Konnichiwa!

Welcome to My Forecasting Project on Temperature & Salinity

IHI

Understanding the Future: Temperature & Salinity Forecasting

Why predicting temperature & salinity is important

climate monitoring

marine life impact

The Data That Drives Our Predictions

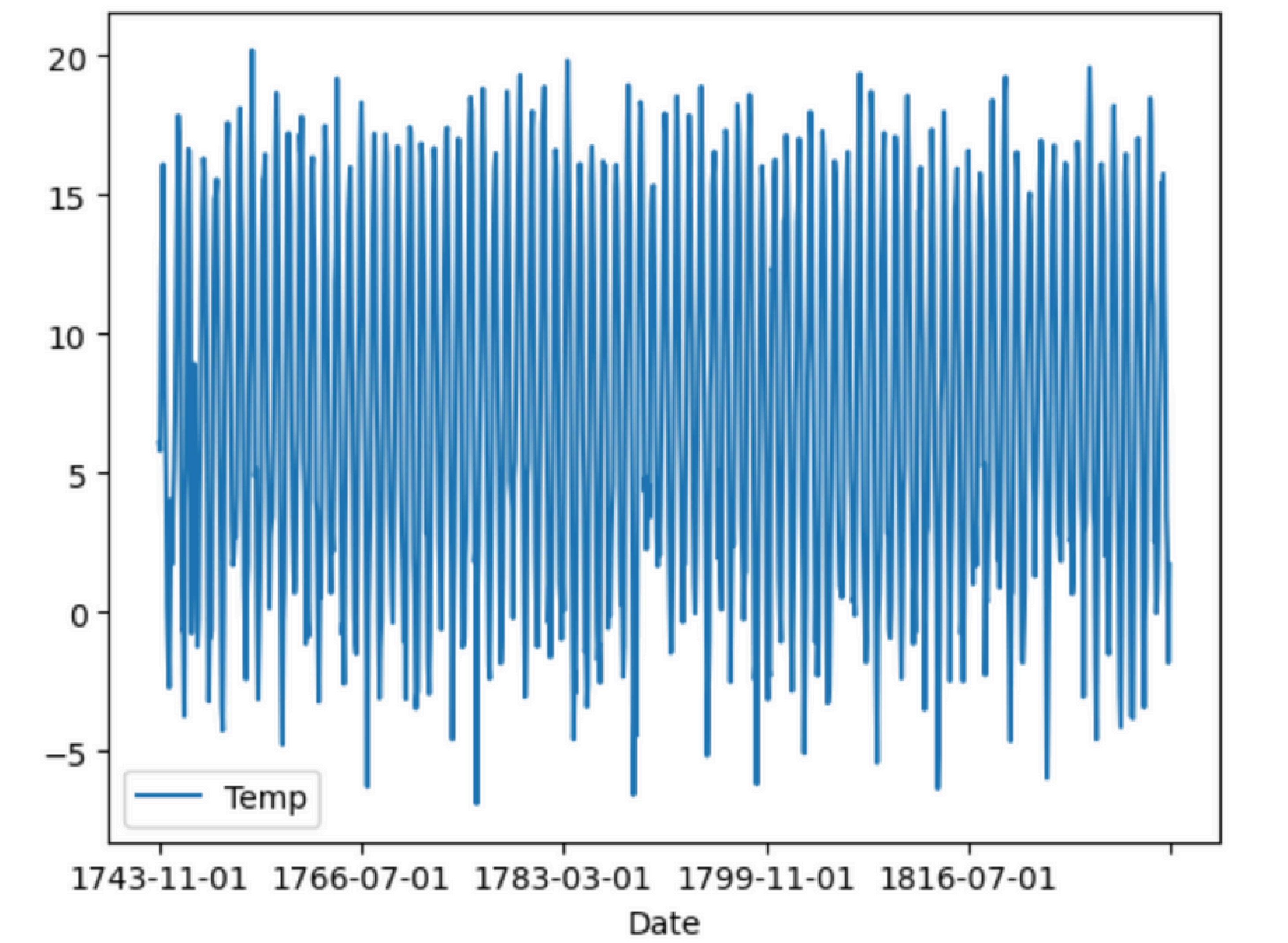
Dataset contains 1000 rows with historical temperature and salinity values.

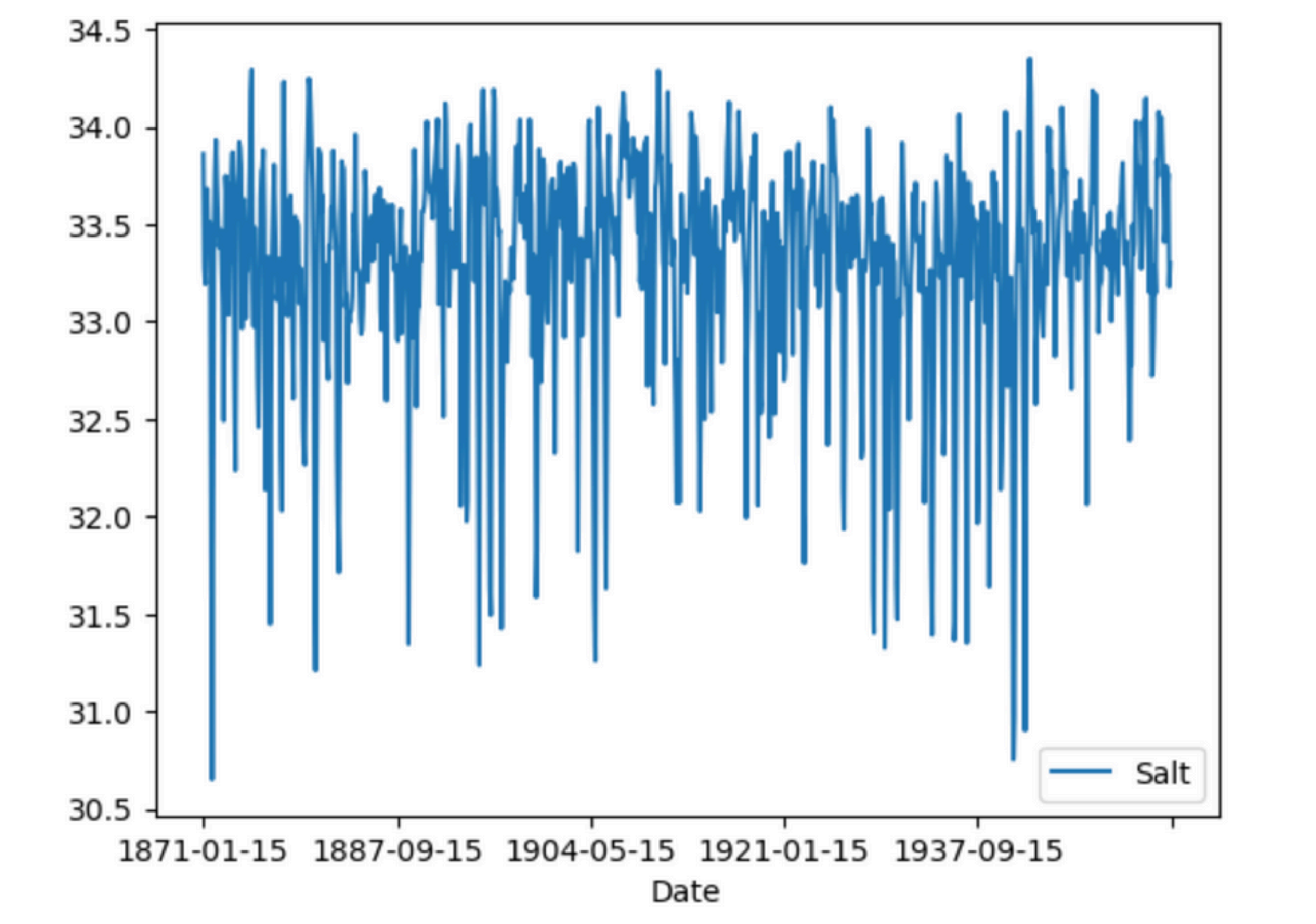
1743-1833

Temperature

Salinity

1871-1954





Choosing the Right Approach

IHI

Linear Regression

Basic benchmark model

Random Forest

Captures non-linearity well.

SARIMA:

Used for time-series trends & seasonality.

Model	Pros	Cons
Linear Regression	- Simple and interpretable.- Fast computation.	- Assumes a linear relationship, which may not hold for complex patterns.- Less accurate for non-linear time series.
Random Forest	- Captures non-linearity.- High accuracy with low MAE & RMSE.- Robust to overfitting with enough trees.	- Computationally expensive.- Less interpretable compared to linear models.
SARIMA	- Good for capturing seasonality and trends.- Works well for time series forecasting.	- High computation time.- Requires extensive parameter tuning.- Poor performance if data is highly irregular.

SARIMA approach

import libraries
-numpy
-pandas
-matplotlib

Load CSV file
set Date as index

Plot Graph to get initial understanding

Perform Augmented Dickey-Fuller (ADF) test

H0--> non stationary
H1--> stationary

p-value<=0.05
(reject H0)
STATIONARY

BRAVO
MOVE AHEAD!!

p-value>0.05
(accept H0)
NON STATIONARY

Do Seasonal Differencing

ADF test again

STATIONARY

NON STATIONARY

/

Other techniques

SARIMA approach

statsmodels library
- Adfuller
- PACF, ACF
- SARIMAX

Why SARIMA ? and not ARIMA

SARIMA is preferred over ARIMA and ARMA for temperature time series forecasting because it captures seasonality, which ARIMA and ARMA cannot model effectively.

Plot PACF and ACF functions

Moving average order (ACF)

Auto regression order (PACF)

Trained SARIMAX model
ARIMA order-(1,1,1)
Seasonal order-(1,1,1,12)

Predicted for last 200 rows

Compared with actual data

Predicted for future 1 year

Predicted for future 5 years

Why (1,1,1,12) and not (2,0,2)?

tried (2,0,2), overfit short-term trends, (1,1,1)(1,1,1,12) improved results - capturing trend and seasonality.

Linear and Random forest

IHI

Linear Regression

pandas, numpy, matplotlib, scikit-learn

extract year, month, & day from the 'Date' column

X (independent)

Temperature- Y (dependent)

Fit model for next 5 years

Random Forest

pandas, numpy, matplotlib, scikit-learn

extract year, month, & day from the 'Date' column

X (independent)

Temperature- Y (dependent)

Fit model for next 5 years

Results

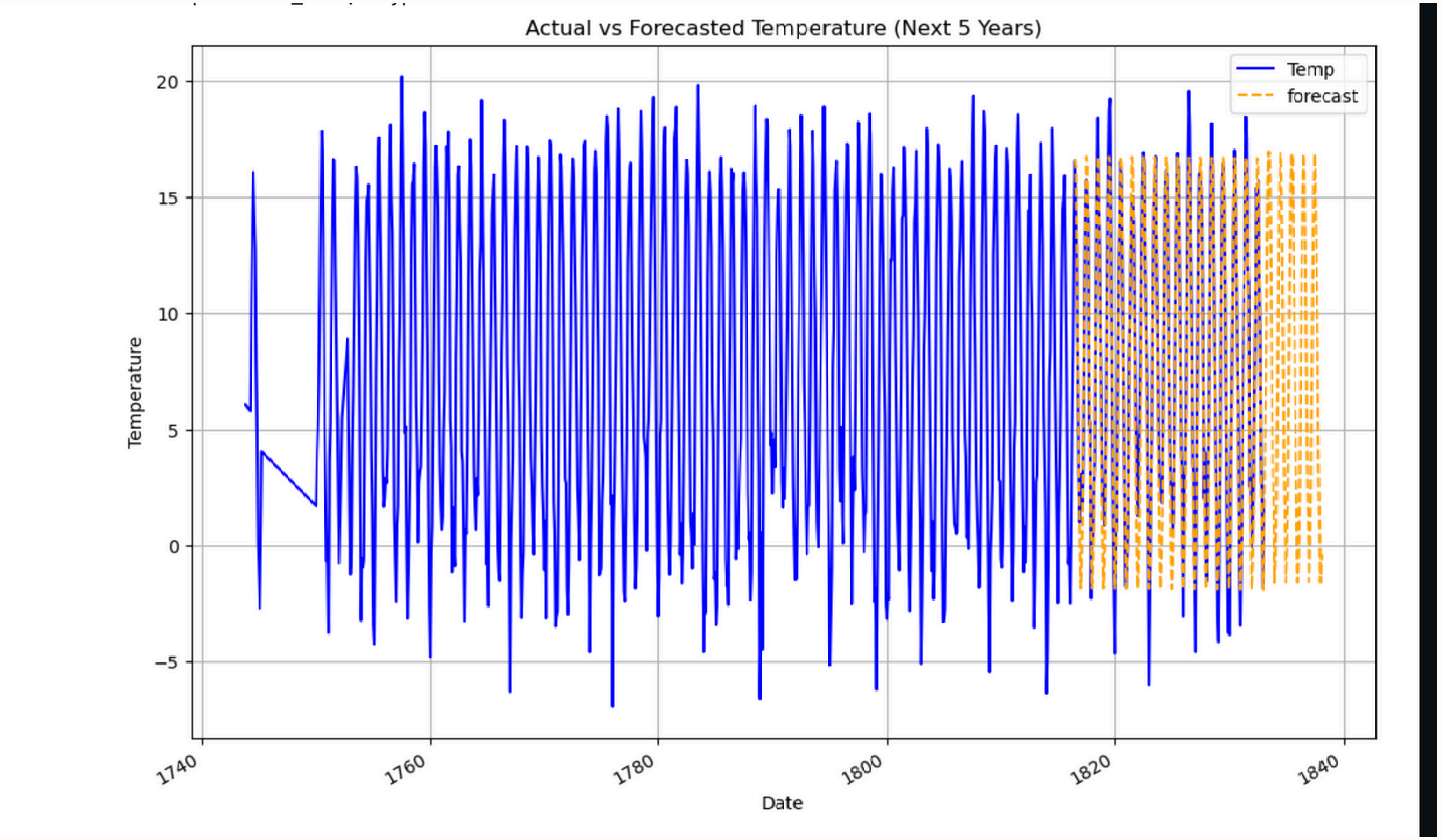
IHI

Model		Temperature	Salinity
Linear Regression	- Training time - Prediction time - MAE - RMSE	-0.0583 seconds -0.0140 seconds -0.3858 -0.5312	-0.0111 seconds -0.0018 seconds -0.3858 -0.5312
Random Forest	- Computation time - MAE - RMSE - R2 score	-0.1449 seconds -0.1101 -0.1497 -0.9290	-0.1556 seconds -0.1101 -0.1497 -0.9290
SARIMA	- Computation time - MAE - RMSE - R2 score	-3.0846 seconds -1.5398 -2.0988 - 0.9036	-2.3560 seconds -0.3311 -1.2525 -(-3.9655)

RESULTS

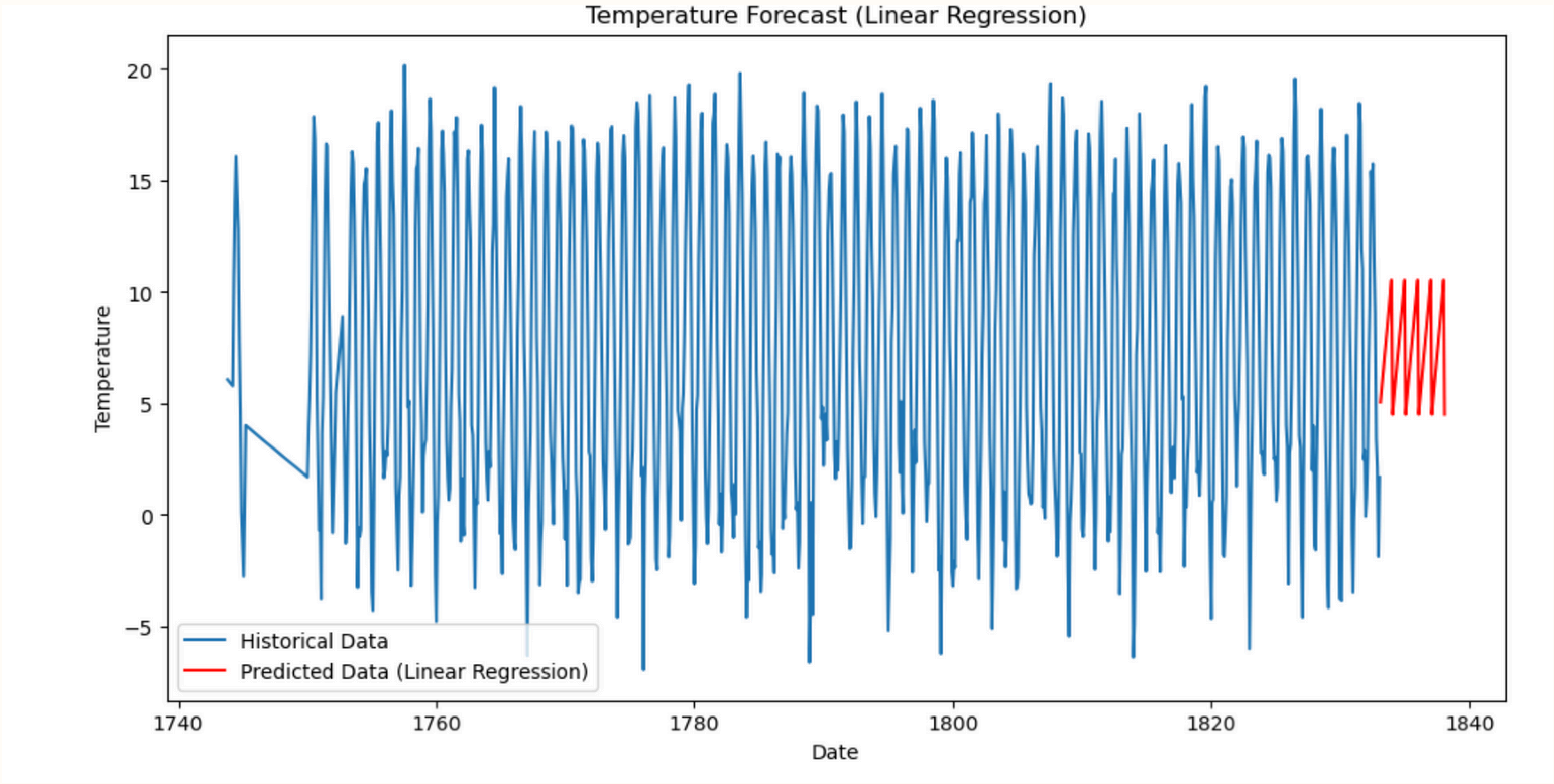
For salinity, Random Forest performed the best with the lowest MAE (0.1101) and RMSE (0.1497), along with a high R^2 (0.9290), making it the most accurate model. For temperature, SARIMA had the highest R^2 (0.9036), but Random Forest still outperformed in terms of lower MAE (0.1101) and RMSE (0.1497), making it the most reliable choice for both variables.

SARIMA

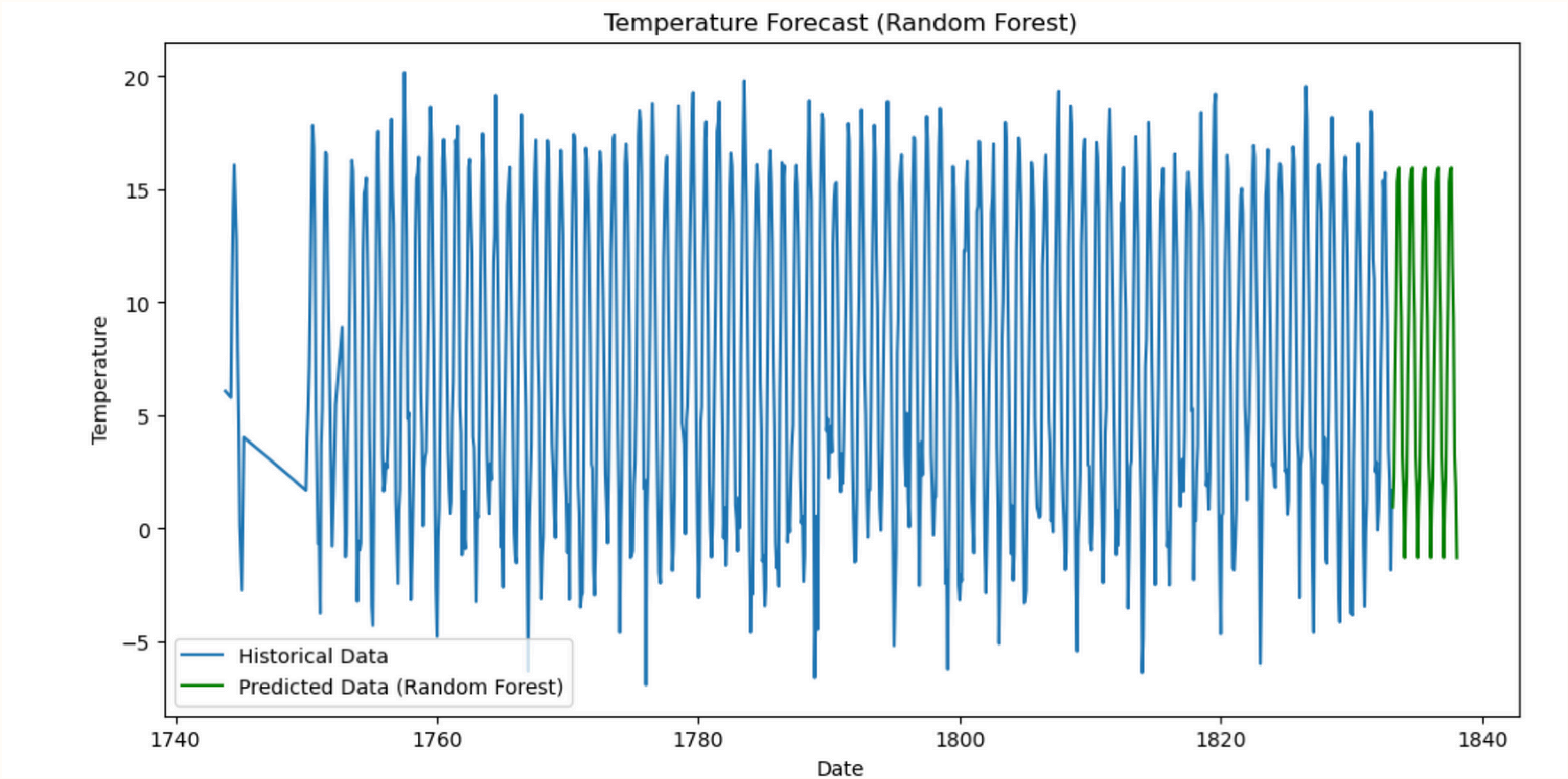


Actual vs Forecasted Temperature (Next 5 Years)
Temp
forecast
Temperature
20
15
10
5
0
−5
Date
1740
1760
1780
1800
1820
1840

Linear


Temperature Forecast (Linear Regression)
20
15
10
5
0
−5
Temperature
1740
1760
1780
1800
1820
1840
Date
Historical Data
Predicted Data (Linear Regression)

Random



Temperature Forecast (Random Forest)
Temperature
Date
20
15
10
5
0
−5
1740
1760
1780
1800
1820
1840
Historical Data
Predicted Data (Random Forest)

IHI

link-(https://www.canva.com/design/DAGea3qHKSU/UCeQGySAZA4oefDzbsesEw/edit?utm_content=DAGea3qHKSU&utm_campaign=designshare&utm_medium=link2&utm_source=sharebutton)

Thank You