



# Modelo SARIMAX para la Predicción de Generación Solar en la Planta Sol del Desierto en Chile: Evaluación de Variables Meteorológicas

Felipe S. Neira Rojas   Angel T. Llanos Herrera   Diego I. Rocha Retamal

Universidad Católica del Maule



## Presentación y Discusión de Literatura

Chile ha consolidado su liderazgo en energía solar gracias a la alta irradiación del Desierto de Atacama, con una creciente necesidad de modelar la variabilidad horaria de generación en plantas específicas como Sol del Desierto. Si bien existen estudios sobre capacidad instalada, persiste una brecha en la predicción operativa a corto plazo usando datos reales y locales. Esta investigación responde a esa carencia, analizando cómo las condiciones meteorológicas influyen en la producción diaria de energía, para así mejorar la planificación operativa mediante modelos estadísticos robustos y de bajo costo computacional.

### Elección de Modelos a Utilizar

El uso de modelos VAR se fundamenta en su capacidad para capturar la dinámica conjunta entre múltiples variables meteorológicas interrelacionadas —como temperatura y presión— tal como lo recomienda Peña (2005), permitiendo una estimación coherente de sus trayectorias temporales. Por otro lado, SARIMAX permite incorporar estos pronósticos como variables exógenas en la predicción de la generación solar, integrando tanto componentes estacionales como efectos externos, lo cual ha demostrado reducir significativamente el error de predicción en estudios previos (Benítez et al., 2023; Herrera et al., 2023). Este enfoque híbrido busca equilibrar precisión y cumplimiento estadístico de supuestos, conforme al objetivo general del estudio.

En este caso, dado que las variables meteorológicas no presentan una relación directa significativa con la generación de energía solar en la planta fotovoltaica, se utilizarán como variables exógenas en el modelo de predicción. En consecuencia, será necesario estimar sus valores futuros mediante un modelo VAR, considerando que dichas variables muestran interdependencia entre sí.

Variable a predecir	VAR SARIMAX	
Meteorológicas	✓	✗
Generación de energía	✗	✓

Table 1. Modelo a utilizar según variable a predecir

## Supuestos de los Modelos

- Estacionariedad:** La serie de tiempo analizada en VAR debe ser estacionaria, con media y varianza constante.
- Variables Endógenas:** No debe haber multicolinealidad perfecta entre ellas, pero teniendo relación entre sí.
- Ruido Blanco:** Los errores deben seguir una distribución de RB con media 0 con varianza y covarianza constante, siendo independientes entre ellos.

Extensión para SARIMAX:

- Variables Exógenas:** No deben tener relación entre sí.

## Generación de Energía de la Planta SV Sol del Desierto

La generación diaria de energía en la planta fotovoltaica Sol del Desierto presenta una alta volatilidad, con valores que oscilan entre 0 y 2.375 MW. Se identificaron 17 días con registros atípicos, incluyendo jornadas con generación nula, posiblemente asociadas a mantenciones o fallas. De forma habitual, la producción se concentra entre 244,7 MW y 2.179,5 MW por día, evidenciando una variabilidad significativa que refuerza la necesidad de modelos predictivos robustos.

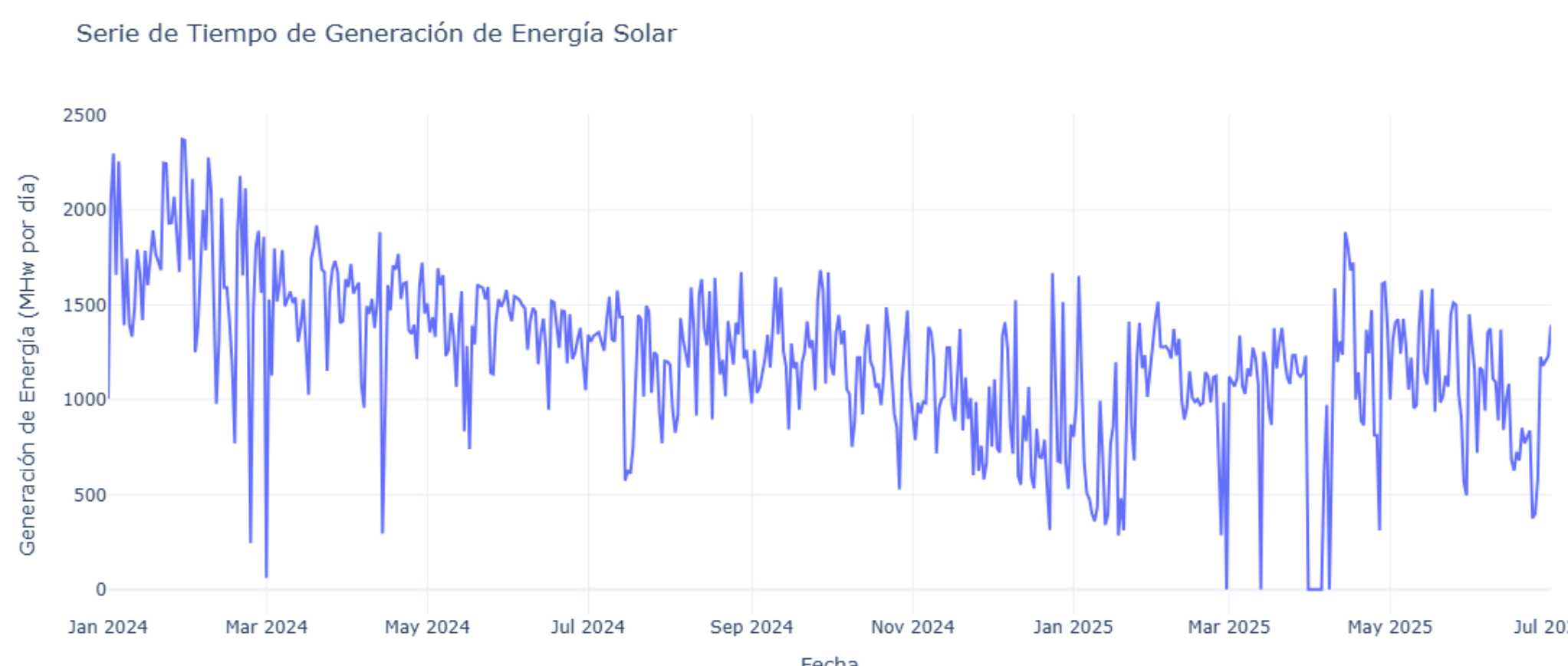


Figure 1. Valores Reales de Generación de Energía de la Planta Sol del Desierto

### Modelo VAR(7) para la Estimación de Variables Meteorológicas

$$y_t = c + A_1y_{t-1} + A_2y_{t-2} + A_3y_{t-3} + A_4y_{t-4} + A_5y_{t-5} + A_6y_{t-6} + A_7y_{t-7} + \epsilon_t \quad (1)$$

Donde:

- $c$ : Vector de constantes (interceptos) de dimensión  $4 \times 1$ .
- $y_t$ : Vector de variables endógenas en el tiempo  $t$ , de dimensión  $4 \times 1$ .
- $A_i$ : Matrices de coeficientes autorregresivos de dimensión  $4 \times 4$  para cada rezago  $i = 1, \dots, 7$ .
- $\epsilon_t$ : Vector de errores aleatorios con media cero y covarianza constante, de dimensión  $4 \times 1$ .

## Predicción de Variables Meteorológicas

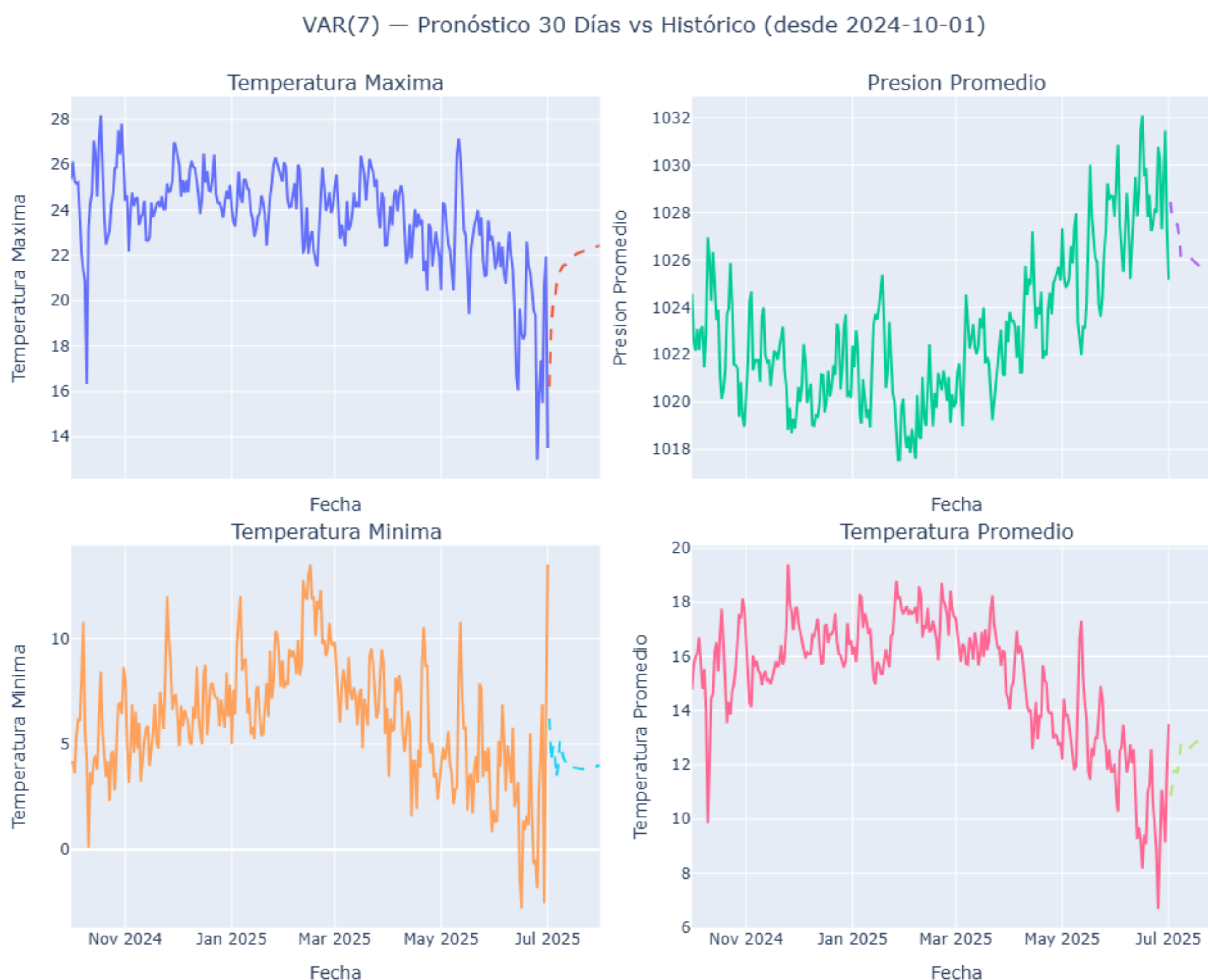


Figure 2. Predicciones de 30 días para Presión a Nivel del Mar y Temperaturas

## Cumplimiento de Supuestos y Métricas

Para evaluar la validez estadística del modelo VAR(7), se consideraron los supuestos clásicos de normalidad, homocedasticidad y ausencia de autocorrelación. Los resultados indican que dichos supuestos se cumplen para las variables temperatura promedio, temperatura mínima y presión a nivel del mar. Sin embargo, los residuos asociados a la temperatura máxima presentaron una leve desviación respecto a la distribución normal.

Las métricas de desempeño del modelo predictivo se resumen en la Tabla 2

Variable	RMSE	MAE
Temperatura Máxima	3.540	2.484
Presión Promedio	2.421	1.982
Temperatura Mínima	3.424	2.673
Temperatura Promedio	2.156	1.815

Table 2. Métricas de error para el modelo VAR(7) en el pronóstico de 30 días (RMSE y MAE).

## Resultados Modelo SARIMAX

Se ajustaron modelos SARIMAX para predecir la generación diaria de energía solar, incorporando como variables exógenas la temperatura máxima y la presión atmosférica promedio. Si bien el modelo que utiliza la presión promedio obtuvo menores errores de predicción, no cumplió con los supuestos estadísticos fundamentales de normalidad, homocedasticidad y ausencia de autocorrelación. En cambio, el modelo con temperatura máxima presentó un rendimiento ligeramente inferior, pero mostró mayor consistencia desde el punto de vista estadístico. Las métricas resultantes se encuentran en la Tabla 3.

Variable Exógena	RMSE	MAE
Temperatura Máxima	3.540	2.484
Presión Promedio	2.421	1.980

Table 3. Métricas de error del modelo SARIMAX en la predicción de generación solar

## Conclusiones y Limitaciones

Los resultados evidencian que el modelo VAR(7) es adecuado para predecir variables meteorológicas interrelacionadas, destacando un mejor desempeño en series menos volátiles como la temperatura promedio. En cambio, los modelos SARIMAX mostraron un mayor poder predictivo al incorporar variables exógenas, aunque con violaciones a supuestos clave, especialmente en el caso de la presión promedio.

- Se trabajó con datos diarios en lugar de horarios, lo que limitó la capacidad para capturar la variabilidad intradiaria. Esto debido a costos computacionales.
- No se incluyó la radiación solar debido a la falta de datos completos y actualizados en la zona de estudio.
- El modelo SARIMAX presentó violaciones a supuestos estadísticos (normalidad, homocedasticidad y autocorrelación), lo que afecta la validez inferencial.