Mini-Project 2 – Interpolation and Curve Fitting

Kayla Oviedo

October 7, 2024

Abstract

Interpolation and curve fitting are fundamental techniques in numerical analysis and data science used to estimate unknown values and analyze data trends. Interpolation entails generating new data points within the confinements of a discrete set of known data points. It guarantees that the curve intersects precisely at those points. Prevalent techniques include linear, polynomial, and spline interpolation. Curve fitting approximates the underlying data trend by identifying a curve that best conforms to a set of data points, frequently using least-squares methods. In contrast to interpolation, curve fitting doesn't require that the curve intersects the actual data points, hence facilitating the modeling of noisy data. These methodologies are extensively utilized in engineering, science, and statistics for data analysis, forecasting, and simulations.

1 Introduction

This research examines the utilization of interpolation and curve fitting methodologies for predicting and modeling the production of electricity from renewable resources across multiple countries. Linear interpolation is used to estimate absent data points between known values, ensuring the curve intersects the real data points. Conversely, curve fitting techniques, such as quadratic, exponential, and logarithmic models, are utilized to approximate the overarching trend of the data, despite the presence of noise or anomalies.

The objective of this study is to implement these methodologies on GDP data pertaining to renewable energy production, offering insights into their precision in forecasting future values, including the GDP for 2020. The methods are assessed using both numerical and visual analysis to determine their efficacy in simulating historical and prospective trends. We utilize linear interpolation to estimate data for 2015 and subsequently investigate curve fitting to forecast GDP figures through 2020.

This paper is organized as follows: Section 1 outlines the theoretical foundation of interpolation and curve-fitting methodologies. Section 2 explores the technique, describing the implementation of the models. Section 3 presents the results and analysis for each country's dataset, while Section 4 ends with important observations and recommendations for future studies.

1.1 Importance of Linear Algebra

Linear algebra is pivotal in solving consistent linear systems of equations, which is essential for both interpolation and linear least squares formulations.

1.2 Interpolation

Interpolation estimates unknown values between known data points, allowing the construction of smooth curves from discrete points. This method is beneficial for filling gaps in time series data, though it is limited to estimating values within the bounds of the given dataset.

1.3 Curve Fitting

Unlike interpolation, linear and nonlinear curve fitting optimizes a function to a dataset by minimizing the overall error between data points and the curve. This technique enables predictions and extrapolations beyond the original data bounds, making it invaluable for analyzing trends in electricity production.

2 Problem Statement

The goal of this project is to analyze the trends in electricity production from renewable resources across multiple countries, using interpolation and curve fitting techniques. Specifically, the project aims to model and compare the growth patterns of renewable energy production from 2000 to 2015 using linear interpolation, exponential fits, and logarithmic fits. This analysis will help determine the rate of growth or decay of electricity production over time, providing insights into each country's progress towards renewable energy targets.

3 Methodology

3.1 Data Collection

The dataset includes annual electricity production (kWh) from renewable sources for the following countries from 2000 to 2015:

- **China**
- **Germany**
- **United Kingdom**
- **India**

- **Japan**

3.2 Math Formulas

Linear Interpolation

The formula for a point between two known points (x_0, y_0) and $(x_1, y_1$ is:

$$y = y_0 + \frac{(x - x_0)}{(x_1 - x_0)}(y_1 - y_0)$$

This formula calculates the interpolated value y at any given point x

Exponential Fit

$$y = ae^{bx} + c$$

Where:

- a is a scaling factor
- b is the growth or decay rate
- x is the independent variable (years)
- c is the vertical shift

Logarithmic Fit

For logarithmic curve fitting, the model used is:

$$y = alog(x) + b$$

Where:

- a is the scaling factor
- x is the independent variable (years)
- b is a constant

3.3 Analysis Process

1. **Linear Interpolation**: Utilizes known data points to estimate GDP for 2015. (Figure 1)

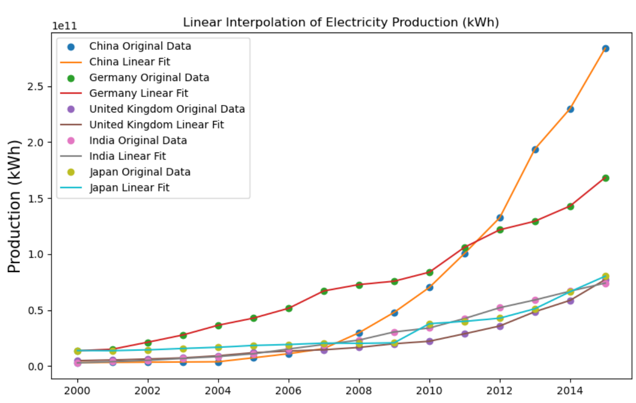

Figure 1: Linear Interpolation of Electricity Production

2. **Exponential Model Fitting**: Normalizes the years and GDP data to fit an exponential model, predicting GDP for 2020. (Figure 2)

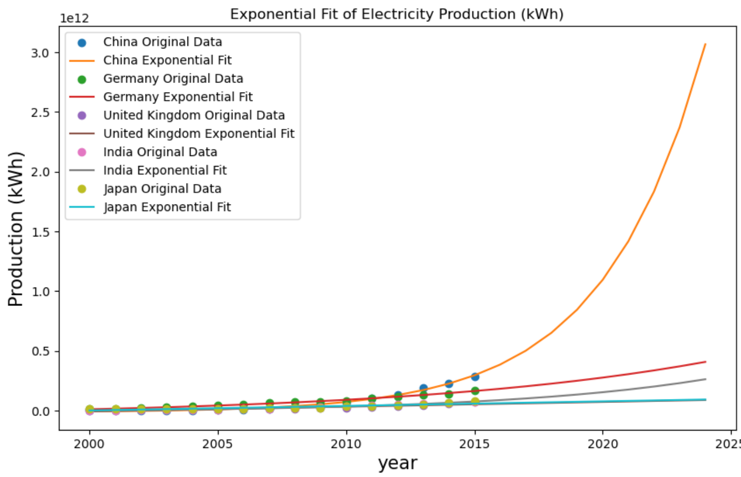

Figure 2: Exponential Fit of Electricity Production (kWh)

3. **Logarithmic Model Fitting**: Fits a logarithmic model to provide an alternative analysis of the data. (Figure 3)

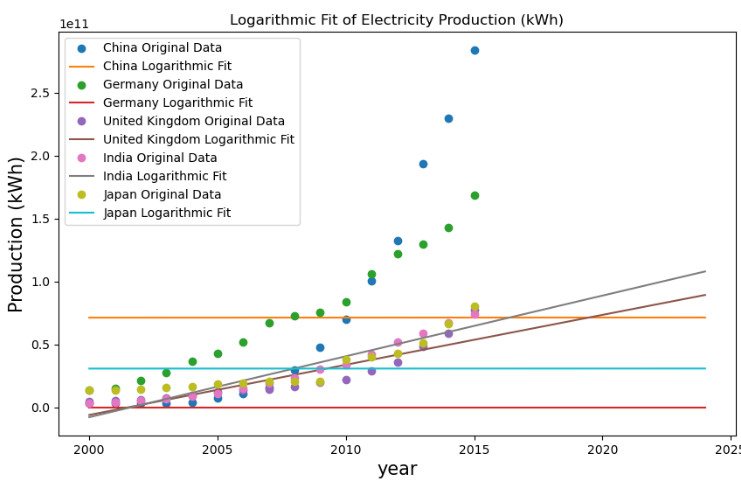

Figure 3: Logarithmic Fit of Electricity Production (kWh)

3.4 Code Implementation

The Python code uses libraries to perform the above analyses, producing visualizations for each country to illustrate trends in electricity production. If b is positive, it indicates that electricity production is experiencing exponential growth, meaning production increases more rapidly as time progresses. If b is negative, it indicates exponential decay, meaning production decreases over time.

Country	**Exponential Fit Parameter** b
China	$2.56e-01$
Germany	$8.50e-02$
United Kingdom	$-1.49e-05$
India	$1.25e-01$
Japan	$-1.16e-05$

Table 1: Exponential Fit Parameter b Values for Electricity Production

4 Results

The exponential model provided insights into the growth rates of renewable energy production. The b values obtained indicate the rate of growth. Countries with positive b values demonstrated a rapid increase in electricity production from renewable sources. Countries with negative b values showed a decline in production, suggesting challenges in renewable energy implementation. Predicted electricity production for 2020 was generated, allowing for comparisons against future targets.

- **Linear Interpolation**: Estimated GDP for 2015 is presented alongside the original data points.
- **Exponential and Logarithmic Fits**: GDP predictions for 2020 are derived from the fitted models, showing variations across countries.

4.1 Example Output

For **China**, the results were as follows:

- **GDP in 2015 (Linear Interpolation)**: $2.84e+11$ kWh
- **GDP in 2020 (Exponential Fit)**: $1.09e+12$ kWh
- **Exponential Fit Parameter b**: $2.56e-01$

5 Conclusion

This project demonstrates the application of linear algebra, interpolation, and curve-fitting techniques in analyzing electricity production from renewable sources. The findings offer insights into production trends and facilitate future forecasting based on historical data. A comparison of the three models (linear, exponential, and logarithmic) revealed different growth patterns: Linear models often provided a straightforward trend but did not capture accelerated growth phases. Exponential fits highlighted rapid growth potential, particularly in countries with aggressive renewable energy policies, and logarithmic models showcased a leveling off, suggesting potential constraints in further growth.