



Universidad Tecnológica De Panamá
Vicerrectoría de Investigación, Postgrado y Extensión Facultad de
Ingeniería de Sistemas Computacionales

Maestría en Analítica de Datos
Modelos Predictivos
Grupo 1AN-212

Proyecto Final:
Análisis y Propuestas de Modelos para Predecir el Consumo de
Energía Eléctrica en Panamá para el año 2024

Profesor:
Juan M Castillo, PhD

Integrante:
William Vásquez

Año: 2024



Introducción

En la actualidad, la predicción del consumo de energía es vital para la planificación y gestión eficiente de los recursos. Este proyecto tiene como objetivo evaluar diferentes modelos de pronósticos y series temporales para estimar el consumo de energía eléctrica en Panamá durante los primeros meses de 2024. Se utilizará datos históricos de consumo mensual en MWh desde 2021 hasta 2023 por distribuidora de energía eléctrica (ENSA, EDEMET, EDECHI), estos datos se obtuvieron del sitio: <https://www.datosabiertos.gob.pa/>.

Se implementarán y compararán varios modelos de pronósticos, incluyendo Prophet, suavización exponencial simple, Holt, Holt-Winters y regresión lineal. Los resultados serán evaluados utilizando métricas como el error porcentual absoluto medio (MAPE) y Desviación media absoluta (MAD)

El análisis se llevará a cabo utilizando Python y diversas librerías especializadas en análisis de series temporales y machine learning. Los resultados obtenidos permitirán conocer el modelo más adecuado para predecir el consumo energético y servir como base para la toma de decisiones informadas en relación con la gestión de la energía y la planificación de la demanda en Panamá.

Análisis de Datos

Los datos ya limpiados de forma manual se encuentran en el archivo “Datos Consumo electrico por distribuidora.xlsx”

ENSA:

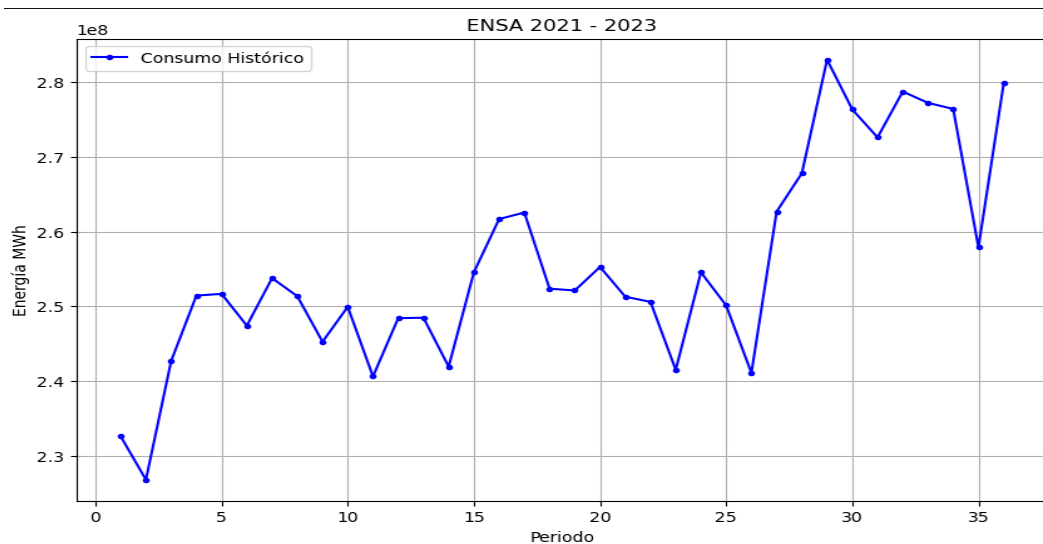


Fig. 1 Gráfica del consumo eléctrico de clientes de ENSA 2021-2023

En la fig.1 se puede observa como varia el consumo eléctrico de clientes de ENSA a lo largo de 3 años o 36 periodos mensuales, en la gráfica se puede ver que mantiene cierta tendencia de aumento.

```
#Promedio Movil
data["Promedio Movil"] = data["Energia MWh"].rolling(window=4).mean()

forecast_ma = data["Energia MWh"].rolling(window=4).mean().iloc[-1]
forecast_ma

272805769.5

# MAD
mad_prom_movil = data["Energia MWh"].sub(data["Energia MWh"].rolling(window=4).mean().shift(1)).abs().mean()
print(f"MAD: {mad_prom_movil}")

# MAPE
mape_prom_movil = data["Energia MWh"].sub(data["Energia MWh"].rolling(window=4).mean().shift(1)).abs().div(data["Energia MWh"]).mean() * 100
print(f"MAPE: {mape_prom_movil}%")

MAD: 7029155.5859375
MAPE: 2.707807282711413%
```

Fig. 3 Código Utilizado para implementar el Promedio móvil a los datos.

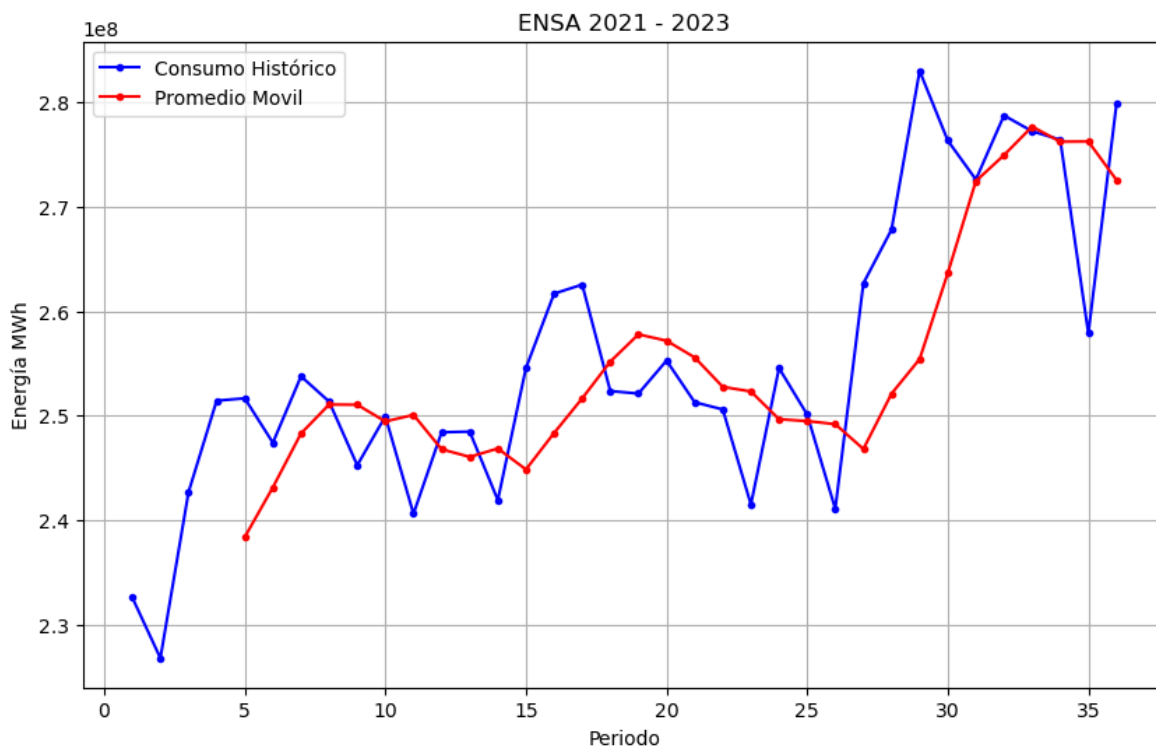


Fig. 2 Gráfica comparativa de promedio móvil y los datos reales

En la fig.2 se puede ver que el promedio móvil suaviza las variaciones a corto plazo, y se puede observar mejor la tendencia de los datos.

```
# Ajustar el modelo de suavización exponencial simple
model_exp_simple = SimpleExpSmoothing(data['Energia MWh'])
model_fit_exp_simple = model_exp_simple.fit()

# Predecir los próximos 6 periodos
forecast_exp_simple = model_fit_exp_simple.forecast(steps = 2)
print(forecast_exp_simple)

# Cálculo de errores
actual = data['Energia MWh'].iloc[-1]
fitted_values_expsimple = model_fit_exp_simple.fittedvalues

# MAD
mad_exp_simple = np.mean(np.abs(data['Energia MWh'] - fitted_values_expsimple))
print(f"MAD: {mad_exp_simple}")

# MAPE
mape_exp_simple = np.mean(np.abs((data['Energia MWh'] - fitted_values_expsimple) / data['Energia MWh'])) * 100
print(f"MAPE: {mape_exp_simple}%")
```

Fig. 5 Código Utilizado para implementar la suavización exponencial simple a los datos.

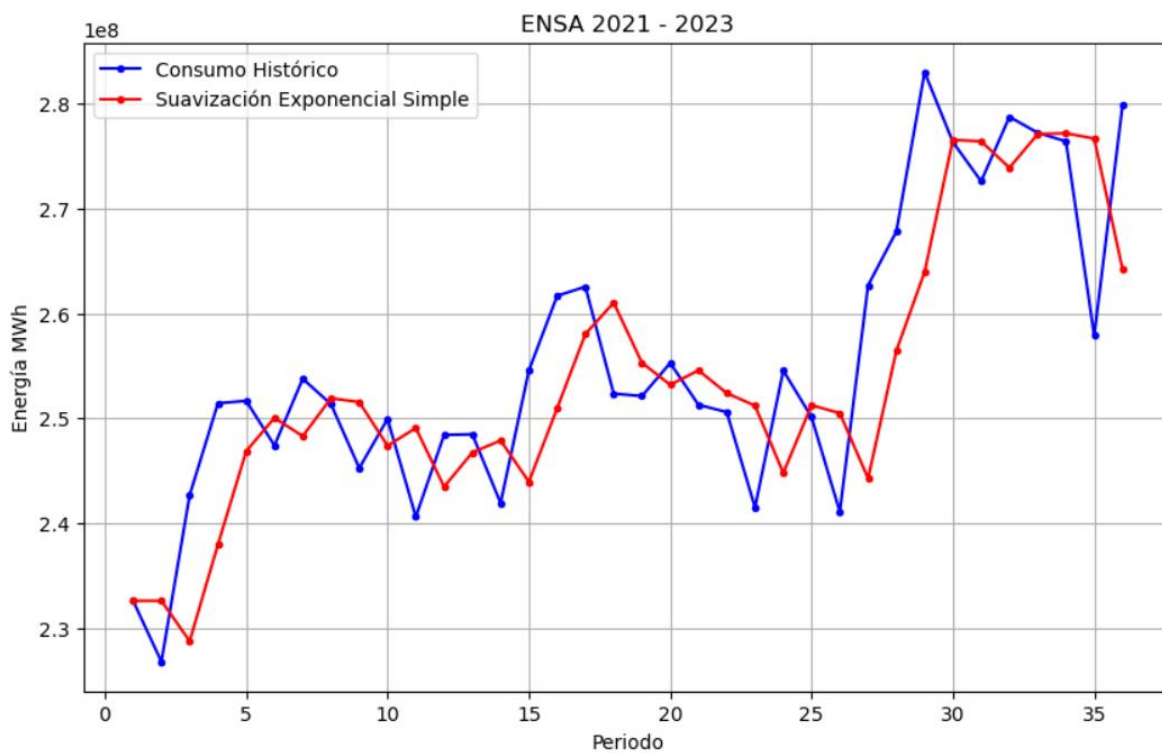


Fig. 4 Gráfica comparativa de suavización exponencial simple y los datos reales

```
# Ajustar el modelo Holt
model_holt = Holt(data['Energia MWh']).fit()

# Pronóstico
forecast_holt = model_holt.forecast(steps=8)
print(forecast_holt)

# Cálculo de errores
fitted_values_holt = model_holt.fittedvalues

# MAD
mad_holt = np.mean(np.abs(data['Energia MWh'] - fitted_values_holt))
print(f"MAD: {mad_holt}")

# MAPE
mape_holt = np.mean(np.abs((data['Energia MWh'] - fitted_values_holt) / data['Energia MWh'])) * 100
print(f"MAPE: {mape_holt}%")
```

Fig. 6 Código Utilizado para implementar el modelo de Holt a los datos

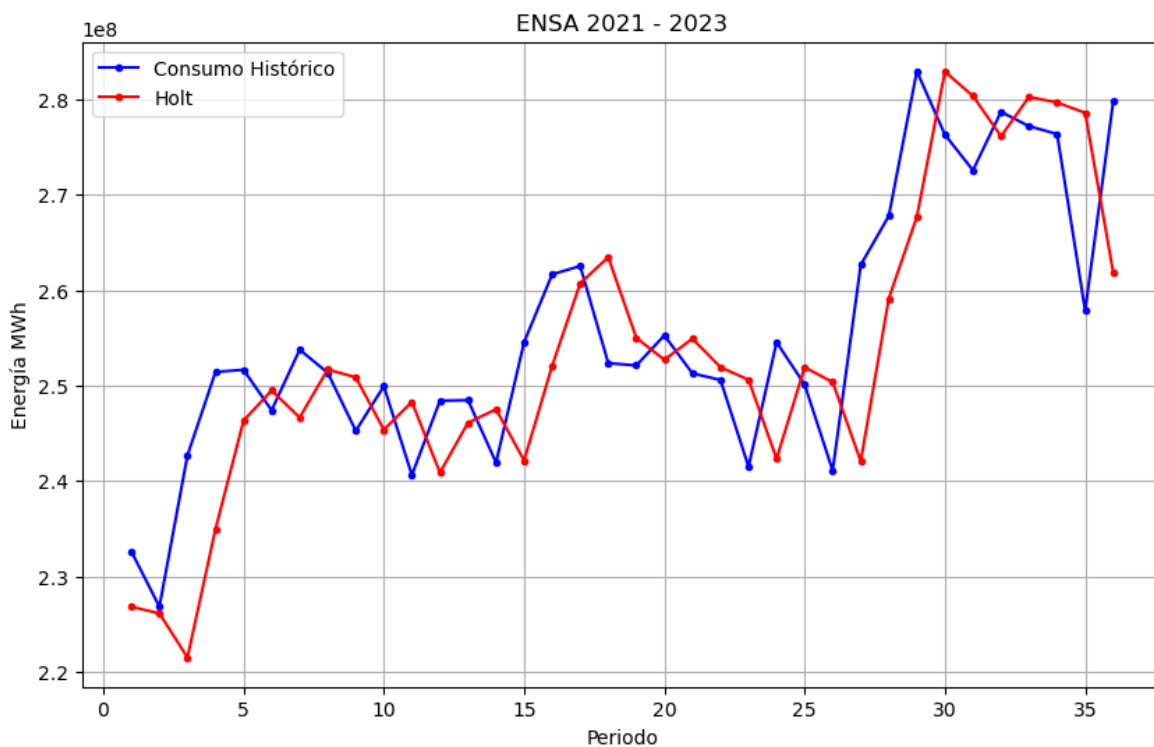


Fig. 7 Gráfica datos reales y valores predichos por el modelo de Holt

En la Fig. 7 El modelo de Holt parece seguir la tendencia general de los datos históricos, pero se puede ver que mantiene un desfase, también se observa que

existen periodos donde la predicción del modelo se desvía significativamente de los datos reales, especialmente en los picos.

```
# Modelo Holt-Winters
model_HW = ExponentialSmoothing(data['Energia MWh'], seasonal='add', seasonal_periods=12).fit()
forecast_HW = model_HW.forecast(steps=8)
print(forecast_HW)

# Cálculo de errores
fitted_values_HW = model_HW.fittedvalues

# MAD
mad_HW = np.mean(np.abs(data['Energia MWh'] - fitted_values_HW))
print(f"MAD: {mad_HW}")

# MAPE
mape_HW = np.mean(np.abs((data['Energia MWh'] - fitted_values_HW) / data['Energia MWh'])) * 100
print(f"MAPE: {mape_HW}%")
```

Fig. 8 Código Utilizado para implementar el modelo de Holt-Winters a los datos

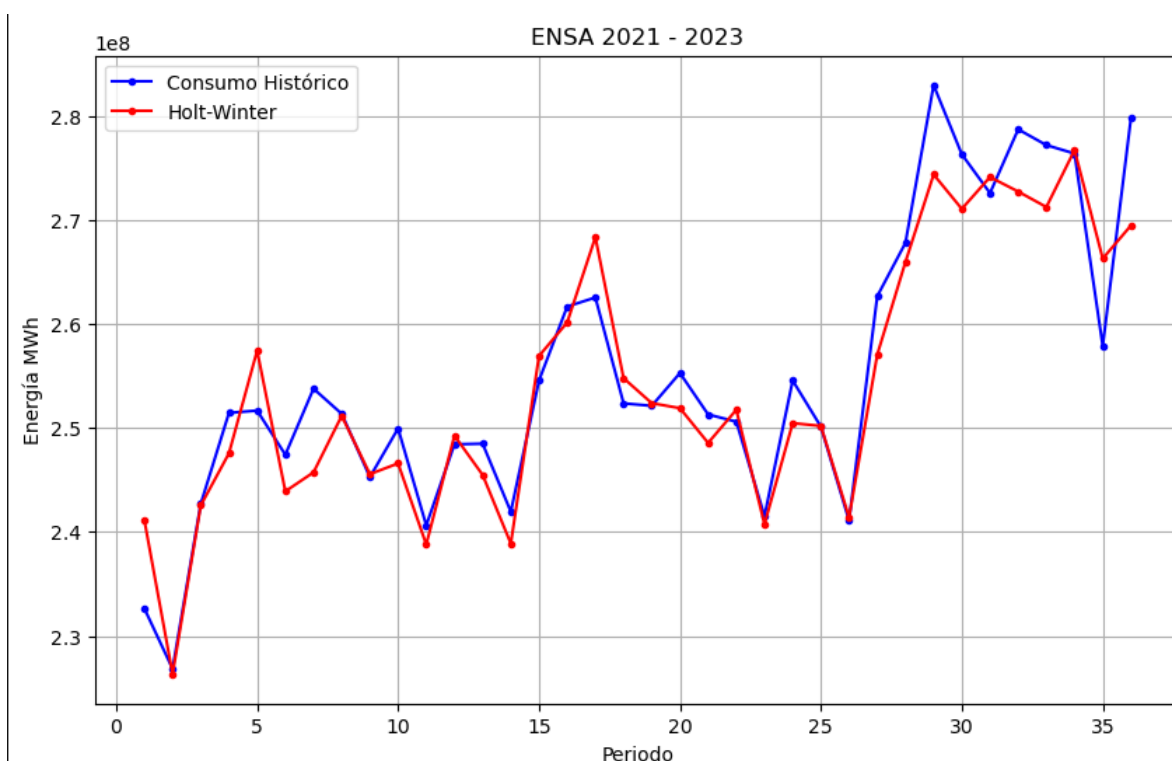


Fig. 9 Gráfica datos reales y valores predichos por el modelo de Holt-Winters

En la Fig. 9 las predicciones de Holt-Winters se ajusta mejor a los picos y valles estacionales de los datos en comparación con el modelo de Holt visto en la Fig.7,

sin embargo, aún se observan algunas desviaciones entre las predicciones y los datos reales, especialmente en algunos puntos específicos.

```
# Regresión Lineal
X = data.index.values.reshape(-1, 1)
y = data['Energia MWh']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)
model_RL = LinearRegression().fit(X_train, y_train)
predictions_RL = model_RL.predict(X_test)
print(predictions_RL)

# Cálculo de errores
fitted_values_RL = model_RL.predict(X)

# MAD
mad_RL = np.mean(np.abs(y - fitted_values_RL))
print(f"MAD: {mad}")

# MAPE
mape_RL = np.mean(np.abs((y - fitted_values_RL) / y)) * 100
print(f"MAPE: {mape}%")
```

Fig. 10 Código Utilizado para implementar Regresión lineal a los datos.

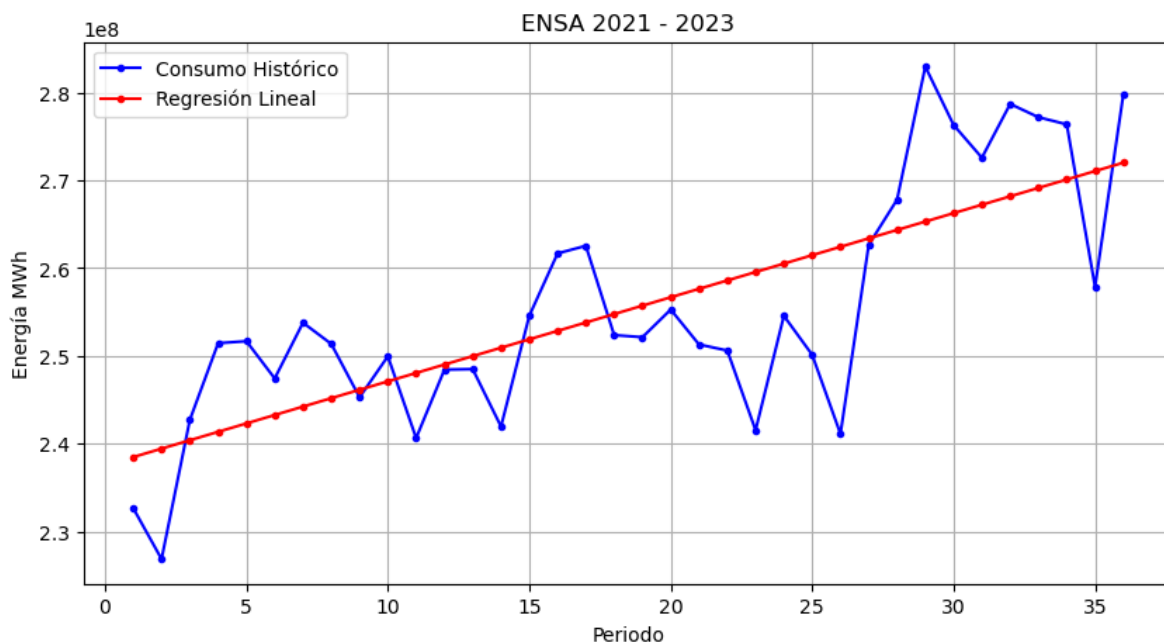


Fig. 11 Gráfica comparativa de la regresión lineal y los datos reales


```
# Modelo SARIMA
model_SARIMAX = SARIMAX(data['Energia MWh'], order=(1, 1, 1), seasonal_order=(1, 1, 1, 12)).fit()
forecast_SARIMAX = model_SARIMAX.forecast(steps=8)
print(forecast_SARIMAX)

# Cálculo de errores
fitted_values_SARIMAX = model_SARIMAX.fittedvalues

# MAD
mad_SARIMAX = np.mean(np.abs(data['Energia MWh'] - fitted_values_SARIMAX))
print(f"MAD: {mad}")

# MAPE
mape_SARIMAX = np.mean(np.abs((data['Energia MWh'] - fitted_values_SARIMAX) / data['Energia MWh'])) * 100
print(f"MAPE: {mape}%")
```

Fig. 12 Código Utilizado para implementar el modelo de Sarima a los datos

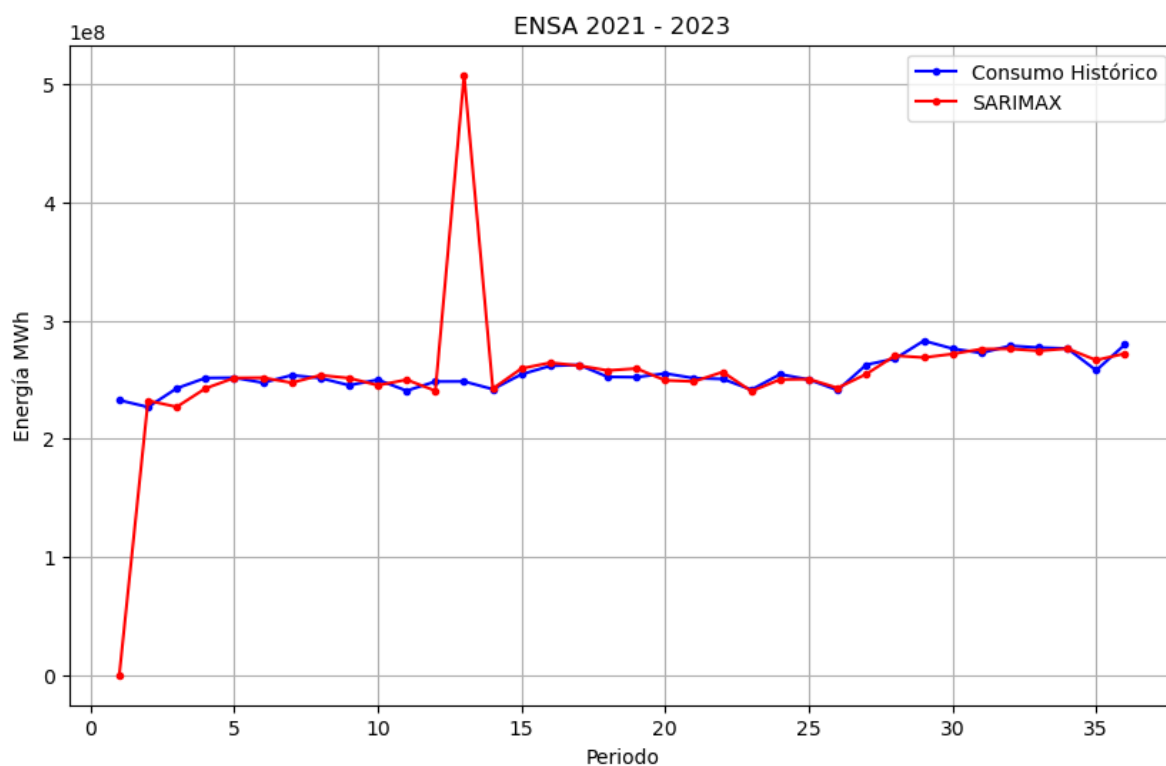


Fig. 13 Gráfica datos reales y valores predichos por el modelo de Sarima

El modelo SARIMA (Fig. 13) es efectivo, ya que sigue de cerca los datos históricos.

No hay desviaciones significativas entre los valores predichos y los reales, excepto por un pico grande en el periodo 13 (enero- 2022) que el modelo predijo sin embargo el valor real fue mucho menor.

El modelo SARIMA parece ser capaz de capturar los patrones generales del consumo eléctrico de ENSA durante el periodo analizado.

```
# Convertir la columna de fechas a tipo datetime
df['ds'] = pd.to_datetime(df['ds'])

# Inicializar y ajustar el modelo
model = Prophet()
model.fit(df)

# Crear un DataFrame para futuras predicciones (ajusta periods según tus necesidades)
future = model.make_future_dataframe(periods=12, freq='M')

# Usar el modelo para hacer predicciones
forecast_prophet = model.predict(future)

# Imprimir o graficar las predicciones según sea necesario
print(forecast_prophet[['ds', 'yhat', 'yhat_lower', 'yhat_upper']])

# Cálculo de errores
fitted_values_prophet = model.predict(df)['yhat']

# MAD
mad_prophet = np.mean(np.abs(df['y'] - fitted_values_prophet))
print(f"MAD: {mad}")

# MAPE
mape_prophet = np.mean(np.abs((df['y'] - fitted_values_prophet) / df['y'])) * 100
print(f"MAPE: {mape}%")
```

Fig. 14 Código Utilizado para implementar el modelo de Prophet a los datos.

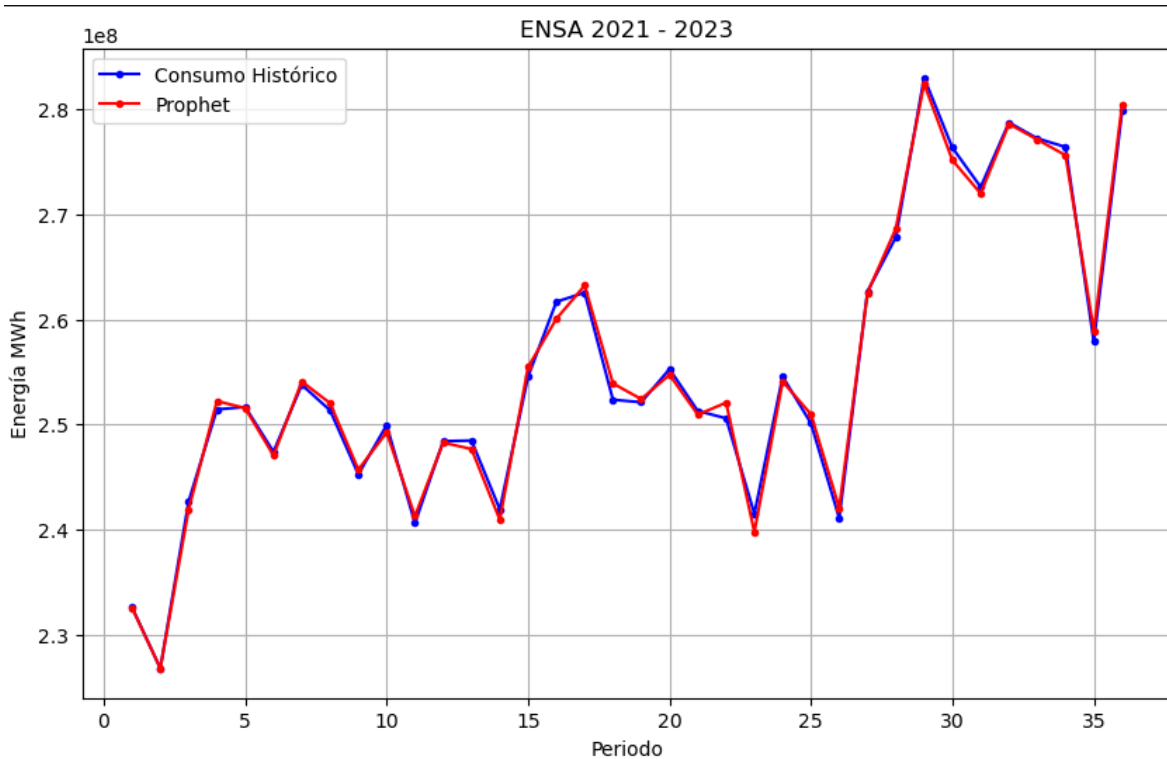


Fig. 15 Gráfica datos reales y valores predichos por el modelo de Prophet

El modelo Prophet (Fig. 15) sigue de cerca los datos históricos en la mayoría de los periodos, lo que indica que captura bien el patrón del consumo de energía eléctrica.

Hay desviaciones, como el del período 30, donde el modelo subestima el consumo real, pero estas discrepancias son menores.

Evaluaciones de Métricas de Error (ENSA)

En todas las imágenes donde se presentó el código utilizado, hay un fragmento donde se calcula dos métricas de error que son MAD y MAPE.

A continuación, se muestra un resumen de estos valores obtenidos, por cada modelo implementado en este caso de estudio:

	Modelo	MAD	MAPE	STD
0	Promedio Movil	7.02916e+06	2.70781	8.78644e+06
1	Suavización Exponencial	6.77762e+06	2.64429	8.47203e+06
2	Holt	7.67868e+06	2.99421	9.59835e+06
3	Holt-Winter	3.37846e+06	1.30256	4.22308e+06
4	Sarimax	1.83722e+07	7.52003	2.29652e+07
5	Regresión	7.33132e+06	2.87748	9.16415e+06
6	Prophet	675533	0.26506	844417

Tal y como se observó gráficamente el modelo de Prophet tiene los valores más bajos de MAD (675,533), MAPE (0.26506), y STD (844,417), lo que indica que es el más preciso y consistente para predecir el consumo eléctrico.

Holt-Winter: También muestra un buen rendimiento con valores relativamente bajos de MAD (3,378,460), MAPE (1.30256), y STD (4,223,080), siendo una opción viable.

El modelo SARIMAX, a pesar de su complejidad, no muestra un mejor desempeño que otros modelos más simples, esto puede deberse al punto extremadamente sobreestimado del modelo en el periodo 13.

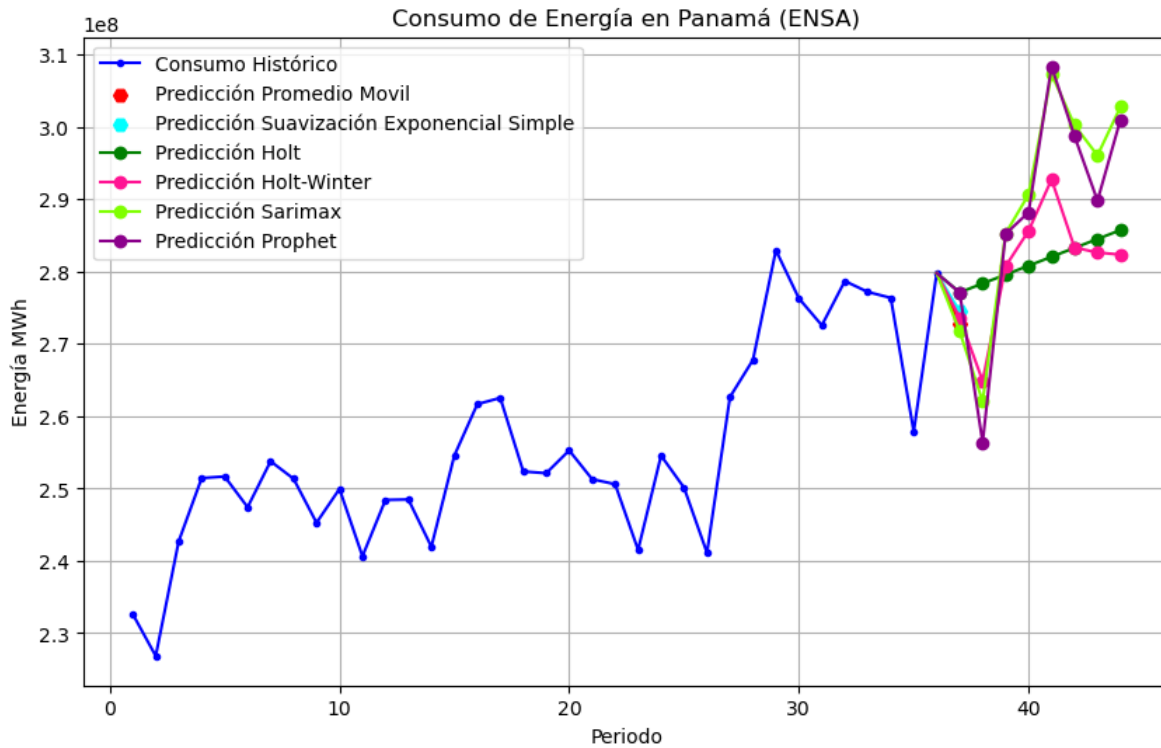


Fig. 16 Predicción de 8 periodos futuros para todos los modelos.

EDEMET:

Observaciones: el código utilizado será el mismo que el de ENSA, por lo que solo se mostrará las graficas y el resumen final.

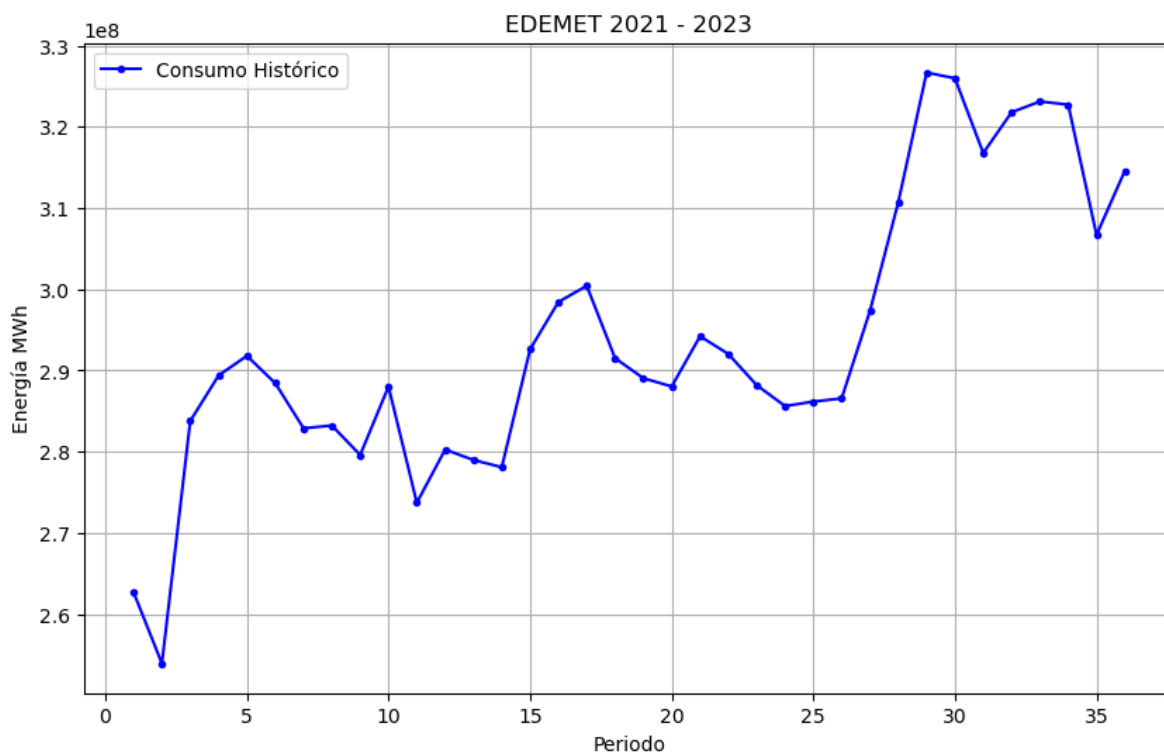


Fig. 17 Consumo eléctrico de Clientes de EDEMET 2021- 2023

En la Fig. 17 se muestra el consumo eléctrico en MWh a lo largo de 36 periodos de clientes de EDEMET, al igual que en la Fig. 1 se puede observar cierta tendencia.

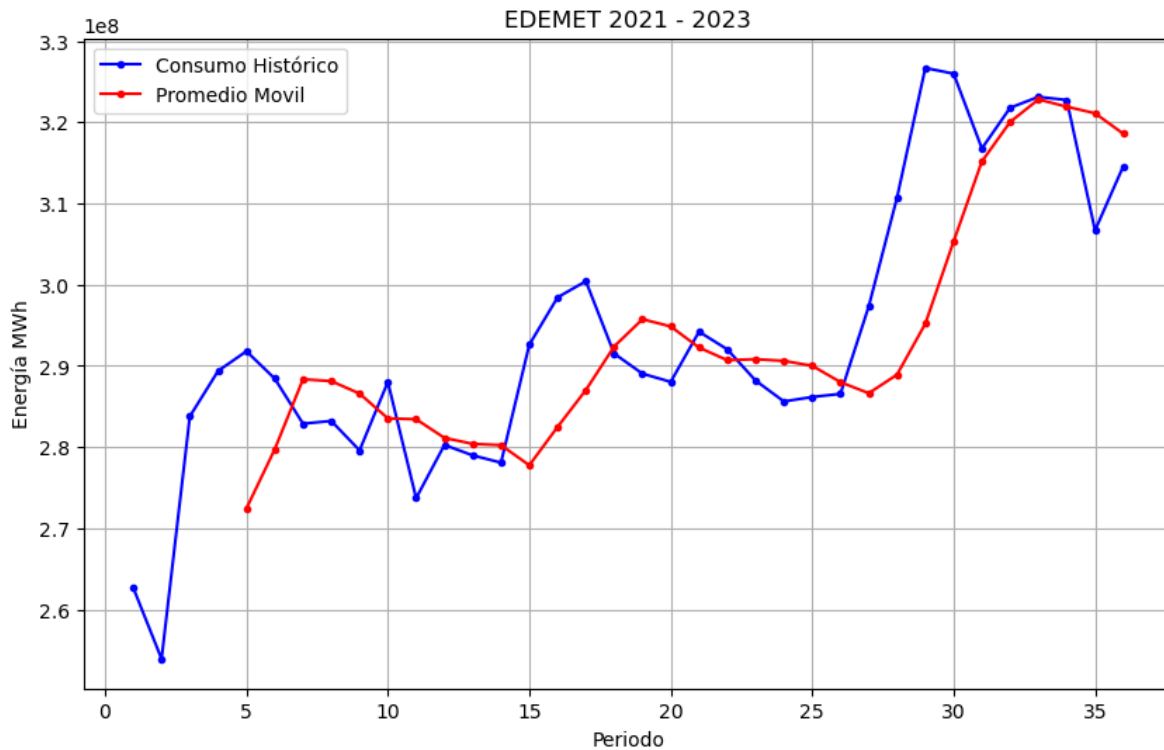


Fig. 18 Gráfica comparativa de promedio móvil y los datos reales de EDEMET

En la Fig. 18 la línea roja del promedio móvil confirma la tendencia al alza y ayuda a identificar los puntos de inflexión o picos debido a factores desconocidos.

También se puede ver cierto desfase al igual que ENSA, además del 27 al 30 subestimo todos los valores en relación con los datos reales.

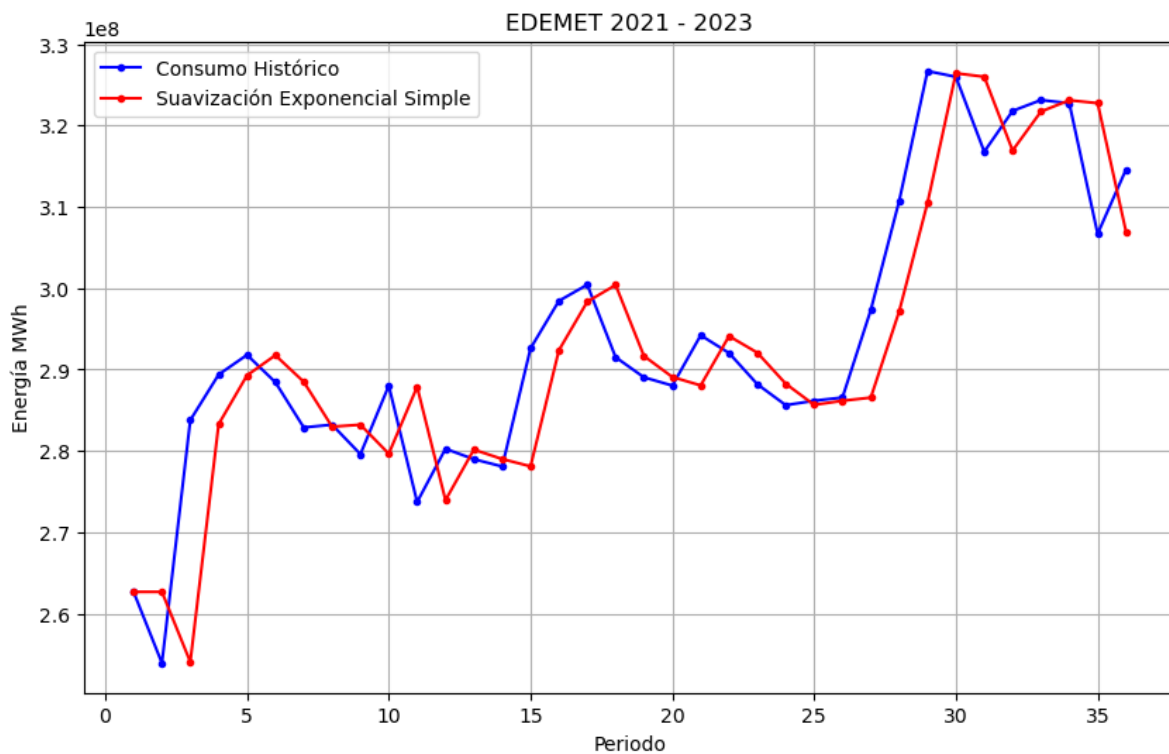


Fig. 19 Gráfica comparativa de suavización exponencial simple y los datos reales

La teoría nos indica que la suavización exponencial, al asignar mayor peso a los datos más recientes, puede reaccionar más rápidamente a cambios en la tendencia en comparación con el promedio móvil y se puede observar esto en la Fig. 19 sin embargo el modelo no predice adecuadamente los datos reales.

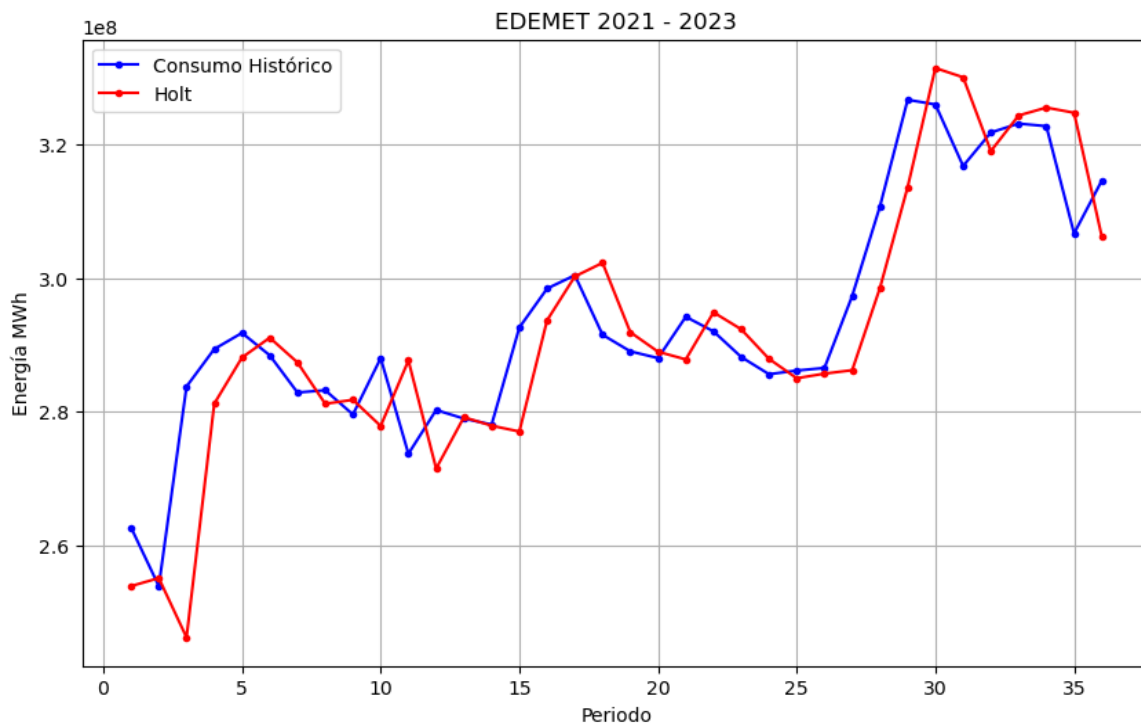


Fig. 20 Gráfica datos reales y valores predichos por el modelo de Holt

El modelo de Holt (Fig. 20) tiene un comportamiento similar al de suavizamiento exponencial simple, aunque hay cierta proximidad entre los datos reales y los predichos por el modelo de holt, no se ajusta adecuadamente.

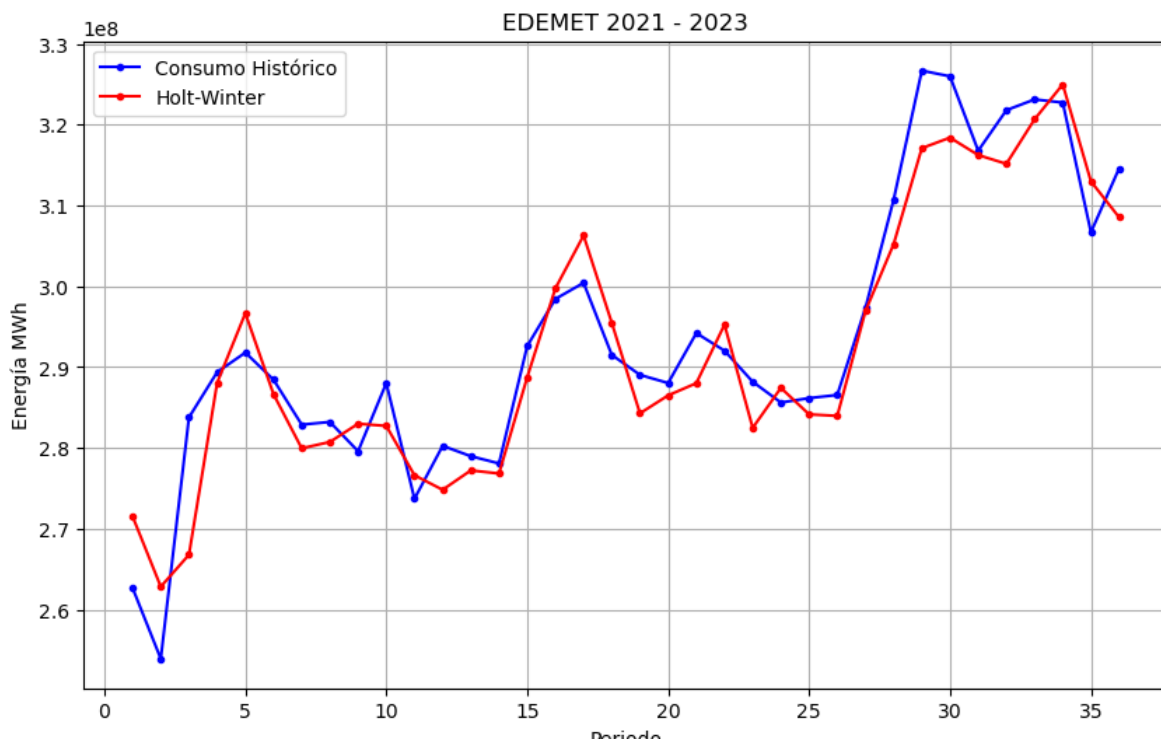


Fig. 21 Gráfica datos reales y valores predichos por el modelo de Holt-Winters

En la Fig. 21 el modelo Holt-Winters se ajusta razonablemente bien a los datos históricos. Esto indica que el modelo es capaz de capturar tanto la tendencia como la estacionalidad del consumo de energía. Sin embargo, existen algunos puntos donde el modelo se desvía ligeramente de los datos reales

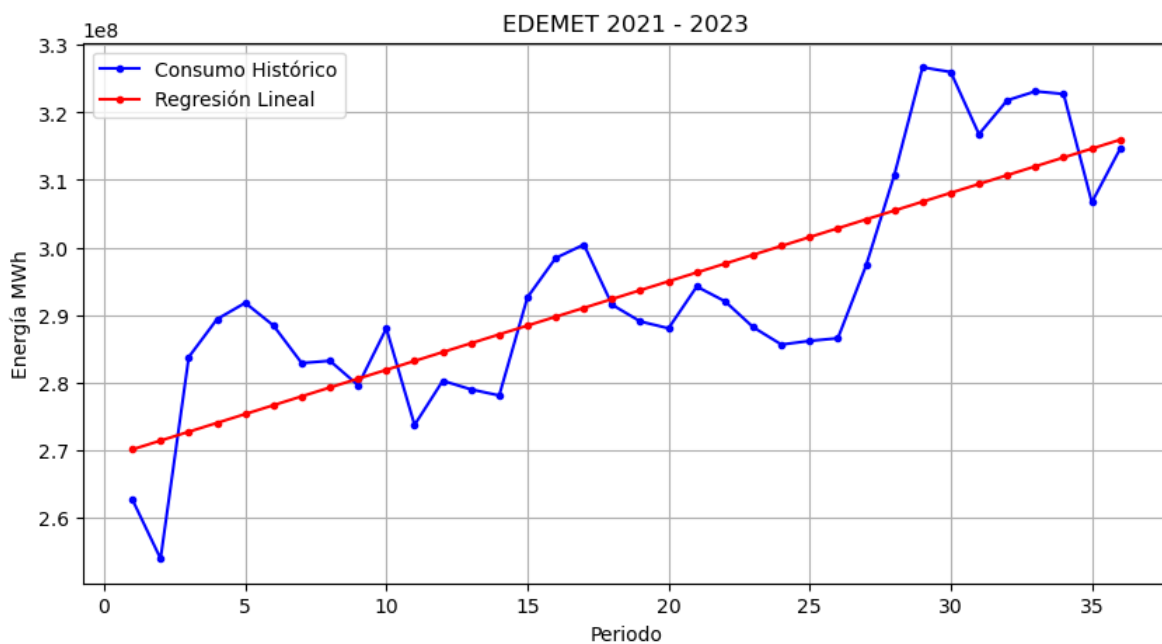


Fig. 22 Gráfica comparativa de la regresión lineal y los datos reales

En la Fig. 22 los puntos de datos históricos se desvían de la línea de regresión lineal, lo que confirma la presencia de factores adicionales que influyen en el consumo más allá de la tendencia general.

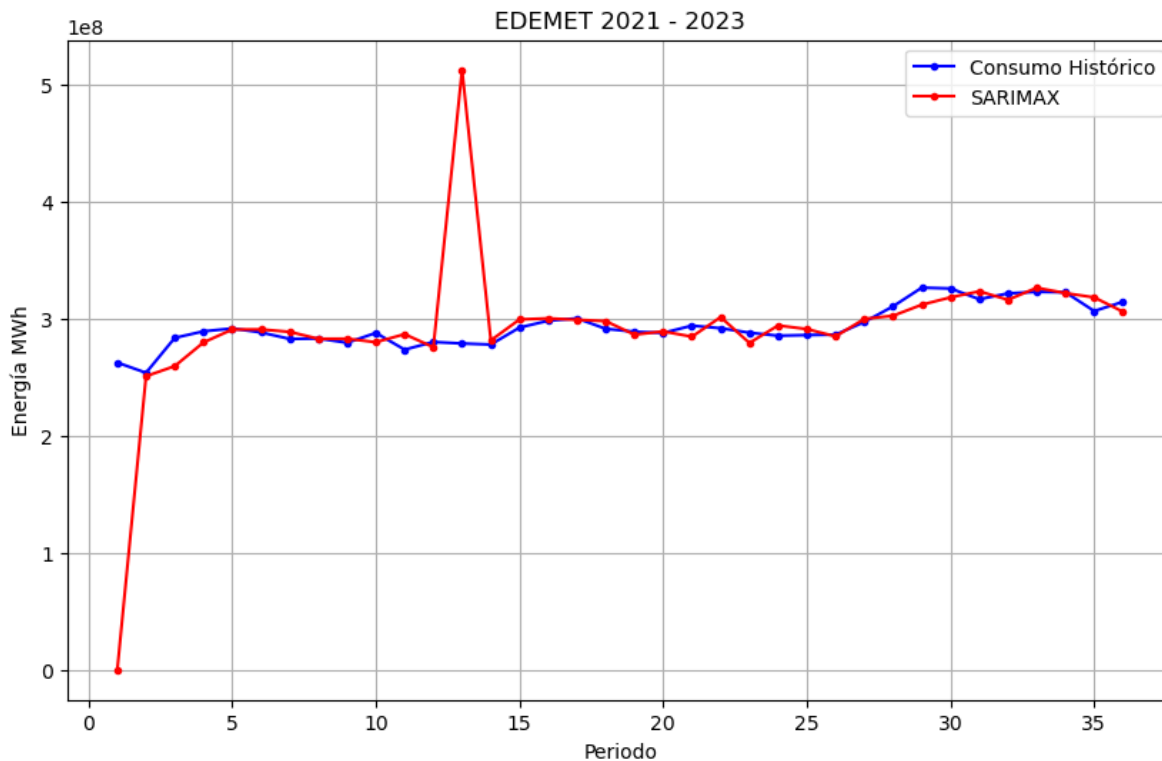


Fig. 23 Gráfica datos reales y valores predichos por el modelo de Sarima

La predicción del Modelo Sarima (Fig. 23) sigue de cerca la tendencia del consumo histórico, lo que indica que el modelo está capturando bien los patrones de consumo, pero a pesar del buen ajuste, existen algunas discrepancias entre la predicción y los datos históricos, especialmente en el pico sobre estimado para el periodo 13, esto hace contraste con ENSA que el modelo tuvo exactamente el mismo comportamiento.

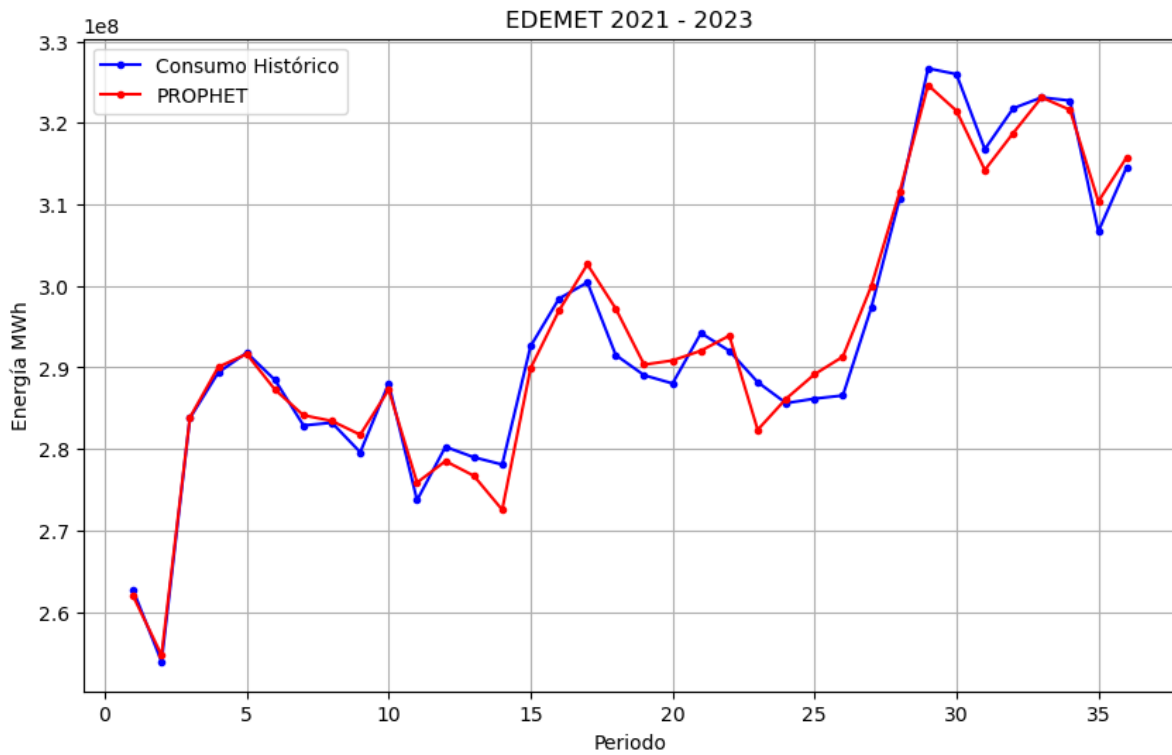


Fig. 24 Gráfica datos reales y valores predichos por el modelo de Prophet

El modelo de prophet para los datos de EDEMET (Fig. 24) también parece capturar de manera razonable la estacionalidad y la tendencia presente en los datos, es decir, los patrones recurrentes de consumo a lo largo del tiempo. Esto se observa en cómo la línea roja sigue los picos y valles de la línea azul.

Evaluaciones de métricas de Error (EDEMET)

	Modelo	MAD	MAPE	STD
0	Promedio Movil	7.78449e+06	2.56312	9.63861e+06
1	Suavización Exponencial	6.17469e+06	2.10066	7.71837e+06
2	Holt	6.80779e+06	2.31337	8.50974e+06
3	Holt-Winter	4.39993e+06	1.50551	5.49991e+06
4	Sarimax	1.96753e+07	7.10035	2.45941e+07
5	Regresión	8.96669e+06	3.05514	1.12884e+07
6	Prophet	2.10175e+06	0.715288	2.62719e+06

Prophet, muestra un excelente desempeño, se ve especialmente en el MAPE, lo que indica una alta precisión en términos porcentuales, sin embargo, en comparación con ENSA el valor de MAD creció bastante.

Sarima presenta el mejor MAD pero el peor MAPE, esto podría decir que en términos absolutos es una buena opción.

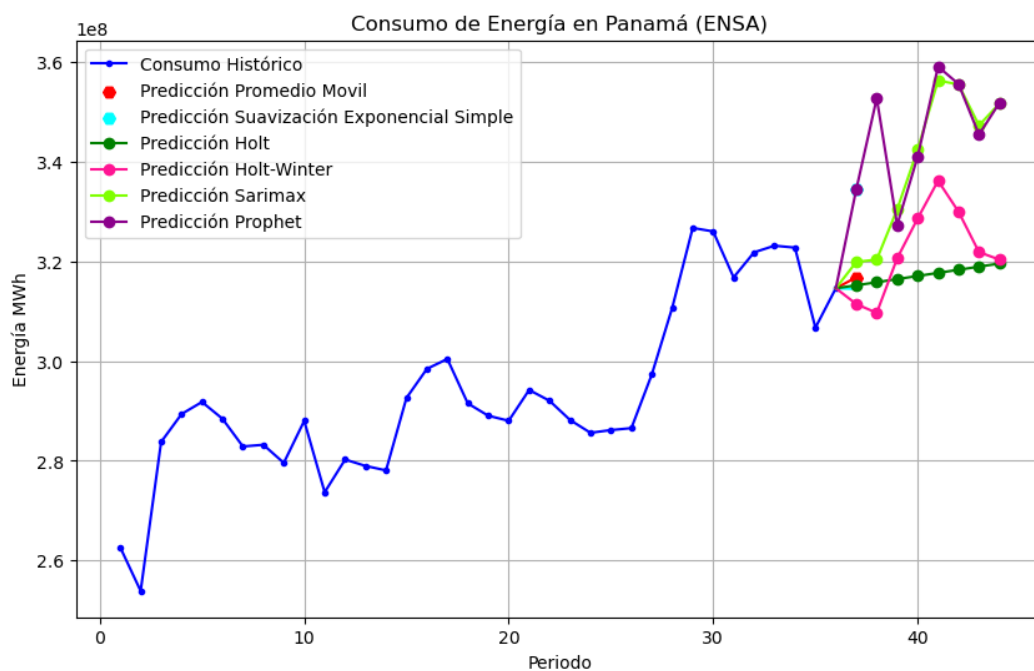


Fig. 25 Predicción de 8 periodos futuros para todos los modelos.

EDECHI:

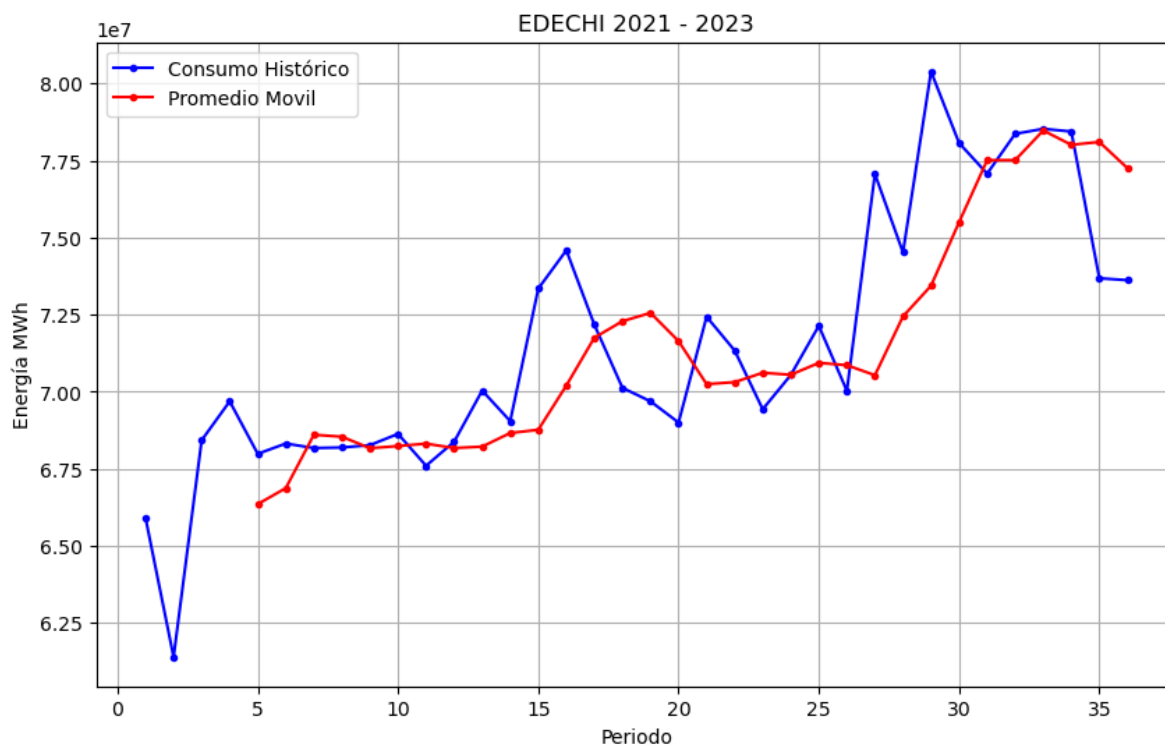


Fig. 26 Gráfica comparativa de promedio móvil y los datos reales de EDECHI

Para EDECHI se tiene el mismo comportamiento del promedio móvil de las otras dos distribuidoras de energía eléctrica en Panamá.

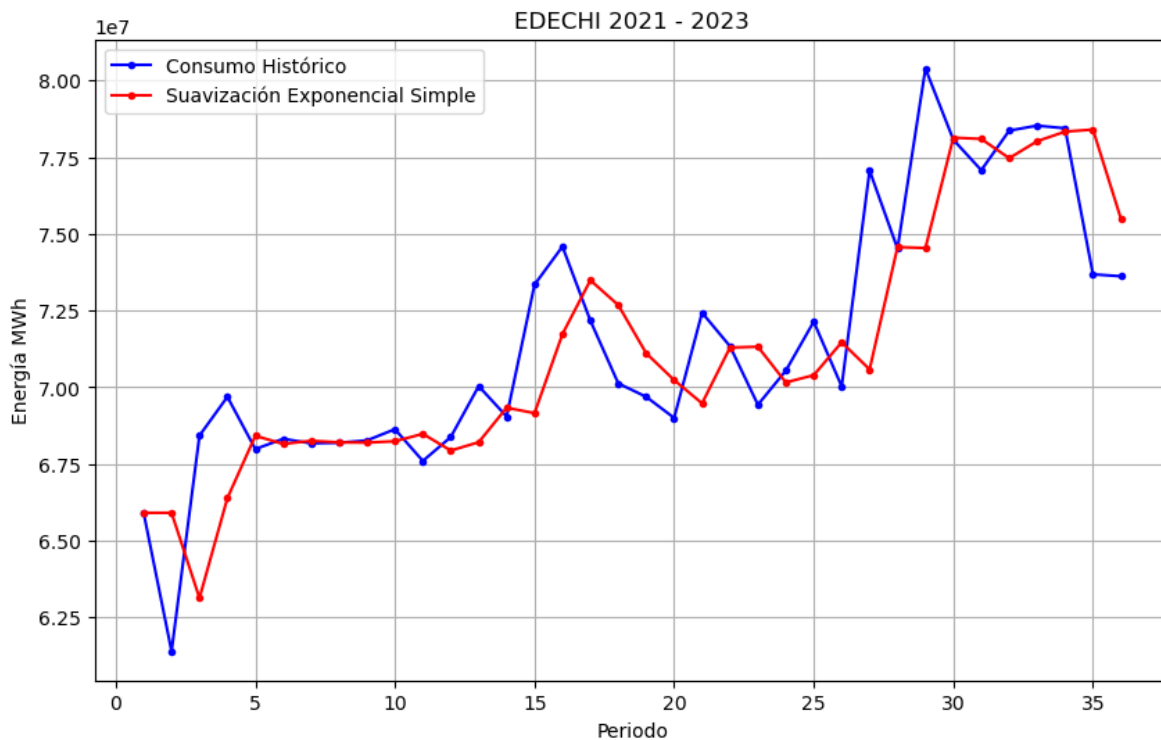


Fig. 27 Gráfica comparativa de suavización exponencial simple y los datos reales

La predicción del modelo de suavización exponencial parece estar bastante cerca de los datos reales, comportamiento similar a las otras distribuidoras, y también mantiene el desfase en los periodos que indica no se ajusta correctamente a los valores reales.

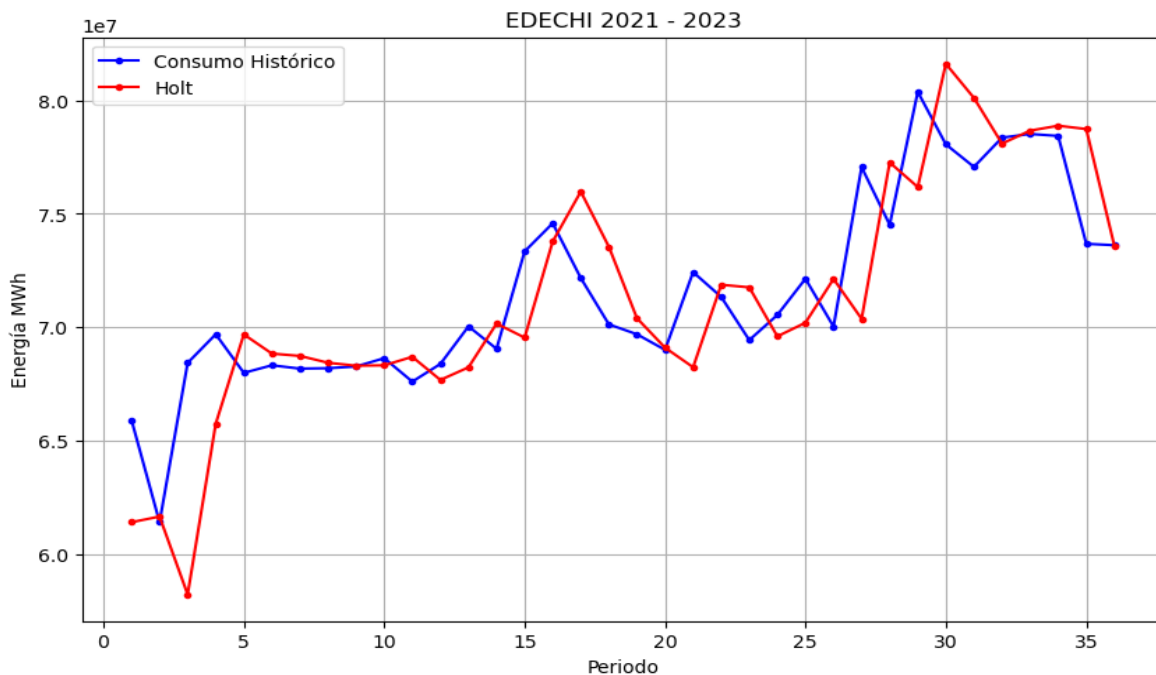


Fig. 28 Gráfica datos reales y valores predichos por el modelo de Holt

El modelo Holt para los datos de EDECHI (Fig. 28) se ajusta de manera relativamente buena a los puntos de datos del consumo histórico, especialmente en los primeros periodos. Sin embargo, hacia el final del período, el modelo parece sobreestimar ligeramente el consumo en algunos puntos.

El modelo Holt para EDECHI no captura con gran precisión los picos y valles del consumo histórico.

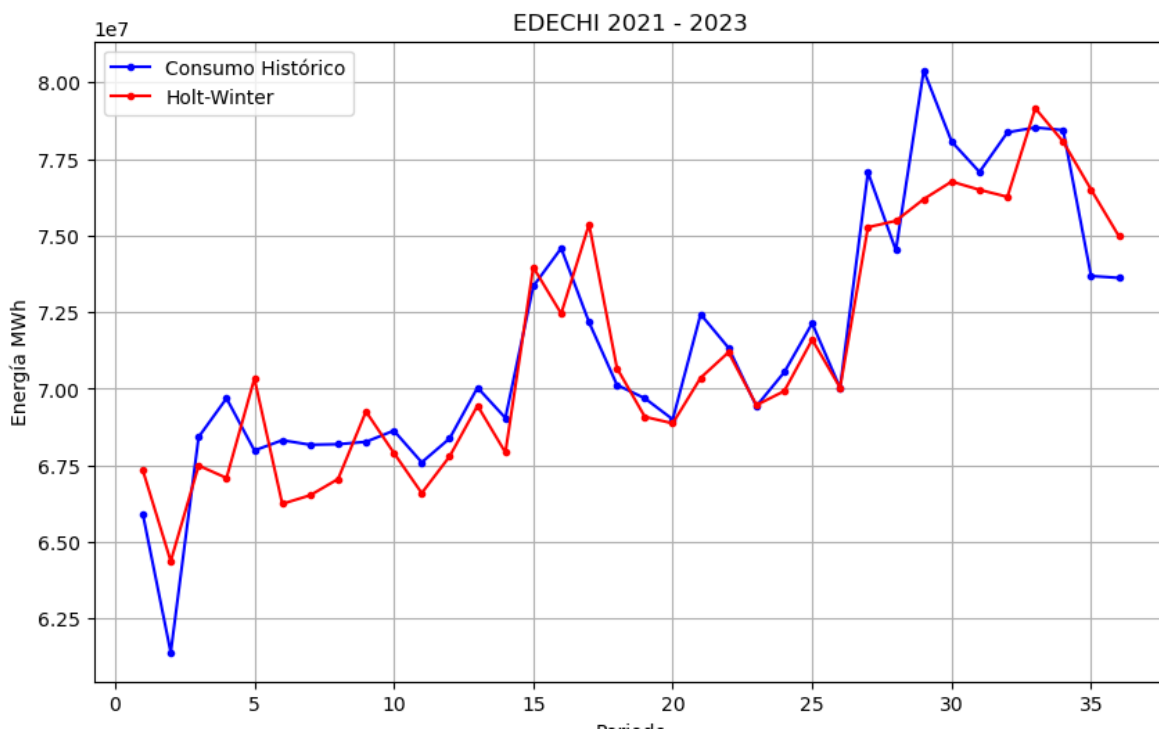


Fig. 29 Gráfica datos reales y valores predichos por el modelo de Holt-Winters

Modelo Holt-Winters para EDECHI Fig. 29, parece seguir de cerca la tendencia general y las fluctuaciones estacionales del consumo histórico, lo que indica que el modelo captura de buena manera tanto la dirección general como los patrones estacionales del consumo.

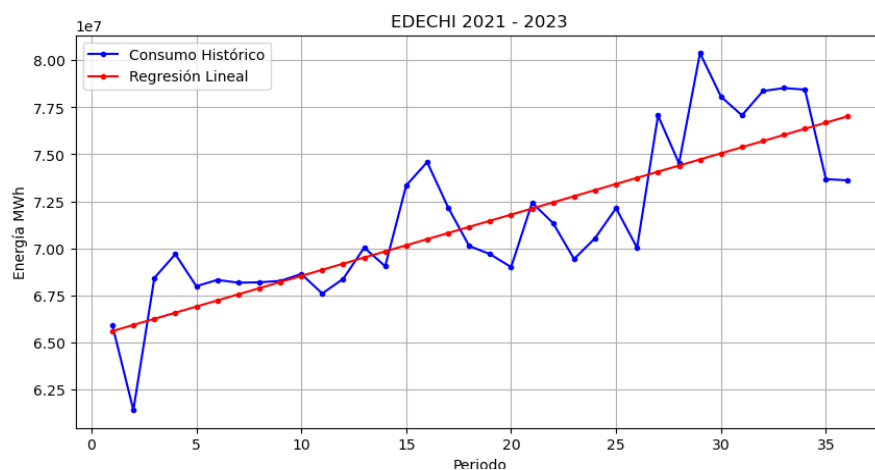


Fig. 30 Gráfica comparativa de la regresión lineal y los datos reales

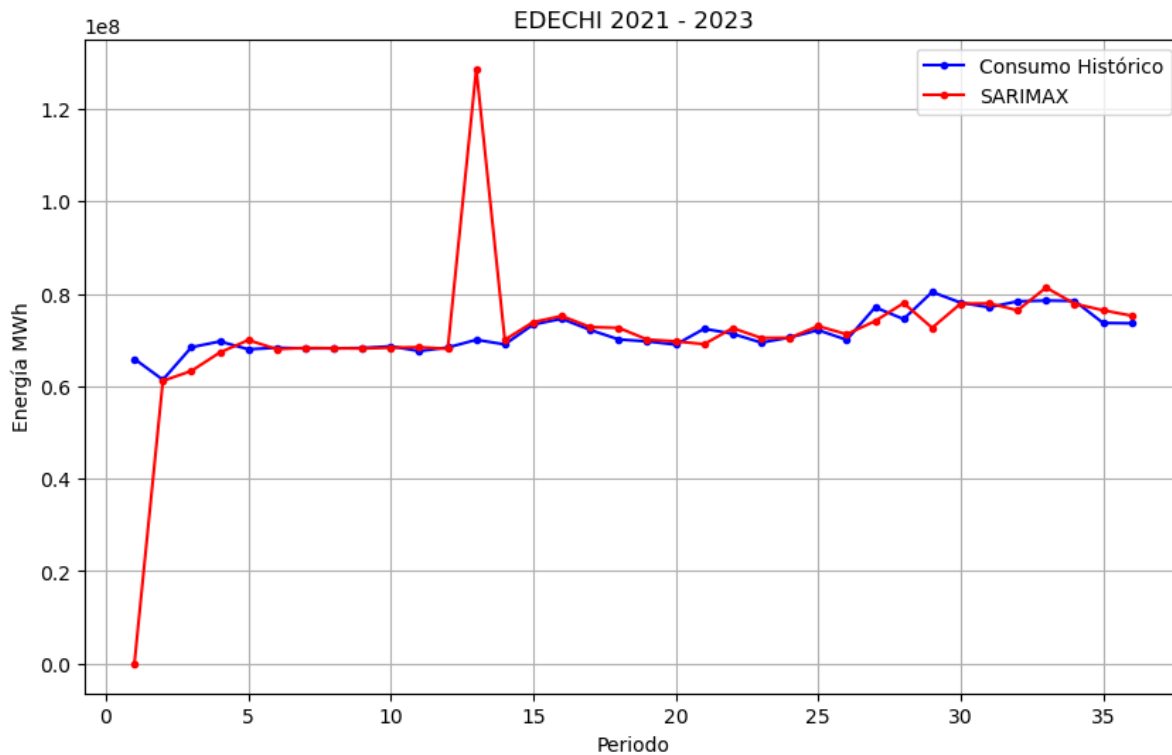


Fig. 31 Gráfica datos reales y valores predichos por el modelo de Sarima

En EDECHI también se observa el valor predicho para el periodo 13 super alto en relación con el valor real de consumo, esto plantea interrogantes en cuanto al ajuste del modelo a nivel de Python ya que puede estar sobre ajustado al ruido aleatorio.

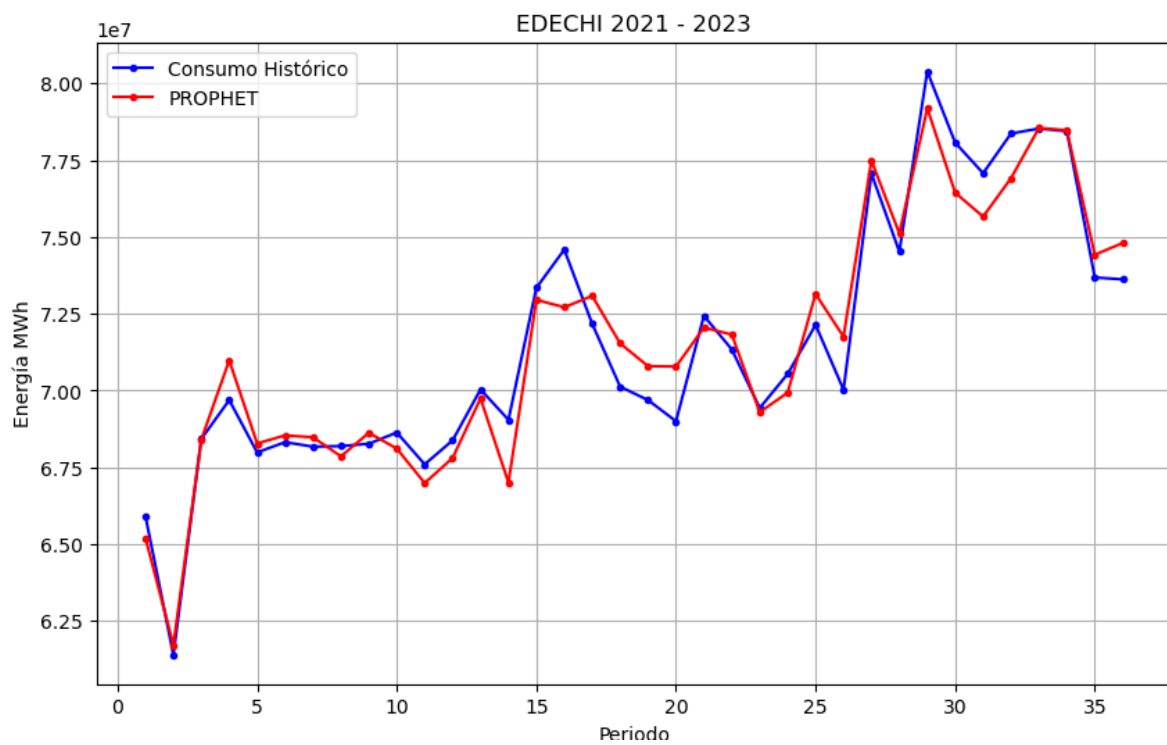


Fig. 32 Gráfica datos reales y valores predichos por el modelo de Prophet

El modelo de prophet para EDECHI (Fig. 32) captura con bastante certeza la tendencia y estacionalidad de los datos reales.

Evaluaciones de métricas de error (EDECHI)

	Modelo	MAD	MAPE	STD
0	Promedio Movil	1.84123e+06	2.50165	2.30154e+06
1	Suavización Exponencial	1.7026e+06	2.36775	2.12825e+06
2	Holt	2.16381e+06	3.00692	2.70476e+06
3	Holt-Winter	1.30376e+06	1.82283	1.6297e+06
4	Sarimax	4.86986e+06	7.02628	6.08732e+06
5	Regresión	1.95629e+06	2.70943	2.44536e+06
6	Prophet	788707	1.0965	985883

El modelo de prophet presenta los valores más bajos en MAD y MAPE, lo que indica que, en promedio, sus predicciones son las más cercanas a los valores reales. Además, su desviación estándar es mucho más baja, sugiriendo una mayor consistencia en sus errores.

Sarimax presenta el peor desempeño, con los valores más altos en todas las métricas, este modelo parece ser el menos preciso en sus predicciones.

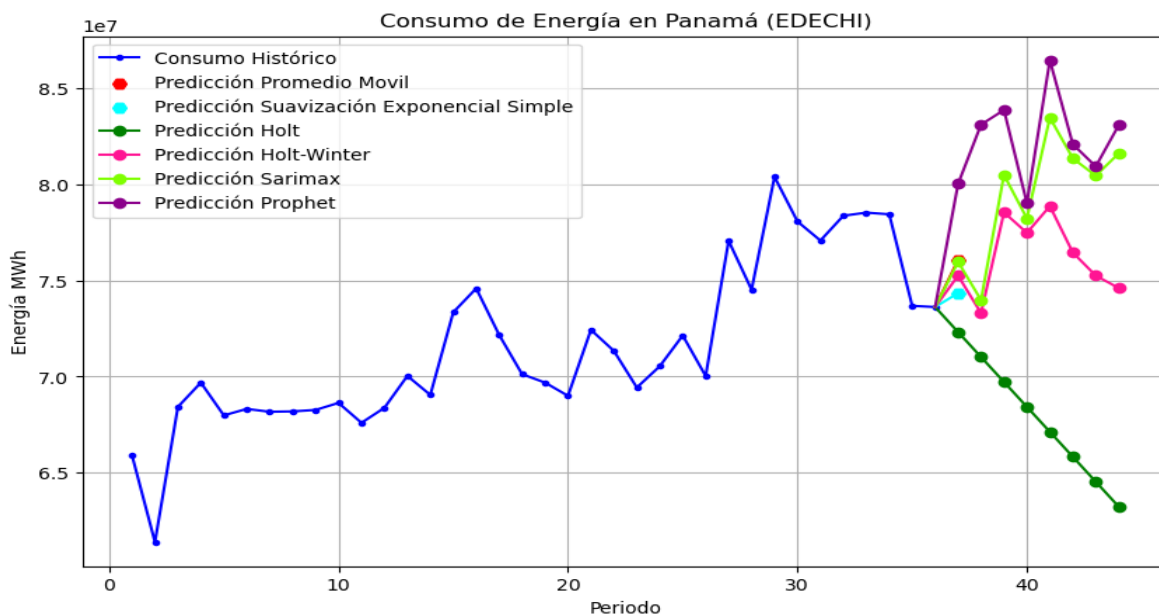


Fig. 33 Predicción de 8 periodos futuros para todos los modelos.



CONCLUSIONES Y RECOMENDACIONES

Basado en los resultados obtenidos, se recomienda utilizar el modelo Prophet como el principal modelo de predicción para el consumo eléctrico en Panamá para el año 2024. Este modelo ha demostrado ser el más adecuado, ofreciendo una alta precisión y un bajo error de predicción. Sin embargo, es fundamental realizar un análisis más profundo y considerar las particularidades de los datos para tomar una decisión final.

Aunque Prophet ha demostrado ser muy bueno, es recomendable comparar su desempeño con otros modelos como redes neuronales recurrentes (RNN) para obtener una perspectiva más amplia y asegurar que se ha seleccionado la mejor opción.

Se debe tener presente que los datos obtenidos de datos abiertos Panamá, no incluían otras variables que pueden ser influyentes en el consumo eléctrico, como la temperatura media para esos meses.

Es crucial monitorear regularmente el desempeño del modelo para asegurarse de que sigue siendo preciso. Esto incluye la revisión de métricas de error y la comparación con datos reales, también es una buena práctica reentrenar el modelo con datos nuevos y actualizados para mantener su precisión y adaptabilidad a cambios en los patrones de consumo.