National Bureau of Statistics (stats.gov.cn), China Energy Statistical Yearbook (nbsti.net), and Statistical Yearbook of Sichuan Province (sc.gov.cn). Missing values and outliers were addressed using the box-plot method, Lagrange polynomial interpolation, and moving average techniques, followed by data standardization to ensure consistency and comparability.

2.2. Methodology

The research methodology and statistical analyses adopted in this study are grounded in the probability theory and mathematical statistics framework established by Mao et al. (2011) [10], thereby ensuring the scientific validity of data processing and the robustness of the models.

2.2.1. Polynomial Fitting Method

The polynomial fitting method employs the least squares principle to establish the relationship f (x, A) between two variables from a given nonlinear dataset {(x_i, y_i), i=0, 1, 2,...,n}, where A=(a_0, a_1,...,a_n) represents undetermined coefficients. The method minimizes the sum of squared residuals between observed values and model predictions, yielding the polynomial curve:

$$y = a_0 + a_1 x^n + a_2 x^{n-1} + \cdots + a_n x \tag{1}$$

This curve effectively captures nonlinear relationships within the measured data.

2.2.2. Multiple Linear Regression

The multiple linear regression model employs two or more explanatory variables to quantify their collective influence on a dependent variable, revealing both the magnitude and direction of these effects. The model is expressed as:

$$y = y_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \varepsilon \tag{2}$$

2.2.3. Pearson Correlation

Pearson correlation measures linear relationships between variables. The correlation coefficient r ranges from [−1,1]:

r<0: indicates a negative linear correlation;
r>0: indicates a positive linear correlation;
r=0: implies no significant linear relationship.

2.2.4. Granger Causality Test

The Granger causality test assumes that all predictive information for variables y and x is embedded within their time series. The test involves estimating the following regressions:

$$y_i = \sum_{i-1} \alpha_i x_{t-i} + \sum_{j-1} \beta_j y_{t-j} + u_{1t} \tag{3}$$

$$x_i = \sum_{i-1} \lambda_i x_{t-1} + \sum_{j-i} \delta_j y_{t-j} + u_{2t} \tag{4}$$

Where u_{1t} and u_{2t} are uncorrelated white noise terms. Equation (3) posits that current y depends on past values of y and x, while Equation (4) assumes analogous behavior for x. The null hypotheses are:

For Equation (3): H_0: $\alpha_1 = \alpha_2 = \cdots = \alpha_q$=0; For Equation (4): H_0: $\delta_1 = \delta_2 = \cdots = \delta$s=0.

Four scenarios are analyzed:

Unidirectional causality (x → y): x Granger-causes y.
Unidirectional causality (y → x): y Granger-causes x.
Bidirectional causality: Mutual causation between x and y.
Independence: No causal relationship between x and y.

The testing procedure is illustrated in Figure 1.

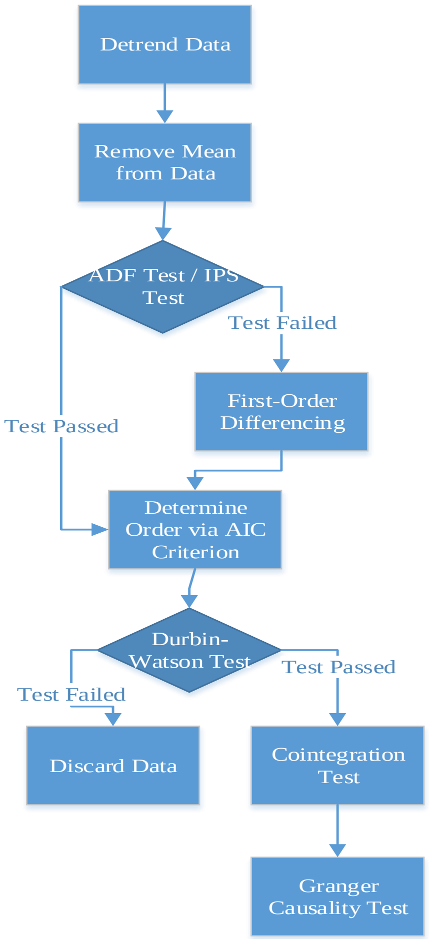


Figure 1. Granger Causality Test Flowchart.

3. Results and Analysis

3.1. Data Processing and Visualization

3.1.1. Data Integration

To facilitate statistical modeling, data on coal, petroleum, natural gas consumption, temperature, precipitation, and CO_2 emissions from diverse sources were integrated. This integration ensured data integrity and consistency for subsequent analyses. The consolidated datasets for energy consumption and climate indicators are presented in Table.1. and Table.2.

Table 1. Primary Energy Consumption by Type (2012–2022)

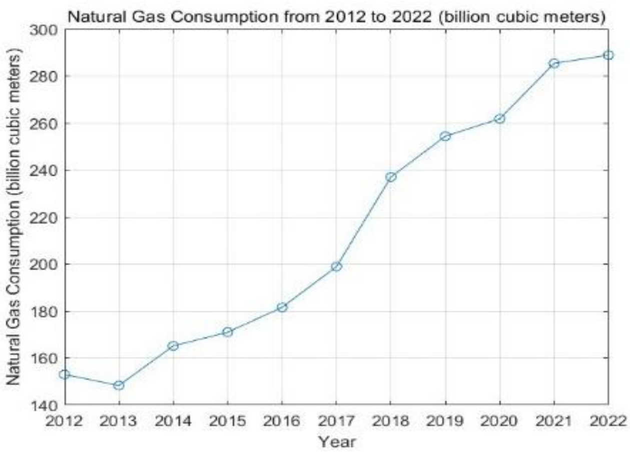
Indicator	Coal Consumption (million tons)	Natural Gas Consumption (100 million cubic meters)	Petroleum Consumption (million tons)
2012	11872	153	2194.55
2013	11678.55	148.3	2428.6
2014	11045.39	165.17	2724.1
2015	9288.9	170.98	2408.7
2016	8869.49	181.57	2474.6
2017	7855.88	198.91	2579.7
2018	7495.78	237	2549
2019	7713.47	254.38	2689.1
2020	7501.6	261.8	2584.3
2021	7796.1	285.4	2660.1
2022	7806.6	288.9	2713.4

Table 2. Climate Indicator Metrics (2012–2022)

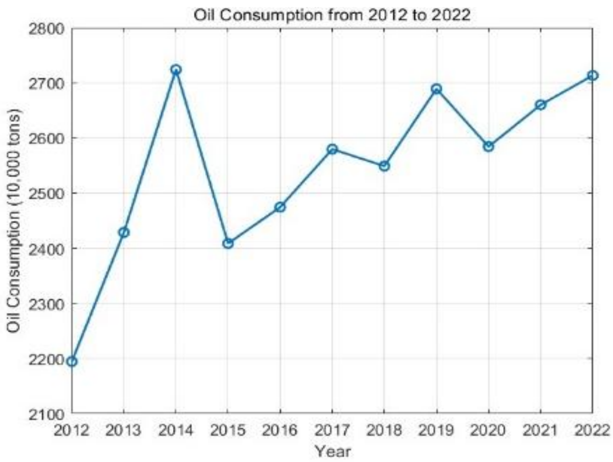
Indicator	Temperature (°C)	Precipitation (mm)	CO_2 Emissions (million tons)
2012	16.3714	22500	39456.48161
2013	17.3952	23152.9	40709.02213
2014	17.39	22598.3	40870.69912
2015	17.381	21092.2	38298.57672
2016	17.1905	21298.6	37747.27273
2017	17.0476	21388.3	36227.96252
2018	16.981	25039.7	34319.86571
2019	16.9464	22919.7	35902.95336
2020	17.0274	24843.1	35509.32282
2021	17.1107	24838.2	35901.28646
2022	17.5936	19291.2	37341.7246

3.1.2. Data Visualization and Anomaly Detection

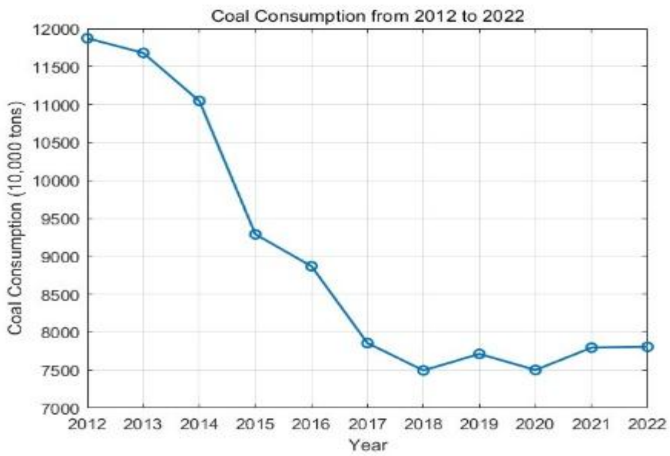
Integrated data tables were read using MATLAB built-in functions (e.g., readtable, readmatrix). Temporal trends and distribution patterns were visualized through graphical tools (plot, scatter, bar). For instance, line charts depicting coal, petroleum, and natural gas consumption over time were generated to analyze their dynamic trends, as illustrated in Figure 2.



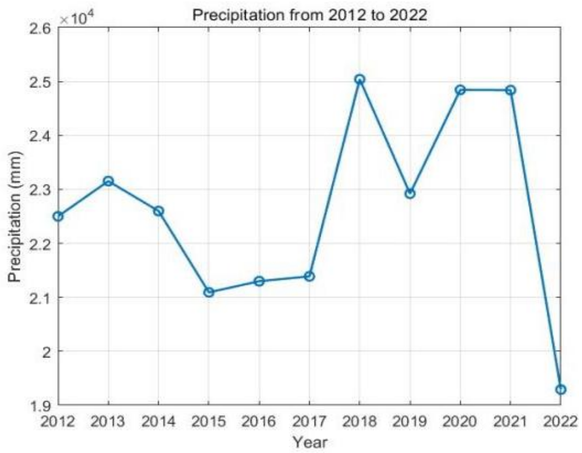
(a) Natural Gas Emissions



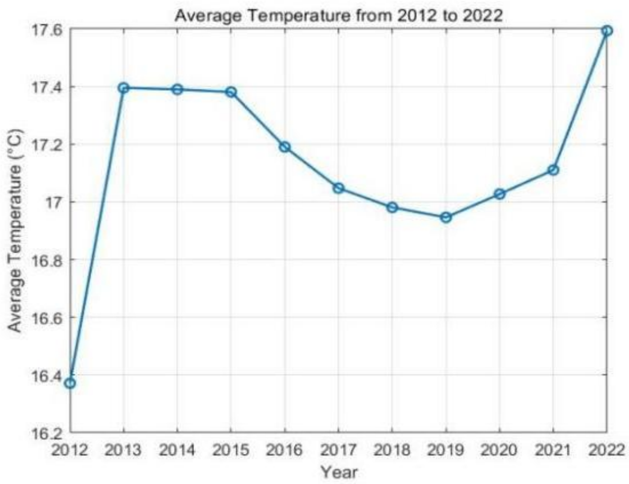
(b) Oil Consumption



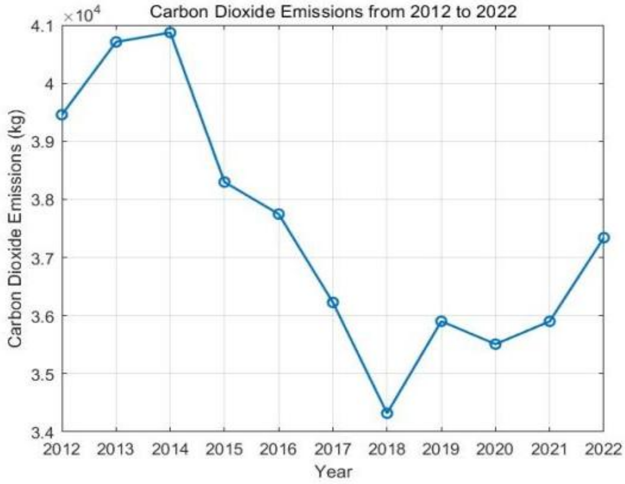
(c) Coal consumption



(d) Precipitation



(e) Average Temperature



(f) Carbon Dioxide Emissions

Figure 2. Dynamic Trends of Energy Consumption and Climate Factors (2012–2022)

Additionally, due to potential registration errors in annual data reporting and insufficient disclosure of external information, detection and treatment of outliers in the dataset were necessary. This study employed the box-plot method to identify potential outliers, ultimately revealing significant anomalies in petroleum consumption and temperature data.

The results are shown in Table 3 and Table 4.

Table 3. Energy Consumption: Box-Plot Analysis Results

Indicator	Coal Consumption (million tons)	Natural Gas Consumption (100 million cubic meters)	Petroleum Consumption (million tons)
Minimum	7495.78	148.3	1960.5
Mean	8993.069091	213.2190909	2524.736364
Standard Deviation	1733.557936	53.57017705	217.1218703
First Quartile (Q1)	7713.47	165.17	2428.6
Sample Median	7855.88	198.91	2579.7
Third Quartile (Q3)	11045.39	261.8	2689.1
Maximum	11872	288.9	2724.1

Table 4. Climatic Factors: Box-Plot Analysis Results

Indicator	Annual Average Temperature (°C)	CO_2 Emissions (million tons)	Precipitation (mm)	Temperature(°C)
Minimum	16.37142857	34319.86571	20140.25	16.3714
Mean	17.06727994	37494.34432	22632.93	17.1028
Standard Deviation	0.338635103	2284.886981	1810.797504	0.3465
First Quartile (Q1)	16.94642857	40870.69912	24838.2	17.0274
Sample Median	17.04761905	36987.61763	22598.3	17.0476
Third Quartile (Q3)	17.38095238	35902.95336	21388.3	17.381
Maximum	17.59365079	40870.69912	25996.25	17.5936

3.1.3. Outlier Treatment

This study employed two methodologies for outlier treatment: the Lagrange interpolation polynomial method and the moving average method. The Lagrange interpolation was applied to temperature anomalies, while the moving average addressed petroleum consumption outliers.

(1) Lagrange Interpolation for Outlier Correction

① Construct Lagrange Interpolation Polynomial: A polynomial was derived from known data points (excluding outliers) using the formula:

$$L(x) = \sum\left(y_i * li(x)\right) \tag{5}$$

Where y_i represents the function values at known data points, and li(x)denotes the Lagrange basis polynomials.

② Estimate Outlier Values: The anomalous data point (e.g., the 2014 value) was substituted into the polynomial to obtain an interpolated estimate, serving as a replacement or diagnostic reference.

③ Replace Outliers: The estimated value was directly substituted for the outlier to ensure data continuity.

(2) Moving Average for Outlier Correction

① Determine Window Size: A window size was selected based on data characteristics and smoothing requirements. Larger windows enhance noise reduction but may obscure short-term trends.

② Compute Moving Average: For each data point, the average of its neighboring values within the window was calculated using:

$$MA_t = \frac{1}{k}\sum_{i=-\frac{k-1}{2}}^{\frac{k-1}{2}} x_{t=i} \tag{6}$$

Where k is the window size (odd integer).

(3) Replace Outliers:

The moving average value replaced the original outlier, preserving temporal coherence.

3.1.4. Data Normalization

Following outlier treatment, Min-Max normalization was applied to the post-processed datasets to eliminate dimensional heterogeneity and facilitate model computation. The procedure involved:

Identify Minimum and Maximum Values: Determine the global minimum (x_{min}) and maximum (x_{max}) of the dataset.

Normalize Data Points: Scale each data point xi to the [0,1] interval using:

$$x_{normalized} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{7}$$

Map to Unified Scale: The normalized data, now dimensionless and uniformly scaled, enhance comparability and algorithmic efficiency in subsequent modeling.

The normalized datasets are summarized in Table 5.

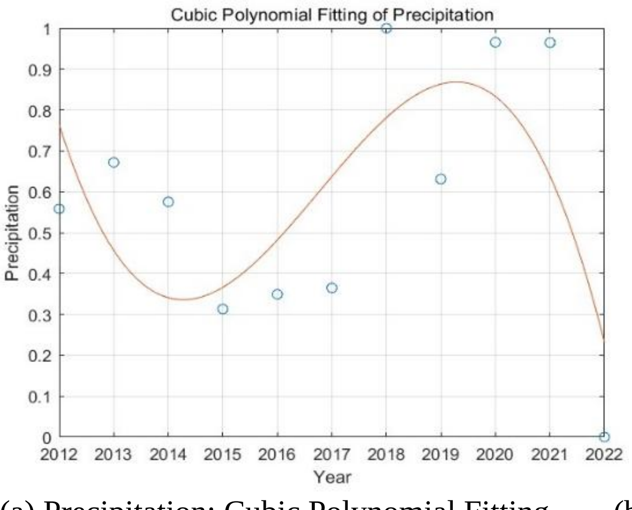
Table 5. Post-Normalization Dataset

Indicator	Coal Consumption	Natural Gas Consumption	Petroleum Consumption	Precipitation	CO_2 Emissions	Temperature
2012	1	0.0334	0	0.5582	0.7841	0
2013	0.9558	0	0.613	0.6718	0.9753	0.8377
2014	0.8111	0.12	1	0.5753	1	0.8334
2015	0.4097	0.1613	0.587	0.3133	0.6074	0.8261
2016	0.3139	0.2366	0.6733	0.3492	0.5232	0.6702
2017	0.0823	0.36	0.8109	0.3648	0.2913	0.5533
2018	0	0.6309	0.7707	1	0	0.4988
2019	0.0497	0.7545	0.9542	0.6312	0.2417	0.4705
2020	0.0013	0.8073	0.8169	0.9658	0.1816	0.5367
2021	0.0686	0.9751	0.9162	0.9649	0.2414	0.6049
2022	0.071	1	0.986	0	0.4613	1

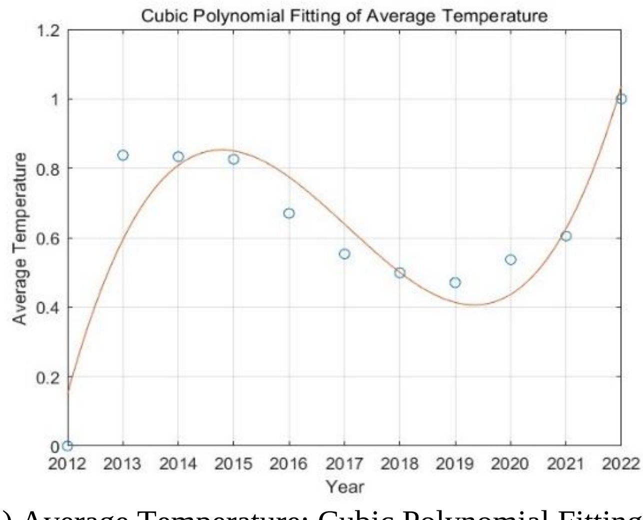
3.2. Model Construction

3.2.1. Polynomial Fitting

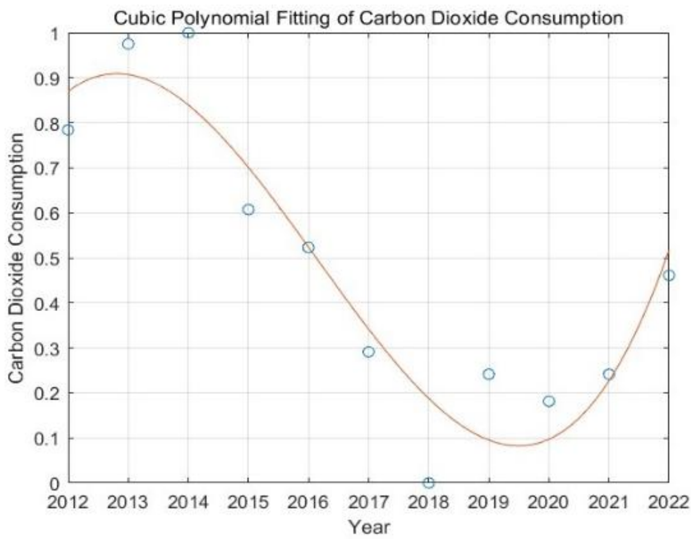
For the normalized datasets, polynomial fitting was employed to establish relationship models between energy consumption and climate indicators. By fitting the data, nonlinear relationships between variables were explored. The fitting results were visualized using MATLAB's plotting functions, as demonstrated in Figure 3.



(a) Precipitation: Cubic Polynomial Fitting



(b) Average Temperature: Cubic Polynomial Fitting



(c) CO_2 Emissions: Cubic Polynomial Fitting

Figure 3. Cubic Polynomial Fitting Curves for Climate Factors

3.2.2. Linear Regression

To investigate the relationship between energy consumption and climate indicators, this study employed multiple linear regression analysis, with coal, natural gas, and petroleum consumption as explanatory variables, and climate indicators (e.g., CO_2 emissions) as the dependent variable. Given that fossil fuels are primary contributors to CO_2 emissions, the regression model quantifies their collective impact on climatic variables. The sample period spanned 2012–2022, with climate indicators designated as dependent variables and primary energy consumption as predictors.

The regression model is formulated as:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon \tag{8}$$

y: CO_2 emissions (dependent variable); x1: Coal consumption; x2: Natural gas consumption; x3: Petroleum consumption; ε: Error term; βi (i =0,1,2,3): Coefficients estimated via the least squares method. Derive the equation:

$$y = -0.1051 + 0.9055x_1 - 0.0558x_2 + 0.4108x_3 \tag{9}$$

Regression Results (from MATLAB outputs):

① Correlation coefficient: 0.9654 (indicating strong linear association);

② F-statistic: 132.4563 >F0.95(4, 13) F0.95(4,13);

③ p-value: 0.001 (< 0.05), confirming model significance.

3.2.3. Pearson Correlation

Building on the regression analysis, Pearson correlation coefficients were computed to assess pairwise relationships between CO_2 emissions and climate indicators:

① CO_2 vs. Precipitation: r=−0.3616, suggesting a weak negative correlation;

② CO_2 vs. Average Temperature: r=0.2106, indicating a negligible positive correlation.

3.2.4. Granger Causality Test

The Granger causality test was conducted to determine whether CO_2 emissions exhibit causal effects on temperature and precipitation. In this analysis, CO_2 emissions data with lag orders of 1, 2, and 3 were utilized to explore their temporal influence on climatic variables.

The results, as shown inTable.6, indicate the statistical significance of lagged CO_2 emissions in predicting temperature variations.

Table 6. Granger Causality Test Results Between CO_2 Emissions and Temperature Under Varying Lag Orders

Null Hypothesis	Lags	F-Statistic	p-Value
y does not Granger-cause x	1	1.5421	0.2495
x does not Granger-cause y		16.1357	0.0039
y does not Granger-cause x	2	7.3956	0.0240
x does not Granger-cause y		38.9509	0.00036571
y does not Granger-cause x	3	15.0389	0.0121
x does not Granger-cause y		22.0017	0.0060

Note: x: CO_2 emissions; y: Temperature

The results confirm that CO_2 emissions are the Granger cause of annual temperature variations, implying that CO_2 emissions not only correlate with but also causally influence temperature trends.

Subsequently, the Granger causality test was extended to examine the impact of CO_2 emissions on both temperature and precipitation. The outcomes for precipitation are summarized in Table.7.

Table 7. Granger Causality Test Results Between CO_2 Emissions and Precipitation Under Varying Lag Orders

Null Hypothesis	Lags	F-Statistic	p-Value
y does not Granger-cause x	1	-2.2172	1
x does not Granger-cause y		-4.1486	1
y does not Granger-cause x	2	6.1366	0.0354
x does not Granger-cause y		1.3067	0.3380
y does not Granger-cause x	3	7.3273	0.0421
y does not Granger-cause x		3.3088	0.1390

Note: x: CO_2 emissions; y: Precipitation.

Although CO_2 emissions were confirmed as the Granger cause of annual temperature variations, no causal relationship was identified between CO_2 emissions and precipitation. The Granger causality test results indicate that CO_2 emissions are not a Granger cause of precipitation changes, suggesting that while CO_2 emissions may influence temperature, their impact on precipitation is statistically insignificant. This implies that other climatic or anthropogenic factors may concurrently drive precipitation variability.

4. Conclusion

This study systematically investigates the dynamic interactions between fossil fuel consumption (coal, petroleum, natural gas) and climate indicators (temperature, precipitation, CO_2 emissions) in Sichuan Province from 2012 to 2022, integrating time series analysis and causal inference methodologies. Anomalies in energy and climate datasets were identified via boxplot analysis and corrected using Lagrange interpolation and moving average techniques. Polynomial fitting, multiple linear regression, Pearson correlation, and Granger causality tests were employed to model these relationships. Key findings include:

Strong Energy-Carbon Linkage: Fossil fuel consumption exhibits a significant positive correlation with CO_2 emissions (correlation coefficient: 0.9654), with coal demonstrating the strongest driving effect (regression coefficient: 0.9055).

Causal Impact on Climate: Granger causality tests confirm CO_2 emissions as a causal driver of temperature variation (lag order: 3, p-value <0.05), though no significant causal relationship with precipitation is observed.

Nonlinear Climate Responses: Weak positive correlation between CO_2 emissions and temperature (Pearson: 0.2106) and weak negative correlation with precipitation (-0.3616) suggest indirect climatic effects mediated through nonlinear pathways.

To mitigate climate risks, Sichuan Province should prioritize accelerating its energy transition by scaling up renewable energy adoption, optimizing water resource management to address precipitation variability, and enhancing technological innovation to improve energy efficiency. While this study focuses on provincial-scale data over a decade, future research should extend to longer-term and cross-regional analyses. Advanced modeling techniques, such as BP neural networks and deep learning, could further unravel the complex mechanisms of energy-climate interactions, providing refined decision-making frameworks for global climate governance.

References

[1] Dai Z P. Research on China's low-carbon energy transition pathways based on the LEAP model [D]. Shanghai University of Finance and Economics, 2021.

[2] Tang X Y. Analysis of the Current Situation and Future Development of CO2 Emissions in China's Provincial Energy Consumption [J]. 2024, 44 (1): 180 – 189.

[3] Wang Y N, Tang J L, Luo X T, et al. Impact of Clean Energy Utilization on Climate Change ——with Carbon Dioxide as the Medium [J]. 2022, 49 (13): 148 – 151.

[4] Xu H N, Xiao T G, Yang M X, Zhang W Y. Analysis of summer precipitation variation characteristics in southwestern China [J]. Frontiers in Earth Science (Hans), 2019, 9 (10): 908 – 920.

[5] Yang M X, Xiao T G, Li Y, et al. Evaluation and Projection of Climate Change in Southwest China Using CMIP6 Models [J]. Plateau Meteorology, 2022, 41 (6): 1557 - 1571.

[6] Xiang B, Yuan Y P, Yang X J, et al. Renewable Energy Utilization for Sichuan Province Based on Carbon Pinch Analysis [J]. Sichuan, 2016, 30 (5): 615 – 621.

[7] Li X. Preliminary Study on Energy Saving and Emission Reduction in Sichuan Province Based on the Relationship among Energy Consumption, Carbon Emission, and Economic Growth [D]. Southwest Jiao tong University, 2020.

[8] China Meteorological Administration Climate Change Centre. China Climate Change Blue Book 2023 [M]. Beijing: Science Press, 2023.

[9] Long T G. Study on environmental impact assessment and monitoring methods under climate change [J]. 2024, 37 (3): 77 – 79.

[10] Mao S S, Cheng Y M, Pu X L. Probability theory and mathematical statistics tutorial [M]. Higher Education Press, 2011.