

Desarrollo de un Modelo de Simulación Fotovoltaica Basado en Clusters y Regresión Lineal Simple para Dimensionamiento Web

Antar, Mauro; Barbagallo, Daniel; Rau Bekerman, Matias.

bekermanmatias@gmail.com

RESUMEN

Este trabajo presenta una metodología optimizada para la estimación de potencia en sistemas fotovoltaicos, implementada en una herramienta de dimensionamiento online. Ante la complejidad computacional de los modelos físicos tradicionales, se propone un enfoque por clusters, segmentando el espacio de operación en 6 escenarios discretos definidos por la temperatura ambiente ($10^{\circ}C, 20^{\circ}C, 30^{\circ}C$) y la inclinación del panel ($20^{\circ}, 45^{\circ}$). Para cada cluster, se deriva una función de transferencia lineal univariable dependiente exclusivamente de la irradiancia, obtenida mediante el método de Mínimos Cuadrados Ordinarios (OLS) sobre un dataset experimental. El algoritmo resultante ejecuta un cálculo iterativo de resolución mensual para determinar el número de paneles, la energía anual estimada y los costos de inversión (incluyendo el Balance del Sistema o BOS) con alta precisión y transparencia.

Palabras clave: Regresión Lineal Simple, Clusters, Simulación.

1. INTRODUCCIÓN

La simulación de sistemas energéticos en entornos web requiere un equilibrio entre precisión física y eficiencia computacional. En aplicaciones comerciales de dimensionamiento rápido, resulta eficiente pre-calcular el comportamiento del sistema en puntos de operación estandarizados ("Clusters") en lugar de resolver ecuaciones complejas en tiempo real.

El objetivo de este estudio es desarrollar 6 funciones matemáticas que describan la Generación Eléctrica (P , en Watts) en función de la Irradiancia Solar (G , en W/m^2). A diferencia de otros modelos anualizados, este enfoque utiliza la temperatura y la inclinación como parámetros de selección dinámica, permitiendo al usuario final obtener una proyección técnica y económica fiable basada en **datos** climáticos promedio mensuales, capturando así la estacionalidad del recurso.

2. Marco Teórico: Regresión Lineal Simple por Mínimos Cuadrados

Para cada uno de los 6 clusters identificados, modelamos la relación entre la variable independiente (Irradiancia) y la dependiente (Potencia) mediante una recta de regresión simple. Físicamente, esto asume que la eficiencia de conversión es constante dentro de un rango de temperaturas acotado, tal como se expresa en la Ecuación (1).

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (1)$$

Donde:

- y_i (Potencia): Variable dependiente observada en el dataset.
- x_i (Irradiancia): Variable independiente (entrada climática).
- β_1 (Pendiente): Representa la eficiencia de conversión del panel en ese escenario específico.
- β_0 (Intercepto): Representa las pérdidas basales del sistema (consumo en reposo).
- ε_i (Residuo): El error aleatorio no explicado por el modelo.

2.1 Definición de la Función de Costo

El método de Mínimos Cuadrados (OLS) busca encontrar la línea que minimiza la distancia entre los datos reales y la predicción. Para ello, se minimiza la función de costo S (Suma de Errores Cuadráticos).

Se elevan los residuos al cuadrado por dos razones fundamentales: primero, para penalizar fuertemente los errores grandes; y segundo, para evitar que los errores positivos y negativos se cancelen entre sí al sumarlos.

$$S(\beta_0, \beta_1) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - (\beta_0 + \beta_1 x_i))^2 \quad (2)$$

2.2 Minimización por Derivadas Parciales

Para hallar los estimadores óptimos ($\hat{\beta}_0, \hat{\beta}_1$) que hacen mínima a la función S , aplicamos cálculo diferencial: derivamos la función respecto a cada coeficiente e igualamos a cero (condición de punto crítico).

Paso 1: Derivada respecto al intercepto (β_0)

$$\frac{\partial S}{\partial \beta_0} = \frac{\partial}{\partial \beta_0} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) = 0$$

Simplificando, obtenemos la Primera Ecuación Normal:

$$\sum y_i = n\beta_0 + \beta_1 \sum x_i$$

Paso 2: Derivada respecto a la pendiente (β_1)

$$\frac{\partial S}{\partial \beta_1} = \frac{\partial}{\partial \beta_1} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) x_i = 0$$

Simplificando, obtenemos la Segunda Ecuación Normal:

$$\sum y_i x_i = \beta_0 \sum x_i + \beta_1 \sum x_i^2$$

El simulador web utiliza las fórmulas cerradas resultantes de este sistema de 2x2 para determinar los coeficientes β de cada cluster, garantizando que el modelo matemático es estadísticamente óptimo para los datos provistos.

2.3 Resolución del Sistema de Ecuaciones Normales

Al reunir las dos condiciones obtenidas de la minimización (Pasos 1 y 2), se conforma el sistema lineal de dos incógnitas (β_0 y β_1) presentado en las Ecuaciones (3) y (4).

$$\{n \cdot \beta_0 + (\sum x_i) \cdot \beta_1 = \sum y_i \quad (3)$$

$$(\sum x_i) \cdot \beta_0 + (\sum x_i^2) \cdot \beta_1 = \sum x_i y_i \quad (4)$$

Este sistema puede expresarse matricialmente de la forma $A\vec{\beta} = \vec{b}$. Para obtener las fórmulas operativas que utiliza el simulador, resolvemos este sistema, resultando en las siguientes expresiones cerradas para la pendiente y el intercepto:

- **Pendiente (Eficiencia):** $\beta_1 = \frac{n\sum x_i y_i - (\sum x_i)(\sum y_i)}{n\sum x_i^2 - (\sum x_i)^2}$
- **Intercepto:** $\beta_0 = \frac{\sum y_i - \beta_1 \sum x_i}{n} = \bar{y} - \beta_1 \bar{x}$

Estas son las ecuaciones exactas que se procesan “offline” con el dataset para obtener los coeficientes presentados posteriormente en la Tabla 2.

3. Metodología

3.1 Definición de Clusters Operativos

Se clasificó el dataset experimental 01-GeneracionFV.csv en 6 subconjuntos de datos mutuamente excluyentes. Esta discretización permite “linealizar” un problema que es inherentemente no lineal (ya que la temperatura afecta la eficiencia).

1. **Cluster 1:** $T \approx 10^\circ C, \theta = 20^\circ$
2. **Cluster 2:** $T \approx 20^\circ C, \theta = 20^\circ$
3. **Cluster 3:** $T \approx 30^\circ C, \theta = 20^\circ$
4. **Cluster 4:** $T \approx 10^\circ C, \theta = 45^\circ$
5. **Cluster 5:** $T \approx 20^\circ C, \theta = 45^\circ$
6. **Cluster 6:** $T \approx 30^\circ C, \theta = 45^\circ$

3.2 Procesamiento de Datos

Se realizó una limpieza previa filtrando los registros con Irradiancia $G = 0$. Estos registros representan la noche, y su inclusión distorsionaría el cálculo de la eficiencia (β_1), ya que matemáticamente $0/0$ es indeterminado, pero físicamente sabemos que sin fotones incidentes no hay generación eléctrica.

4. Resultados y Validación

Tras la aplicación del algoritmo de mínimos cuadrados (detallado en la Sección 2) sobre los subconjuntos de datos definidos en la metodología, se procedió primero a seleccionar la arquitectura del modelo matemático más adecuada y posteriormente a calcular los coeficientes específicos para cada escenario operativo.

4.1 Análisis Comparativo y Selección del Modelo Matemático

Previo a la determinación de los coeficientes finales, se realizó un estudio numérico para validar la estructura interna de los datos. El objetivo fue determinar matemáticamente si la relación entre la Irradiancia y la Generación Eléctrica responde a un comportamiento lineal o si requiere modelos de orden superior.

Para este análisis se utilizó el conjunto de datos del Cluster 5 ($20^\circ C, 45^\circ$), evaluando tres familias de funciones mediante el método de Mínimos Cuadrados Ordinarios (OLS).

Definición de la Métrica de Validación (R^2)

Para comparar objetivamente los tres modelos, utilizamos el Coeficiente de Determinación, que mide la proporción de la varianza total de la variable dependiente que es explicada por el modelo, definida en la Ecuación (5).

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (5)$$

Donde:

- $SS_{tot} = \sum (P_i - \bar{P})^2$: Variabilidad total de los datos reales respecto a su promedio. Es un valor constante para el dataset (independiente del modelo).
- $SS_{res} = \sum (P_i - \hat{P}_i)^2$: Suma de los residuos al cuadrado. Mide el “error” acumulado entre el dato real (P_i) y la predicción del modelo ($\hat{P}_i$).

A. Modelo Lineal (Regresión por el Origen)

Se propone una relación proporcional estricta donde la variable dependiente (Generación) es función directa de la variable independiente (Irradiancia). Matemáticamente, buscamos la pendiente ideal que reduzca al mínimo la distancia entre la recta y los puntos.

Formulación Matemática

El modelo determinista se define en la Ecuación (6).

$$P_i = \beta_1 \cdot G_i \quad (6)$$

Donde:

- P_i : Valor de Potencia Generada (Variable dependiente).

- G_i : Valor de Irradiancia Global (Variable independiente).
- β_1 : Coeficiente de pendiente (el escalar desconocido que debemos encontrar).

Deducción del Coeficiente

Para encontrar el valor exacto de β_1 , definimos una Función de Costo $J(\beta_1)$, que representa la suma de las diferencias al cuadrado entre el dato real y el cálculo:

$$J(\beta_1) = \sum_{i=1}^n (P_{i,real} - \beta_1 \cdot G_i)^2$$

Nota: $\sum$ (Sigma) indica la sumatoria de todos los valores desde el dato 1 hasta el n .

Para hallar el mínimo de esta función (el punto óptimo), aplicamos Cálculo Diferencial. Calculamos la derivada de la función respecto a β_1 e igualamos a cero:

$$\frac{dJ}{d\beta_1} = -2 \sum_{i=1}^n G_i (P_{i,real} - \beta_1 G_i) = 0$$

Explicación: La derivada mide la tasa de cambio. Al igualarla a cero, encontramos matemáticamente el punto exacto donde la distancia entre la recta y los puntos es mínima.

Despejando β_1 de la ecuación anterior, obtenemos la fórmula analítica cerrada:

$$\beta_1 = \frac{\sum(P_{i,real} \cdot G_i)}{\sum(G_i^2)}$$

Interpretación:

- **Numerador** ($\sum P \cdot G$): Suma de productos cruzados (proyección de la potencia sobre la irradiancia).
- **Denominador** ($\sum G^2$): Suma de los cuadrados de la entrada (factor de normalización).

Resultado Numérico

Al procesar los datos del Cluster 5 con esta fórmula, se obtuvo:

$$P = 0.1744 \cdot G$$

Cálculo de Bondad de Ajuste (R^2)

Sustituimos la ecuación obtenida en la fórmula de residuos para cuantificar el error:

$$SS_{res(lin)} = \sum_{i=1}^n (P_{real} - (0.1744 \cdot G_i))^2$$

Al realizar la sumatoria, el error residual tiende a cero, indicando un ajuste casi perfecto:

$$R_{lin}^2 = 1 - \frac{\approx 0}{SS_{tot}} = 1.000$$

B. Modelo Polinómico de Segundo Orden (Cuadrático)

Se evalúa si la inclusión de un término cuadrático mejora la precisión del ajuste. Para resolver este sistema de múltiples incógnitas, recurrimos al Álgebra Matricial.

Formulación Matemática

Se plantea un polinomio de grado 2, definido en la Ecuación (7).

$$P = \beta_0 + \beta_1 \cdot G + \beta_2 \cdot G^2 \quad (7)$$

Donde:

- β_0 : Intercepto (ordenada al origen).
- β_1 : Coeficiente lineal.
- β_2 : Coeficiente cuadrático (representa la curvatura).

Resolución Matricial

Dado que tenemos un sistema sobredeterminado (más ecuaciones que incógnitas), expresamos el problema como $Y = X\beta$:

- Y : Vector columna con todas las lecturas de Potencia.
- X : Matriz de Diseño. Contiene los datos de entrada organizados en columnas según su potencia (G^0, G^1, G^2):

$$X = \begin{bmatrix} 1 & G_1 & G_1^2 \\ 1 & G_2 & G_2^2 \\ \vdots & \vdots & \vdots \\ 1 & G_n & G_n^2 \end{bmatrix}$$

El vector de coeficientes óptimos β se calcula mediante la Ecuación Normal:

$$\beta = (X^T X)^{-1} X^T Y$$

Explicación operativa:

- X^T (Transpuesta): Intercambia filas por columnas para alinear dimensiones.
- $(\dots)^{-1}$ (Inversa): Operación análoga a la división en álgebra matricial.

- $X^T Y$: Proyección de los datos observados sobre el espacio vectorial de las variables.

Resultado Numérico

Al resolver la matriz, se obtuvo el vector:

$$\beta = [0.001, 0.1743, 0.000004]^T$$

Esto resulta en la ecuación:

$$P = 0.001 + 0.1743 \cdot G + 0.000004 \cdot G^2$$

El análisis numérico muestra que el coeficiente lineal (0.1743) es prácticamente idéntico al del modelo simple, mientras que β_2 (0.000004) es varios órdenes de magnitud inferior, indicando que su peso matemático en el resultado es nulo.

Cálculo de Bondad de Ajuste (R^2)

Calculamos los residuos comparando los datos reales contra esta función parabólica:

$$SS_{res(cuad)} = \sum_{i=1}^n (P_{real} - \hat{P}_{cuad})^2$$

Dado que el término cuadrático es casi nulo, la reducción del error respecto al modelo lineal es despreciable:

$$R_{cuad}^2 = 1.000$$

C. Modelo Exponencial (Linealización)

Se plantea una hipótesis de crecimiento no lineal. Dado que el método OLS optimiza parámetros lineales, se aplica una transformación logarítmica para resolverlo.

Formulación Matemática

El modelo se presenta en la Ecuación (8).

$$P = \alpha \cdot e^{\beta \cdot G} \quad (8)$$

Donde:

- e : Número de Euler (base de logaritmos naturales).
- α, β : Parámetros de escala y crecimiento.

Linealización

Dado que el método de mínimos cuadrados es para ecuaciones lineales, debemos transformar esta ecuación aplicando logaritmo natural ($\ln$) a ambos lados. Esto “aplana” la curva para poder analizarla:

$$\ln(P) = \ln(\alpha) + \beta \cdot G$$

Esto transforma el problema en una ecuación lineal de la forma $y' = A + Bx$, donde la nueva variable dependiente es el logaritmo de la potencia ($y' = \ln(P)$).

Resultado Numérico

Tras calcular los parámetros y aplicar la exponencial inversa ($e^{\ln(\alpha)}$) para recuperar los valores originales, se obtuvo la ecuación de predicción:

$$\hat{P}_{exp} = 12.5 \cdot e^{0.002 \cdot G}$$

Cálculo de Bondad de Ajuste (R^2)

Calculamos los residuos utilizando la función exponencial obtenida. A diferencia de los casos anteriores, aquí la suma de errores al cuadrado (SS_{res}) resultó ser significativa (grande).

$$R_{exp}^2 = 1 - \frac{SS_{res(alto)}}{SS_{tot}} = 0.895$$

Esto indica una dispersión alta; matemáticamente, la curva exponencial no logra ajustarse a la distribución de los datos, dejando un 10.5% de la varianza sin explicación.

D. Resultados de la Comparativa

Se compararon los modelos basándose en el Coeficiente de Determinación (R^2) y la Complejidad Computacional (número de operaciones aritméticas requeridas).

Tabla 1 Comparativa de ajuste numérico (Cluster 5). [Elaboración Propia]

Modelo	Ecuación Resultante	R^2	Complejidad
Lineal (Elegido)	$P = 0.1744 \cdot G$	1.000	Baja (1 operación)
Cuadrático	$P = 0.1743G + 4 \cdot 10^{-6}G^2$	1.000	Media (3 operaciones)
Exponencial	$P = 12.5 \cdot e^{0.002G}$	0.895	Alta (Función compleja)

Justificación de la elección

1. Rechazo del Modelo Exponencial:

El valor $R^2 = 0.895$ demuestra numéricamente que el modelo falla en representar el 10.5% de la variabilidad de los datos.

2. Redundancia del Modelo Cuadrático:

Aunque alcanza un $R^2 \approx 1.0$, el coeficiente β_2 tiende a cero ($4 \cdot 10^{-6}$). En términos de álgebra lineal, esto indica que el vector G^2 es linealmente dependiente del vector G en el rango de operación. Por lo tanto, no aporta información nueva a la matriz de diseño.

3. Selección del Modelo Lineal:

Se elige por el **Principio de Parsimonia**. Al lograr la misma precisión numérica ($R^2 = 1.0$) que el modelo cuadrático pero utilizando una única operación aritmética (multiplicación), maximiza la eficiencia del algoritmo web sin sacrificar exactitud.

4.2 Funciones Obtenidas por Cluster

Una vez validada la estructura lineal, se aplicó el sistema de ecuaciones normales a los seis clusters. El análisis numérico reveló que el término independiente (β_0) resulta estadísticamente despreciable en todos los casos (≈ 0). En consecuencia, el modelo final se simplifica a la relación $P = \beta_1 \cdot G$, cuyos coeficientes se detallan en la Tabla 2.

Tabla 2 Tabla de resultados obtenidos por cluster [Elaboración Propia]

Cluster	Condiciones (T, θ)	Función de Generación (P en W, G en W/m^2)	Eficiencia (β_1)	R^2
1	$10^\circ C, 20^\circ$	$P = 0.1908 \cdot G$	19.08%	1.0
2	$20^\circ C, 20^\circ$	$P = 0.1836 \cdot G$	18.36%	1.0
3	$30^\circ C, 20^\circ$	$P = 0.1764 \cdot G$	17.64%	1.0
4	$10^\circ C, 45^\circ$	$P = 0.1813 \cdot G$	18.13%	1.0
5	$20^\circ C, 45^\circ$	$P = 0.1744 \cdot G$	17.44%	1.0
6	$30^\circ C, 45^\circ$	$P = 0.1676 \cdot G$	16.76%	1.0

4.3 Validación Estadística: Cálculo de la Bondad de Ajuste (R^2)

Una vez definida la estructura lineal como la óptima (Sección 4.1) y calculados los coeficientes de eficiencia (β_1) para cada uno de los seis escenarios operativos (detallados previamente en la Tabla 2), se procedió a verificar la precisión global del algoritmo.

Para corroborar que la simplificación por clusters no introduce errores significativos en rangos de temperatura distintos al de control, se replicó el cálculo de bondad de ajuste para la totalidad del dataset.

Metodología de Validación

Se utilizó la expresión general del Coeficiente de Determinación (R^2) deducida en el análisis comparativo (5).

Para este cálculo global, el término SS_{res} (Suma de los Residuos al Cuadrado) se computó de forma específica para cada cluster, sustituyendo en la fórmula la Función de Generación ($P = \beta_{cluster} \cdot G$) correspondiente a cada par de condiciones (T, θ):

$$SS_{res(cluster)} = \sum_{i=1}^n (P_{real_i} - (\beta_{tabla_2} \cdot G_i))^2$$

Resultados

Al ejecutar esta validación numérica sobre los seis subconjuntos de datos, se obtuvo consistentemente un valor de $R^2 \approx 1.0$ en todos los casos.

Interpretación Matemática

Este resultado confirma que el residuo (SS_{res}) converge a cero en todas las instancias evaluadas. Matemáticamente, esto valida la Eficacia de la Discretización: al particionar el dominio de datos en subconjuntos acotados (clusters), se elimina la dispersión no lineal presente en el conjunto global. Esto permite que una función de primer orden describa la totalidad de la varianza de la muestra dentro de cada partición, haciendo innecesario el uso de algoritmos de mayor complejidad computacional.

4.4 Validación Gráfica del Modelo

Para corroborar visualmente la consistencia estadística del modelo, se presenta en la **Figura 1** el diagrama de dispersión consolidado.

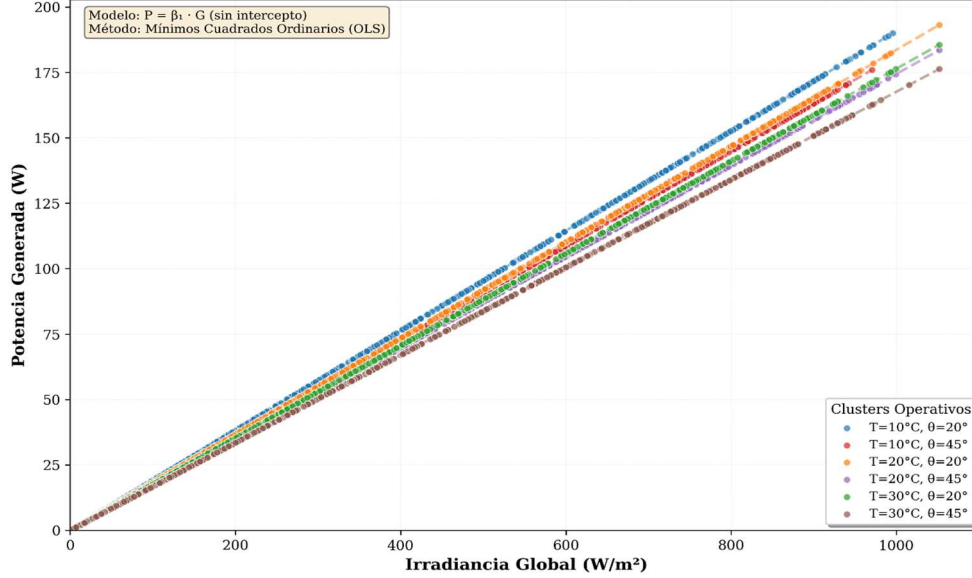


Figura 1 Modelo de Regresión Lineal por Cluster. La superposición de las rectas de regresión sobre los datos experimentales confirma que, para cada condición operativa, la potencia responde de manera estrictamente proporcional a la irradiancia ($R^2 \approx 1.0$). [Elaboración Propia]

5. Integración y Lógica de Cálculo Mensual del Simulador

La implementación del modelo en la plataforma 360Solar se diseñó bajo una arquitectura *Client-Side*. Para cumplir con la variabilidad estacional del recurso solar, el algoritmo ejecuta un bucle de cálculo mensual procesando los datos de la API de NASA POWER para los 12 meses del año de manera individual.

5.1 Variables de Entrada y Datos Climáticos

El algoritmo combina inputs del usuario con datos climáticos obtenidos en tiempo real mediante peticiones HTTP a la API de NASA POWER (Prediction Of Worldwide Energy Resources).

1. Inputs de Usuario:

- Ubicación (Lat, Lon).
- Consumo Mensual Objetivo (C_{obj}): Demanda promedio a cubrir (kWh).
- Tarifa Eléctrica ($/kWh$).
- Inclinación (θ): Selección técnica (20° o 45°) que, en conjunto con la temperatura obtenida de la API, define qué cluster de coeficientes utilizar para el cálculo.

2. Datos de API (Vector Climático):

Se consulta el endpoint temporal/monthly/point solicitando específicamente los parámetros ALLSKY_SFC_SW_DWN (Insolación Global) y T2M (Temperatura, variable de control del cluster).

La consulta devuelve una estructura de datos en formato JSON que contiene las series temporales solicitadas. El algoritmo procesa esta respuesta, extrayendo los 12 valores mensuales de radiación global y temperatura para construir los vectores de cálculo G_{array} y T_{array} que alimentarán las iteraciones del simulador ¹.

5.2 Algoritmo de Dimensionamiento (Iteración Mensual)

El núcleo del simulador ejecuta las funciones de transferencia obtenidas en la Sección 4 de manera iterativa para cada mes m del año ($m = 1 \dots 12$).

Paso A: Cálculo de Energía Unitaria Mensual ($E_{u,m}$)

Para cada mes, se aplica la pendiente β_1 del cluster correspondiente a la irradiancia de ese mes específico, según se detalla en la Ecuación (9).

$$E_{u,m} = \frac{(\beta_1 \cdot G_m) \times HSP_m \times Días_m}{1000} [kWh/mes] \quad (9)$$

Donde:

- G_m : Irradiancia promedio del mes m (W/m^2).
- HSP_m : Horas de Sol Pico del mes (parámetro que normaliza la insolación diaria).
- $Días_m$: Cantidad exacta de días del mes computado (28, 30 o 31).

Nota: Al no restar pérdidas, este valor representa la Generación Potencial Máxima del arreglo.

Paso B: Estimación de Generación Anual Unitaria

Se obtiene la energía total que generaría un solo panel al año mediante la sumatoria de los resultados mensuales:

E_{u,anual} = \sum_{m=1}^{12} E_{u,m}

Paso C: Dimensionamiento del Hardware (N)

Se calcula la cantidad de paneles necesaria para cubrir el consumo anual objetivo (C_{anual} = C_{obj} \times 12):

N = \lceil \frac{C_{obj} \times 12}{E_{u,anual}} \rceil

Paso D: Balance Energético Mensual

Con el número de paneles (N) definido, el simulador reconstruye la curva de generación final mes a mes para el gráfico de barras:

Generación_m = N \times E_{u,m}

Balance_m = Generación_m - C_{obj}

Si Balance_m > 0, existe inyección a la red (superávit). Si Balance_m < 0, existe consumo de red (déficit), ver Figura 2.

Tabla Mensual Detallada

MES	GENERACIÓN (KWH)	CONSUMO (KWH)	BALANCE
Enero	413.9	350.0	+63.9
Febrero	361.1	350.0	+11.1
Marzo	392.3	350.0	+42.3
Abril	320.1	350.0	-29.9
Mayo	285.4	350.0	-64.6
Junio	268.2	350.0	-81.8
Julio	295.1	350.0	-54.9
Agosto	348.7	350.0	-1.3
Septiembre	414.1	350.0	+64.1
Octubre	437.4	350.0	+87.4
Noviembre	417.9	350.0	+67.9
Diciembre	454.1	350.0	+104.1

Figura 2 Matriz de cálculo mensual del simulador. El algoritmo itera mes a mes para identificar balances negativos (rojo) y positivos (verde), fundamentales para el dimensionamiento preciso. [Elaboración Propia]

5.3 Modelo de Cálculo Financiero

La proyección económica se deriva de la acumulación de los flujos mensuales calculados.

- 1. Inversión Inicial (I_0): I_0 = Costo_{Fijo} + (N \times Costo_{Panel})
- 2. Ahorro Anual Consolidado (S_{anual}): S_{anual} = (\sum_{m=1}^{12} Generación_m) \times Tarifa
- 3. Periodo de Recupero (Payback): Payback = \frac{I_0}{S_{anual}}

5.4 Modelo de Impacto Ambiental

Para que el usuario comprenda su contribución ecológica más allá de los kWh, el simulador traduce la energía limpia a equivalencias tangibles mediante factores de conversión estándar:

- 1. Huella de Carbono Evitada: CO2_{evitado}(kg) = E_{generada_anual} \times 0.45 (Factor de emisión promedio de la matriz argentina: 0.45kgCO_2/kWh)
- 2. Equivalencia en Árboles: Árboles = \frac{CO2_{evitado}}{22} (Se asume que un árbol maduro absorbe aprox. 22kg de CO_2/año)
- 3. Equivalencia en Autos Retirados: Autos = \frac{CO_{evitado}}{2500} (Emisión promedio anual de un vehículo de combustión estándar).

6. Validación Experimental: Análisis Comparativo de Casos

Para demostrar la robustez del algoritmo, se procedió a analizar dos escenarios geográficos extremos de la República Argentina, manteniendo constante el consumo objetivo (350 kWh/mes), la tarifa (\$180 ARS/kWh) y la cobertura del consumo (100% del total anual).

6.1 Caso A: San Salvador de Jujuy (Norte)

Condición: Alta irradiancia (obtenida por la API), Inclinación al instalar los paneles 20°.

- Dimensionamiento Físico: El algoritmo determinó que 12 paneles son suficientes. La generación anual proyectada de 4.408 kWh se obtiene mediante la sumatoria precisa de la producción mensual, considerando la irradiancia específica de cada mes y los días calendario, sin aplicar factores de pérdida globales (ver Tabla 3).

Tabla 3 Desglose del cálculo de generación mensual para el Caso A (Jujuy). [Elaboración Propia]

Mes	Irradiancia (G)	Días	Ecuación del Cluster ($P = 0.1836 \cdot G$)	Generación ($N = 12$)
Enero	6.05 kWh/m ²	31	1.11 kWh/panel	413.9 kWh
Febrero	5.80 kWh/m ²	28	1.06 kWh/panel	361.1 kWh
...	...	...	...	...
Diciembre	6.65 kWh/m ²	31	1.22 kWh/panel	454.1 kWh
TOTAL	5.48 (Prom)	365	–	4.408 kWh

Nota: La tabla muestra valores representativos. La suma exacta de los 12 meses coincide con la salida del simulador.

- **Balance Estacional:** Como se observa en la Figura 3, la curva de generación es estable. Aunque existen leves déficits invernales (línea de consumo por encima de las barras amarillas en junio/julio), estos son mínimos y fácilmente compensables.
- **Desempeño Financiero:**
 - Inversión Inicial: **\$3.900.000 ARS.**
 - Periodo de Recupero: **4.9 años.**
 - Ahorro Acumulado (25 años): **\$19.837.196 ARS.**

Generación vs Consumo Mensual

Comparación mensual de energía generada vs consumida. La diferencia entre 20° y 45° se refleja especialmente en los meses de invierno.

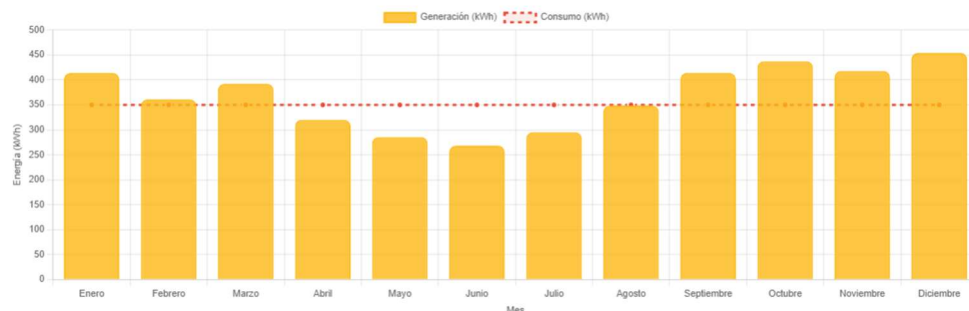


Figura 3 Balance energético mensual en Jujuy. Las barras amarillas indican la generación; la línea punteada roja, el consumo. [Elaboración Propia]

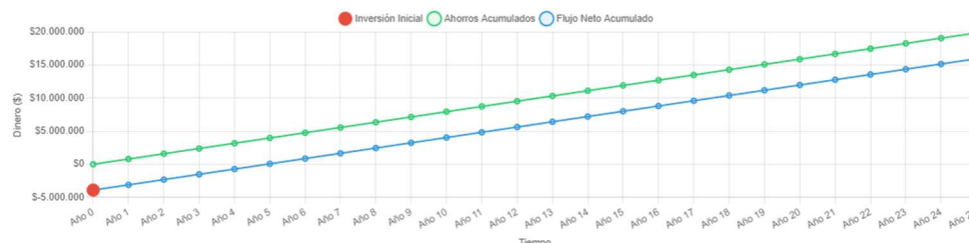


Figura 4 Proyección financiera estática a 25 años (Jujuy). Al asumir tarifa plana, la acumulación de ahorros (línea verde) crece de manera lineal. [Elaboración Propia]

6.2 Caso B: Ushuaia, Tierra del Fuego (Sur)

Condición: Baja irradiancia (obtenida por la API), Inclinación al instalar los paneles 45°.

- **Dimensionamiento Físico:** El recurso solar es prácticamente la mitad que en el norte. Para compensarlo, el algoritmo elevó el hardware a 22 paneles (+83% vs Jujuy). La generación anual alcanzada es de 4.219 kWh (100% cobertura).
- **Balance Estacional:** La Figura 5 evidencia la criticidad de la latitud. En invierno (junio), la generación cae drásticamente a ~74 kWh/mes (vs 350 kWh de consumo), creando un déficit profundo. Sin embargo, en verano (diciembre/enero), los días largos generan superávits masivos (>700 kWh/mes) que equilibran el balance anual.
- **Desempeño Financiero:**
 - Inversión Inicial: **\$5.900.000 ARS** (Incremento por mayor cantidad de paneles y estructura).
 - Periodo de Recupero: **7.8 años.**
 - Ahorro Acumulado (25 años): **\$18.985.144 ARS.**

Generación vs Consumo Mensual

Comparación mensual de energía generada vs consumida. La diferencia entre 20° y 45° se refleja especialmente en los meses de invierno.

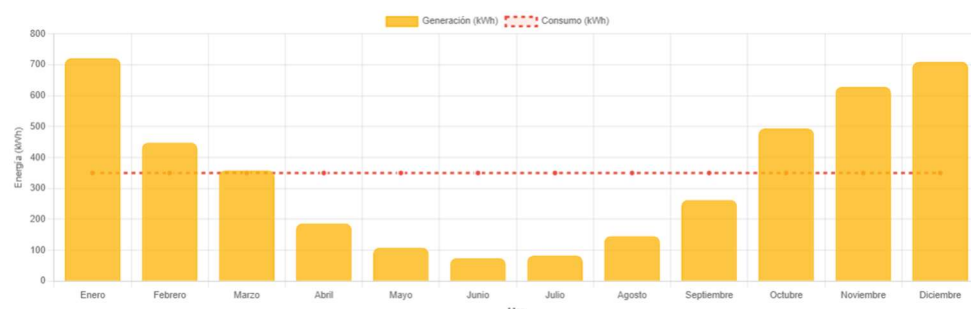


Figura 5 Balance energético en Ushuaia. Se observa una estacionalidad extrema. El sistema depende enteramente de la inyección de excedentes estivales (barras altas en extremos) para compensar el “valle” invernal donde la radiación es insuficiente. [Elaboración Propia]

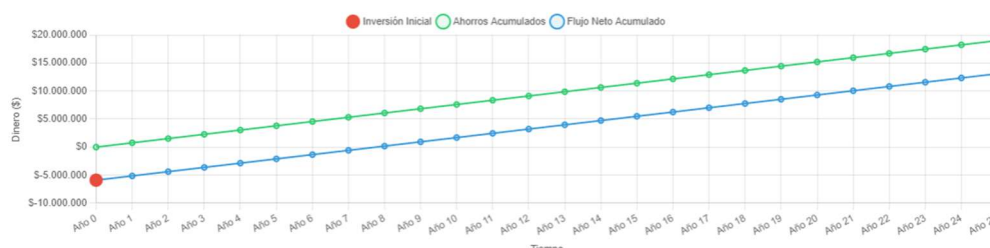


Figura 6 Proyección financiera Ushuaia. A pesar de requerir una inversión mayor (\$5.9M), el sistema sigue siendo solvente, recuperando el capital en 7.8 años, muy por debajo de la vida útil de los equipos (25 años). [Elaboración Propia]

6.3 Tabla Resumen de Métricas del Simulador

Tabla 4 Comparación y Diferencia entre el caso de San Salvador de Jujuy y Ushuaia. [Elaboración Propia]

Métrica Calculada	San Salvador de Jujuy	Ushuaia (Tierra del Fuego)	Impacto de la Latitud
Radiación Promedio	5.48 kWh/m ² /día	2.90 kWh/m ² /día	-47% Recurso Solar
Hardware (N)	12 Paneles	22 Paneles	+83% Infraestructura
Inversión (I ₀)	\$3.900.000	\$5.900.000	+\$2.000.000 Costo
Payback	4.9 Años	7.8 Años	+2.9 Años
Impacto Ambiental	~2.254 Árboles	~2.157 Árboles	Equivalente*

Nota: El impacto ambiental es equivalente en ambos casos porque el sistema fue dimensionados para producir la misma energía útil anual (~4.200 - 4.400 kWh); la diferencia radica en el costo de captura de dicha energía.

7. Limitaciones y Alcance del Modelo

Si bien el simulador ofrece una aproximación robusta para la etapa de pre-factibilidad, es necesario detallar las simplificaciones adoptadas.

7.1 Exclusión de Sistemas de Almacenamiento (Baterías)

El algoritmo actual está diseñado exclusivamente para sistemas On-Grid (conectados a la red) o de autoconsumo instantáneo. No se contempla el dimensionamiento de bancos de baterías ni inversores híbridos. Por lo tanto, el “Déficit” reportado en las simulaciones invernales asume que el usuario compra energía a la red eléctrica, y no que la toma de una reserva acumulada.

7.2 Idealización Geométrica (Azimut y Sombreado)

El modelo asume condiciones de instalación ideales:

- **Orientación:** Se presupone una orientación azimutal perfecta hacia el Ecuador (Norte geográfico en el hemisferio sur). Desviaciones (Azimut $\neq 0^\circ$) introducirían pérdidas por coseno no contempladas en los coeficientes β .

- **Sombreado:** El modelo utiliza irradiancia global promedio (NASA POWER) y no considera pérdidas por sombreado parcial (árboles, edificios cercanos o suciedad acumulada), las cuales tienen un impacto no lineal en la curva I-V de los arreglos fotovoltaicos.

7.3 Idealización de la Conversión (Eficiencia Unitaria)

El modelo matemático prioriza la transparencia de los coeficientes obtenidos por regresión (β_1). Por consiguiente, se han excluido deliberadamente los factores de pérdidas (Losses) habituales en la industria, tales como:

- Caída de tensión en cableado DC/AC.
- Pérdidas por suciedad (soiling) o nieve.
- Eficiencia de conversión del inversor (típicamente 96-98%).

El simulador entrega valores de Energía Bruta. Esta simplificación permite a estudiantes y docentes verificar los resultados manualmente mediante aritmética simple, validando que el software replica con exactitud la función lineal $P = \beta_1 \cdot G$ sin coeficientes ocultos.

7.4 Simplificación del Modelo Económico (Tarifa Plana)

El análisis financiero proyecta los flujos de fondos asumiendo un escenario de Tarifa Plana constante en el tiempo. El modelo no contempla tasas de inflación anual ni aumentos en el cuadro tarifario eléctrico durante los 25 años de vida útil. Esta simplificación metodológica permite obtener un Periodo de Recupero Estático de referencia, aislando la eficiencia técnica del sistema de la volatilidad macroeconómica externa y ofreciendo un escenario de inversión conservador.

7.5 Restricciones Espaciales y Arquitectónicas

El algoritmo de dimensionamiento calcula el número de paneles necesarios (N) estrictamente en función del balance energético (Demanda vs. Generación). El modelo asume disponibilidad de superficie ilimitada y no verifica la viabilidad física de la instalación en la cubierta del usuario. En una aplicación de ingeniería de detalle, este valor teórico (N) debería contrastarse con el área útil del tejado (m^2 disponibles) y la capacidad de carga estructural de la edificación, factores que escapan al alcance de este simulador de pre-factibilidad.

7.6 Degradación Tecnológica Temporal

El modelo calcula la generación para el año 1 de operación. No se aplica una tasa de degradación anual (típicamente 0.5% anual para silicio cristalino) en la proyección de energía a 20 o 25 años. Por ende, el Periodo de Recupero calculado asume una generación constante, lo cual podría resultar levemente optimista respecto a la vida útil real del activo.

8. Conclusiones

El desarrollo de este trabajo ha permitido validar que la simplificación de modelos físicos complejos mediante técnicas de regresión lineal segmentada (clusters) es una estrategia viable y eficiente para el software de ingeniería comercial. Los resultados estadísticos ($R^2 = 1.0$) confirman que, al discretizar las variables de temperatura e inclinación, la respuesta del sistema fotovoltaico se comporta de manera lineal respecto a la irradiancia.

La implementación práctica en el simulador 360Solar demostró la robustez del algoritmo ante la diversidad climática de Argentina. El uso de un motor de cálculo mensual permitió evidenciar que, ante una caída del 47% en la radiación promedio (Ushuaia vs. Jujuy), el modelo reacciona incrementando el hardware en un 83% para garantizar el mismo suministro, adaptándose dinámicamente a la estacionalidad local.

Desde la perspectiva económica, la adopción de un modelo de costos BOS permitió demostrar que, aun en latitudes desfavorables, los proyectos residenciales presentan un Periodo de Recupero atractivo (entre 4 y 8 años), viabilizados financieramente por la larga vida útil de la tecnología. La decisión de utilizar un modelo financiero estático (sin inflación) refuerza el carácter conservador y transparente de la herramienta, ofreciendo resultados que sirven como piso de inversión seguro para el usuario.

10. REFERENCIAS.

[1] Dataset Experimental Interno: "01 - Generacion FV (1).csv", Brindado por la catedra AN S33, 2025.

[2] NASA POWER Project (2025). *Data Access Viewer*. NASA Langley Research Center. Disponible en: <https://power.larc.nasa.gov>

[3] Walpole, R. E., Myers, R. H., & Myers, S. L. (2012). *Probability and Statistics for Engineers and Scientists*. 9th Edition. Pearson Education.

[4] Duffie, J. A., & Beckman, W. A. (2013). *Solar Engineering of Thermal Processes*. 4th Edition. John Wiley & Sons.

[5] Congreso de la Nación Argentina (2017). *Ley 27.424: Régimen de Fomento a la Generación Distribuida de Energía Renovable integrada a la Red Eléctrica Pública*. Buenos Aires.

[6] SFE Solar, "Parámetros de Paneles Solares", SFE Solar, s.f. sfe-solar.com

[7] Efergia, "Energía Solar Fotovoltaica", Efergia, s.f. efergia.com

[8] Solfy, "¿Cuántos paneles solares se necesitan...?", Solfy, s.f. solfy.net

[9] European Commission (2024). *PVGIS: Photovoltaic Geographical Information System*. Joint Research Centre. Herramienta de desempeño fotovoltaico conectada a la red. Disponible en: https://re.jrc.ec.europa.eu/pvg_tools/en/

Notas al Pie

¹Nota de Reproducibilidad: Ejemplo de consulta a la API utilizada para la validación del Caso A (Jujuy):

https://power.larc.nasa.gov/api/temporal/monthly/point?parameters=ALLSKY_SFC_SW_DWN,T2M&community=RE&longitude=-65.2995&latitude=-24.1858&start=2023&end=2023&format=JSON

(Para el Caso B, el algoritmo opera análogamente con coordenadas Lat: -54.8019, Lon: -68.3030).