



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA MECÁNICA

**EVALUACIÓN DE TÉCNICAS DE APRENDIZAJE PROFUNDO PARA EL  
PRONÓSTICO DE GENERACIÓN ELÉCTRICA EN CORTO PLAZO DE UN  
SISTEMA FOTOVOLTAICO DE 16 KW**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL MECÁNICO

**CRISTIÁN JAVIER HERRERA BUSTAMANTE**

PROFESOR GUÍA:  
ENRIQUE LÓPEZ DROGUETT

MIEMBROS DE LA COMISIÓN:  
JOSÉ CARDEMIL  
JUAN TAPIA

SANTIAGO, CHILE  
2020

RESUMEN DE LA MEMORIA PARA OPTAR  
AL TÍTULO DE MAGÍSTER EN CIENCIAS  
DE LA INGENIERÍA  
POR: **CRISTIÁN JAVIER HERRERA BUSTAMANTE**  
FECHA: 2020  
PROF. GUÍA: ENRIQUE LÓPEZ DROGUETT

## **EVALUACIÓN DE TÉCNICAS DE APRENDIZAJE PROFUNDO PARA EL PRONÓSTICO DE GENERACIÓN ELÉCTRICA EN CORTO PLAZO DE UN SISTEMA FOTOVOLTAICO DE 16 KW**

El acelerado y súbito crecimiento que ha experimentado la tecnología solar fotovoltaica (PV) en el último tiempo en Chile, supone un desafío a futuro para los agentes del sector eléctrico encargados de la coordinación y distribución de la electricidad en el sistema interconectado. Dada la directa dependencia que este tipo de tecnología posee respecto a las condiciones climáticas y meteorológicas que influyen sobre la irradiación incidente en la superficie, significa que una parte importante de la generación en el país estará sujeta a fluctuaciones irregulares que dependerán principalmente de la disponibilidad variable del Sol en las distintas zonas de la región. Por tanto, el desarrollo de modelos que permitan pronosticar confiablemente la generación energética de una planta solar PV en el corto plazo será de suma relevancia, en los próximos años, para lograr mantener una operación balanceada e integral del sistema interconectado.

Así, el objetivo en este trabajo de título consiste en el desarrollo, mediante técnicas de aprendizaje profundo, de modelos de pronóstico de generación energética, en corto plazo, del sistema fotovoltaico situado en el complejo de Beauchef 851 de la Facultad de Ciencias Físicas y Matemáticas. Para cumplir lo anterior, tres arquitecturas de redes neuronales: LSTM, JANet y JANet-ConvJANet, son evaluadas con respecto a su capacidad de pronóstico de la potencia fotovoltaica. En particular, se realiza una evaluación de desempeño de los modelos desarrollados frente a distintos contextos meteorológicos. Por otro lado, con el fin de contar con una base de referencia, una serie de métodos estocásticos (Persistencia, AR y KNN) son empleados para comparar los resultados de los modelos. Por último, se estudian las distribuciones del error de pronóstico obtenidas para cada modelo y se analiza su correlación con la variabilidad atmosférica.

De los resultados, se obtiene que las tres arquitecturas desarrolladas presentan desempeños considerablemente superiores a los obtenidos con métodos estocásticos de pronóstico. En particular, la arquitectura JANet-ConvJANet es la que presenta la mayor capacidad de pronóstico en todos los contextos meteorológicos analizados, alcanzando errores RMSE hasta un 9% menor respecto a los modelos LSTM y JANet. Así, se verifica la capacidad de la estructura ConvJANet de interpretar secuencias de imágenes satelitales para un mejor ajuste de las estimaciones de potencia eléctrica generada.

*A mis padres*

# Tabla de Contenidos

<b>1. Introducción</b>	<b>1</b>
1.1. Objetivos y Alcances . . . . .	2
1.1.1. Objetivo General . . . . .	2
1.1.2. Objetivos Específicos . . . . .	2
1.1.3. Alcances . . . . .	3
<b>2. Antecedentes</b>	<b>4</b>
2.1. Formulación del pronóstico fotovoltaico . . . . .	4
2.2. Irradiación solar . . . . .	4
2.3. Modelos estocásticos . . . . .	7
2.3.1. Autoregresión (AR) . . . . .	7
2.3.2. k-Nearest-Neighbors (kNN) . . . . .	7
2.3.3. Método de persistencia . . . . .	8
2.4. Aprendizaje profundo . . . . .	9
2.4.1. Densely Connected Neural Networks . . . . .	9
2.4.2. Long Short-Term Memory RNN . . . . .	10
2.4.3. Just Another Network . . . . .	11
2.4.4. Convolutional JANet . . . . .	12
<b>3. Revisión bibliográfica</b>	<b>13</b>
3.1. Redes ANN para pronóstico fotovoltaico . . . . .	13
3.2. Redes LSTM para pronóstico fotovoltaico . . . . .	14
3.3. Imágenes satelitales para pronóstico solar . . . . .	15
<b>4. Metodología</b>	<b>17</b>
4.1. Datos y recursos disponibles . . . . .	17
4.1.1. Sistema fotovoltaico . . . . .	17
4.1.2. Estación solarimétrica . . . . .	18
4.1.3. Imágenes satelitales NOAA . . . . .	20
4.2. Pre-procesamiento de datos . . . . .	20
4.2.1. Exploración de datos . . . . .	21
4.3. Clustering para clasificación solar de días . . . . .	22
4.4. Generación de conjuntos de datos de entrenamiento y validación . . . . .	24
4.4.1. Datos del sistema y de irradiación . . . . .	25
4.4.2. Datos satelitales del Advanced Baseline Imager . . . . .	26
4.5. Configuración de arquitecturas . . . . .	27
4.5.1. Modelos LSTM y JANet . . . . .	29

4.5.2. Modelo JANet-ConvJANet . . . . .	29
4.6. Entrenamiento . . . . .	30
4.7. Métricas de evaluación . . . . .	31
<b>5. Resultados y discusión</b>	<b>32</b>
5.1. Análisis modelos LSTM y JANet . . . . .	33
5.1.1. Desempeño en días soleados . . . . .	33
5.1.2. Desempeño en días nublados . . . . .	36
5.1.3. Desempeño en días parcialmente nublados . . . . .	37
5.2. Análisis modelo JANet-ConvJANet . . . . .	39
5.2.1. Desempeño en días soleados . . . . .	40
5.2.2. Desempeño en días nublados . . . . .	41
5.2.3. Desempeño en días parcialmente nublados . . . . .	44
5.3. Resumen . . . . .	45
<b>6. Conclusiones</b>	<b>48</b>
<b>Bibliografía</b>	<b>50</b>

# Índice de Tablas

4.1.	Resumen técnico correspondiente a la banda espectral Cirrus del ABI. . . . .	20
4.2.	Características de los clusters identificados por el algoritmo DBSCAN. . . . .	24
4.3.	Características de los conjuntos de datos generados. . . . .	25
4.4.	Proporción de <i>samples</i> asociados a tiempos meteorológicos en los conjuntos de entrenamiento y validación. . . . .	26
4.5.	Descripción de los atributos ( <i>features</i> ) utilizados en la generación de los conjuntos de entrenamiento y validación. . . . .	27
5.1.	Métricas estadísticas obtenidas en la evaluación de desempeño de los métodos de estocásticos de referencia. . . . .	32
5.2.	Métricas estadísticas obtenidas en la evaluación de desempeño del modelo LSTM. . . . .	33
5.3.	Métricas estadísticas obtenidas en la evaluación de desempeño del modelo JANet. . . . .	33
5.4.	Métricas estadísticas obtenidas en la evaluación de desempeño del modelo JANet-ConvJANet. . . . .	40
5.5.	Métricas estadísticas de evaluación para cada uno de los métodos empleados . . . . .	47

# Índice de Ilustraciones

2.1.	Comparación entre el espectro de la irradiancia incidente en la superficie terrestre (a nivel del mar) y la extraterrestre, en un día soleado [7]. . . . .	6
2.2.	Comparación entre las irradiancias en superficie horizontal registradas entre el 14 y 17 de Septiembre del 2018 y la extraterrestre correspondiente. . . . .	6
2.3.	Esquema de una neurona artificial. . . . .	10
2.4.	Esquema de una celda LSTM. . . . .	11
2.5.	Esquema de una celda ConvJANet [12] . . . . .	12
3.1.	Registro temporal de imágenes satelitales sobre el área estudiada [4]. . . . .	16
3.2.	Selección de píxeles con mayor correlación dentro de la secuencia de imágenes satelitales con los datos registrados en terreno [4]. . . . .	16
4.1.	Imágen satelital del sistema fotovoltaico de 16.12 kW estudiado. La etiqueta A corresponde a la ubicación de instalación fotovoltaica, mientras que la etiqueta B corresponde a la ubicación de la estación solarimétrica. . . . .	18
4.2.	Gráfico de potencia generada por el sistema fotovoltaico (todos los inversores) durante el periodo analizado. . . . .	18
4.3.	Gráfico de temperatura en superficie del sistema fotovoltaico durante el periodo analizado. . . . .	19
4.4.	Gráfico de irradiación GHI ( $Wh/m^2$ ) registrado por la estación solarimétrica. . . . .	19
4.5.	Gráfico de temperatura ambiente durante el periodo analizado. . . . .	20
4.6.	Gráfico de fracción difusa versus índice de claridad. Horas en UTC. . . . .	21
4.7.	Gráfico de potencia generada versus irradiancia global en el plano horizontal. Horas en UTC. . . . .	22
4.8.	Gráfico de performance ratio versus índice de claridad. Horas en UTC. . . . .	22
4.9.	Gráfico del clustering realizado sobre los parámetros. (a) índice de nubosidad versus índice de claridad, (b) índice de suavidad versus índice de claridad y (c) índice de suavidad versus índice de nubosidad. . . . .	23
4.10.	Gráfico de irradiancia global en los días correspondientes a cada cluster. (a) Cluster A, (b) Cluster B y (c) Cluster C. . . . .	24
4.11.	Muestra de procedimiento de segmentación de periodos para la generación de los conjuntos de datos de entrenamiento. Los datos sombreados de color morado son procesados para el set de entrenamiento, mientras que los sombreados de color amarillos son procesados para el set de validación. (Periodo entre el 15 y 25 de Septiembre del 2018). . . . .	25
4.12.	Muestra de procedimiento de generación de ventanas de datos para la creación de los conjuntos $X$ e $Y$ . . . . .	26
4.13.	Muestra de procedimiento de generación de ventanas de datos para la creación de los conjuntos $X$ de secuencias de imágenes. . . . .	28

4.14.	Muestra de una secuencia de imágenes satelitales procesadas. Periodo entre las 14:00 y 20:00 del 2 de Diciembre del 2018. . . . .	28
4.15.	Esquema de las arquitecturas empleadas para los modelos LSTM y JANet. . .	29
4.16.	Esquema de la arquitectura empleada para el modelo JANet-ConvJANet. . . .	30
5.1.	Dispersión de los pronósticos del modelo JANet con respecto a los valores reales para cada uno de los intervalos del horizonte de pronóstico en días soleados. .	34
5.2.	Distribución del error de pronóstico del modelo JANet para cada uno de los intervalos del horizonte de pronóstico en días soleados. . . . .	35
5.3.	Error absoluto promedio del modelo JANet versus la variabilidad atmosférica para cada uno de los intervalos del horizonte de pronóstico en días soleados. .	35
5.4.	Dispersión de los pronósticos del modelo JANet con respecto a los valores reales para cada uno de los intervalos del horizonte de pronóstico en días nublados. .	36
5.5.	Distribución del error de pronóstico del modelo JANet para cada uno de los intervalos del horizonte de pronóstico en días nublados. . . . .	37
5.6.	Error absoluto promedio del modelo JANet versus la variabilidad atmosférica para cada uno de los intervalos del horizonte de pronóstico en días nublados. .	38
5.7.	Dispersión de los pronósticos del modelo JANet con respecto a los valores reales para cada uno de los intervalos del horizonte de pronóstico en días parcialmente nublados. . . . .	38
5.8.	Distribución del error de pronóstico del modelo JANet para cada uno de los intervalos del horizonte de pronóstico en días parcialmente nublados. . . . .	39
5.9.	Error absoluto promedio del modelo JANet versus la variabilidad atmosférica para cada uno de los intervalos del horizonte de pronóstico en días parcialmente nublados. . . . .	40
5.10.	Dispersión de los pronósticos del modelo JANet-ConvJANet con respecto a los valores reales para cada uno de los intervalos del horizonte de pronóstico en días soleados. . . . .	41
5.11.	Distribución del error de pronóstico del modelo JANet-ConvJANet para cada uno de los intervalos del horizonte de pronóstico en días soleados. . . . .	42
5.12.	Error absoluto promedio en el modelo JANet-ConvJANet versus la variabilidad atmosférica para cada uno de los intervalos del horizonte de pronóstico en días soleados. . . . .	42
5.13.	Dispersión de los pronósticos del modelo JANet-ConvJANet con respecto a los valores reales para cada uno de los intervalos del horizonte de pronóstico en días nublados. . . . .	43
5.14.	Distribución del error de pronóstico del modelo JANet-ConvJANet para cada uno de los intervalos del horizonte de pronóstico en días nublados. . . . .	43
5.15.	Error absoluto promedio del modelo JANet-ConvJANet versus la variabilidad atmosférica para cada uno de los intervalos del horizonte de pronóstico en días nublados. . . . .	44
5.16.	Dispersión de los pronósticos del modelo JANet-ConvJANet con respecto a los valores reales para cada uno de los intervalos del horizonte de pronóstico en días parcialmente nublados. . . . .	45
5.17.	Distribución del error de pronóstico del modelo JANet-ConvJANet para cada uno de los intervalos del horizonte de pronóstico en días parcialmente nublados. . . . .	46

5.18. Error absoluto promedio del modelo JANet-ConvJANet versus la variabilidad atmosférica para cada uno de los intervalos del horizonte de pronóstico en días parcialmente nublados. . . . .	46
--	----

# Capítulo 1

## Introducción

En la última década, ha sido evidente el acelerado y súbito desarrollo que ha experimentado el sector eléctrico de Chile en lo que respecta a la integración de tecnologías renovables no convencionales (ERNC) en la matriz energética del país. Cerca del 18.9 % de la capacidad instalada actualmente proviene de este tipo de tecnologías, donde la solar fotovoltaica (solar PV) es la que más destaca con una contribución de un 10.2 % para Abril del 2019 (equivalente a unos 2.38 GW). Esta tecnología, seguida por la eólica y, próximamente, por la solar de concentración (CSP), ha demostrado tener el potencial de liderar la expansión y transformación energética del país en los próximos años. Se estima que para el año 2025, el crecimiento de la solar PV sea del orden de 2 GW en capacidad instalada, elevando considerablemente y en pocos años su contribución a la matriz nacional en conjunto con el resto de las tecnologías renovables, con las cuales se espera alcanzar una integración del 50 % para el mismo año respecto a la capacidad total del país [1].

Ahora bien, este rápido crecimiento en los próximos años por parte de la solar PV supondrá un desafío importante para los agentes encargados de la coordinación y distribución de la electricidad en el país, e incluso para el resto de los participantes del sector y de la industria. Producto de la directa influencia de las condiciones climáticas sobre la irradiación en la superficie y, por tanto, sobre la producción misma en este tipo de plantas, parte importante de la generación en el país estará sujeta a fluctuaciones que dependerán principalmente de la disponibilidad temporal del sol en las distintas zonas de la región. De no manejarse adecuadamente este tipo de fluctuaciones en lo que concierne a la distribución de energía, se pone en un riesgo significativo la estabilidad del sistema eléctrico. Esto último, respecto a su disponibilidad, como también respecto al costo marginal de producción, debido a que para balancear estas variaciones se requiere en ocasiones de una rápida respuesta por parte de las generadoras de carga base como lo son las termoeléctricas y las hidroeléctricas.

Así, modelos numéricos o bien, computacionales que permitan reducir la incertidumbre en tiempo real respecto a la operación, generación y/o disponibilidad de las plantas solares PV en el corto plazo, resultan en herramientas sumamente útiles a la hora de mantener una operación eficiente e integral del sistema interconectado. Mientras más información se posea y mejor sean las estimaciones respecto a los pronósticos de las próximas horas, mejores decisiones podrán ser evaluadas e implementadas en cuanto a la coordinación de la red: planificación de demanda y carga a otros medios de generación, programación de sistemas de respaldo, optimización del balance y estabilidad en la matriz, y estimación del futuro comportamiento

del costo marginal y la venta de energía en el mercado spot, entre otros [2].

En los últimos años, múltiples métodos han sido desarrollados y estudiados para el pronóstico o forecasting de la irradiación solar con el fin de mejorar las estimaciones de producción energética en plantas PV y CSP. Existen modelos basados en el procesamiento y análisis del movimiento de las nubes a partir de imágenes registradas por cámaras de cielo [3] y mediante satélites geoestacionarios [3, 4] utilizando técnicas tanto estadísticas como numéricas de predicción; y otros, en su gran mayoría, basados en el procesamiento temporal de datos históricos mediante técnicas de regresión tradicionales (lineales como no-lineales) y de aprendizaje profundo (Deep Learning) con diversas arquitecturas de redes neuronales (ANN), introduciendo tanto redes convolucionales (CNN) [5], como recurrentes (RNN) como las Long-Short Term Memory (LSTM) [2, 6].

No obstante, en este tipo de métodos de forecasting no se suelen considerar en el modelamiento los datos operativos de las plantas PV, pues el pronóstico se limita al de la irradiación solar en la región correspondiente y no al de la producción energética, siendo que en ocasiones otros factores pudiesen afectar ésta última: malfuncionamiento de trackers, temperatura y grado de ensuciamiento en la superficie de los paneles, degradación de las celdas fotovoltaicas, y sombras de estructuras aledañas, solo por mencionar algunos. Así, para el desarrollo de modelos más robustos de pronóstico de generación resulta necesario incorporar variables relacionadas a los aspectos físicos y operativos de la planta.

Considerando esto último, en este trabajo de título, se evaluarán distintas arquitecturas de redes neuronales profundas para el desarrollo de modelos de pronóstico de una planta PV de pequeña escala que permitan la estimación de generación eléctrica en el corto plazo a partir de datos meteorológicos y operacionales.

## **1.1. Objetivos y Alcances**

### **1.1.1. Objetivo General**

Desarrollar, mediante técnicas de aprendizaje profundo, modelos de pronóstico de generación de energía eléctrica, en corto plazo, del sistema fotovoltaico instalado en el techo del edificio oriente del complejo Beauchef 851 de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile.

### **1.1.2. Objetivos Específicos**

- Evaluar el desempeño de las arquitecturas desarrolladas frente a distintos contextos y tiempos meteorológicos: días soleados, nublados y nublados parciales.
- Comparar la capacidad de pronóstico de los modelos resultantes con respecto a técnicas y métodos estocásticos tradicionales.
- Estudiar la distribución del error de pronóstico resultante de cada modelo y analizar su correlación con la variabilidad atmosférica (respecto a irradiación solar).

### 1.1.3. Alcances

Dado el objetivo principal del trabajo, el alcance del proyecto radica en el modelamiento específico de la operación del arreglo de paneles fotovoltaicos ya mencionado. Así, no se espera del trabajo generar un modelo generalizable de *forecasting* que pueda ser aplicado directamente en otras plantas, sino que se busca la construcción de un modelo altamente ligado al desempeño del sitio estudiado.

Por otro lado, al tratarse de una evaluación de distintas técnicas y arquitecturas de aprendizaje profundo, el trabajo constituye un estudio principalmente exploratorio respecto al modelamiento específico del sistema. En este sentido, si bien durante el desarrollo de la investigación se procura lograr los mejores resultados sobre cada modelo, no se busca superar métricas de pronóstico específicas o previamente definidas.

# Capítulo 2

## Antecedentes

### 2.1. Formulación del pronóstico fotovoltaico

El objetivo del pronóstico fotovoltaico es utilizar los datos observados durante el periodo previo al intervalo temporal que se desea pronosticar, este último denominado horizonte de pronóstico, para estimar la potencia eléctrica generada por el sistema en el periodo correspondiente. Dependiendo de la extensión del horizonte de pronóstico se puede distinguir entre pronóstico inmediato (*nowcasting*) para aplicaciones en el rango 1-15 minutos, pronóstico de corto plazo para estimaciones en el rango 1-6 horas, y pronóstico de largo plazo que opera con horizontes desde 1 día hasta semanas en el futuro.

Considerese que la información del sistema en un momento determinado  $t$  puede ser representada por un vector  $\mathcal{X}_t$  de dimensión  $N$ , donde cada componente del vector corresponde a una variable o parámetro distinto. Si se tiene un registro periódico de estas  $N$  variables, entonces se cuenta con una serie de vectores  $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_t$ . Luego, el problema de pronóstico puede ser formulado como la predicción de la secuencia temporal de tamaño  $K$  (horizonte de pronóstico) futura más probable, dada la información de las  $M$  observaciones previas. Sea  $\mathcal{P}_t$  la potencia generada por el sistema en el instante  $t$ , entonces lo anterior puede ser expresado como:

$$\hat{\mathcal{P}}_{t+1}, \dots, \hat{\mathcal{P}}_{t+K} = \arg \max p(\mathcal{P}_{t+1}, \dots, \mathcal{P}_{t+K} | \mathcal{X}_{t-M+1}, \mathcal{X}_{t-M+2}, \dots, \mathcal{X}_t) \quad (2.1)$$

Donde se utiliza la notación  $\hat{\mathcal{P}}_t$  para indicar que se trata de una estimación, a diferencia de los valores registrados mediante instrumentación en terreno.

### 2.2. Irradiación solar

Si bien no es la única, sin lugar a dudas, la irradiación solar es la variable más relevante en cuanto a producción fotovoltaica, pues la potencia eléctrica que entrega cada celda fotovoltaica depende directamente del nivel de irradiancia que éstas reciben en cada momento.

La irradiancia solar en la superficie terrestre, que se define como la cantidad de energía, por unidad de tiempo, que incide sobre esta ( $W/m^2$ ), depende fundamentalmente de dos factores: la posición relativa entre la Tierra y el Sol, y la condiciones atmosféricas y meteorológicas por sobre la superficie. Así, se distingue la irradiancia extraterrestre  $G_o$  y la

irradiancia en la superficie  $G$ .

La irradiancia extraterrestre, por su lado, corresponde a la tasa de radiación solar que incide sobre la atmosfera terrestre y que, por tanto, no atraviesa esta última. Si bien la radiación emitida por el Sol no es del todo constante, sus fluctuaciones son lo suficientemente pequeñas como para ser despreciadas ( $<1.5\%$ ), por lo que en general se considera que esta depende únicamente de la posición relativa entre la Tierra y el Sol. Así, la irradiancia extraterrestre correspondiente al día  $n$  del año (1-365) puede ser estimada mediante la siguiente expresión:

$$B = (n - 1) \frac{360}{365} \quad (2.2)$$

$$G_o(n) = G_{sc} (1.000110 + 0.034221 \cos B + 0.00128 \sin B + 0.000719 \cos 2B + 0.000077 \sin 2B) \quad (2.3)$$

Donde  $G_{sc}$  esponde a la constante solar, que consiste en la irradiancia solar que incide perpendicularmente sobre la atmosfera cuando la Tierra y el Sol se encuentran a una distancia promedio respecto a la órbita, y equivale a unos  $1367 \text{ W/m}^2$ . Luego, si se quisiese calcular la irradiación extraterrestre sobre una región y a una hora en particular,

$$G_{o,loc} = G_o(n) (\cos \delta \cos \varphi \cos \omega + \sin \delta \sin \varphi) \quad (2.4)$$

Donde  $\delta$  es la declinación solar del día correspondiente al día  $n$  el año,  $\varphi$  es la latitud de la región, y  $\omega$  corresponde al ángulo horario solar relativo a la hora para la que se desea calcular la irradiancia. Si se posee la hora solar correspondiente  $t_{solar}$ ,  $\omega$  puede determinarse mediante,

$$\omega = (t_{solar} - 12) \cdot 15^\circ \quad (2.5)$$

Luego, por otro lado, se tiene la irradiancia en la superficie  $G$ , que corresponde a la tasa de radiación solar que incide en la superficie terrestre luego de haber atravesado la atmosfera. Así, esta irradiancia, producto de la absorción y dispersión que ocurre en su paso por la atmosfera, es menor a la extraterrestre correspondiente, donde el grado de atenuación depende tanto de las características atmosféricas y meteorológicas de la región (nubosidad, concentración de aerosoles y gases de efecto invernadero, nivel de particulado, etc.), como de la elevación del Sol, pues mientras menor sea, más atmosfera deberá atravesar la irradiación [7]. En la Figura 2.1 se presenta la comparación entre el espectro de la irradiación incidente en la superficie y la extraterrestre, donde se observa el efecto de la atenuación atmosférica incluso en un día soleado.

Es por esta razón que el forecasting de la irradiación solar resulta en un problema tan complejo desde el punto de vista técnico. Dado que son múltiples las variables y dinámicas que influyen en la atenuación atmosférica de la irradiación (el movimiento y el grosor de las nubes, la concentración del particulado, la contaminación atmosférica, la concentración de vapor de agua en el aire, etc.), la diferencia entre  $G_o$  y  $G$  suele seguir un comportamiento sumamente variable. En la Figura 2.2, por ejemplo, se presentan las radiaciones registradas y calculadas entre el 14 y 17 de Septiembre del 2018, en la estación solarimétrica de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile. En este gráfico se aprecia que en cada día

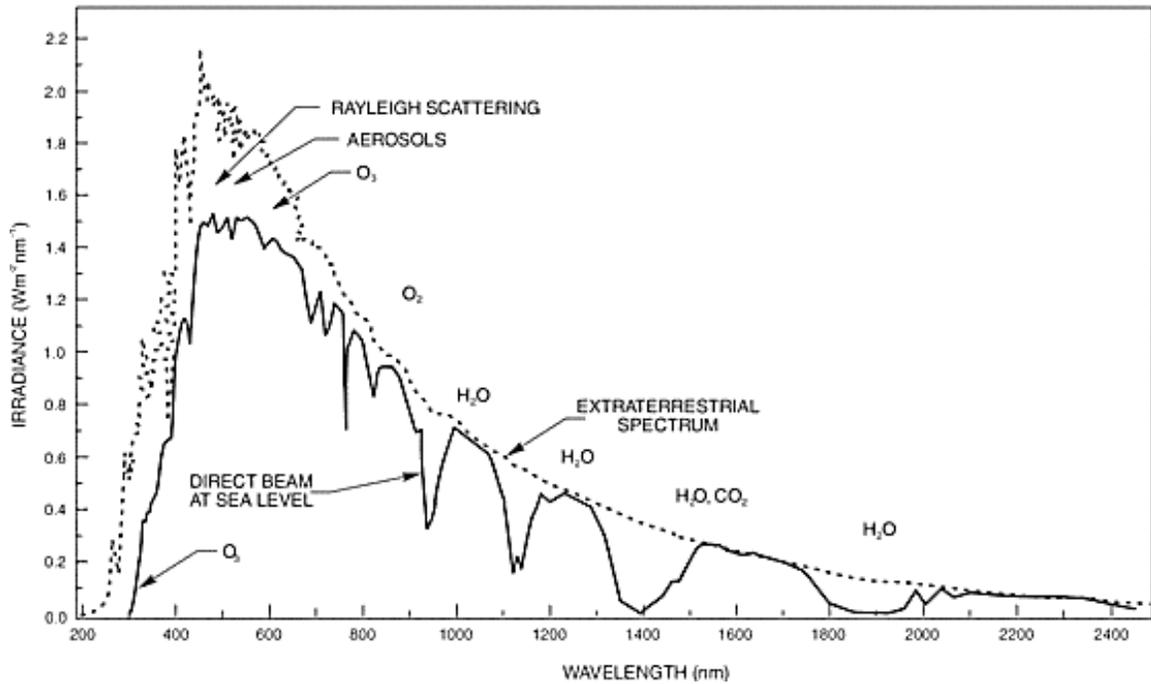


Figura 2.1: Comparación entre el espectro de la irradiancia incidente en la superficie terrestre (a nivel del mar) y la extraterrestre, en un día soleado [7].

la irradiancia de superficie sigue un comportamiento distinto de acuerdo a las condiciones meteorológicas respectivas: el primer día corresponde a un día soleado, el segundo a un día nublado, el tercer a un día soleado con nubosidad parcial en la tarde, y el cuarto a un día con nubosidad parcial. No obstante, si bien es clara la correlación entre el comportamiento de la irradiancia durante el día con las condiciones meteorológicas generales, esto no es suficiente para predecir por ejemplo las fluctuaciones que se observan en el cuarto día, y que serían de relevancia para un operador de una planta PV, por lo que se requiere de un modelo bastante más complejo.

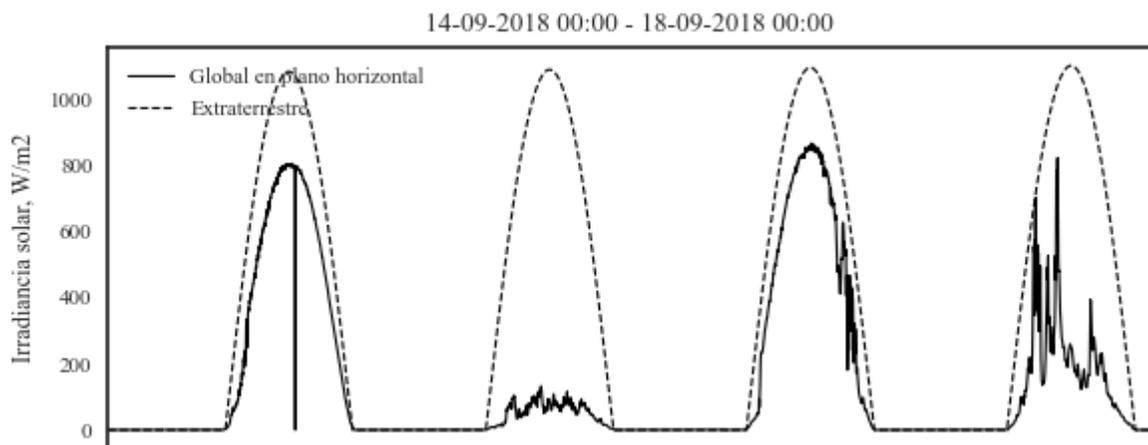


Figura 2.2: Comparación entre las irradiancias en superficie horizontal registradas entre el 14 y 17 de Septiembre del 2018 y la extraterrestre correspondiente.

## 2.3. Modelos estocásticos

En esta sección se describen una serie de metodologías estocásticas empleadas para el pronóstico de generación fotovoltaica. Estas ofrecen una base de referencia para la comparación y evaluación de nuevas metodologías, permitiendo la identificación de las fortalezas y debilidades de estas últimas. Los dos primeros métodos a tratar consisten en modelos univariados o endógenos, donde el pronóstico se realiza únicamente en base a los valores pasados de la potencia generada. Por otro lado, el tercer método requiere de información y datos adicionales sobre los componentes de la irradiación solar (GHI y DI).

### 2.3.1. Autoregresión (AR)

Los modelos autoregresivos consisten en regresiones lineales sobre los valores pasados de la variable que se desea pronosticar. Generalmente, estos son útiles cuando existe una correlación entre el valor de la variable en un determinado tiempo  $t$  y los valores que la preceden [8]. La cantidad de valores pasados que son considerados en la regresión está dada por el parámetro  $p$ , también llamado *orden* de la autoregresión.

Matemáticamente, el pronóstico de un modelo AR de orden  $p$  sobre la serie de valores  $X$  está dado por la ecuación:

$$\hat{X}_i = \sum_{j=1}^p \phi_j X_{i-j} \quad (2.6)$$

Los coeficientes de ponderación  $\phi_j$  son determinados mediante la minimización de los errores cuadrados del modelo sobre un subconjunto de entrenamiento  $X_{train}$ .

### 2.3.2. k-Nearest-Neighbors (kNN)

Es uno de los algoritmos más sencillos dentro del machine-learning. Este método consiste en la clasificación de atributos o series de datos en base a la similitud entre la serie de interés y un conjunto de series de referencia [8]. En el caso de pronósticos, el modelo kNN estima los valores futuros de una serie de tiempo (sea, por ejemplo,  $\vec{x}$ ) a partir de las  $k$  series "más cercanas" dentro de un subconjunto de entrenamiento  $X_{train}$ . A su vez, este último está pareado a un subconjunto  $Y_{train}$  que contiene los valores futuros correspondientes a cada una de las series de entrenamiento  $\vec{X}_i$ .

$$X_{train} = \{\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{N-1}, \vec{X}_N\}, \quad Y_{train} = \{\vec{Y}_1, \vec{Y}_2, \dots, \vec{Y}_{N-1}, \vec{Y}_N\} \quad (2.7)$$

Las  $k$  series más similares pueden ser determinadas a partir de la distancia euclidiana  $e_i$  entre la serie de tiempo de interés  $\vec{x}$  y cada una de las series  $\vec{X}_i$  contenidas en el subconjunto de entrenamiento.

$$e_i = \sqrt{\sum_j (x_j - X_{ij})^2} \quad (2.8)$$

Una vez identificadas las  $k$  menores distancias  $e_i$ , el pronóstico de la serie de interés está

dado por el promedio de las  $k$  series correspondientes  $\vec{Y}_i$ .

$$\vec{\hat{y}} = \frac{1}{k} \sum_{i=1}^k \vec{Y}_i \quad (2.9)$$

### 2.3.3. Método de persistencia

Considerado la referencia estándar a la hora de determinar el *forecast skill* de alguna nueva metodología, el método de persistencia resulta en uno de los modelos más simples para el pronóstico de series de tiempo. En el caso particular de pronóstico solar, este modelo se basa en el supuesto de que las condiciones de irradiación solar se mantienen constantes durante todo el horizonte de pronóstico ( $K$ ). En particular, se supone que tanto el índice de claridad ( $I_g/I_o$ ) como la fracción difusa ( $I_d/I_g$ ) se mantienen constantes.

$$\begin{aligned} \frac{I_g}{I_o}(t) &= \frac{I_g}{I_o}(t + \Delta t \cdot i), \quad i = 1, \dots, K \\ \frac{I_d}{I_g}(t) &= \frac{I_d}{I_g}(t + \Delta t \cdot i), \quad i = 1, \dots, K \end{aligned} \quad (2.10)$$

Así, calculando la irradiación extraterrestre en los tiempos ( $t + \Delta t \cdot i$ ), las ecuaciones anteriores permiten calcular las componentes de la irradiación (difusa  $I_d$  y directa  $I_b$ ) en el horizonte de pronóstico. Luego, estimando la irradiación total incidente sobre el arreglo fotovoltaico ( $I_T$ ) en cada uno de los timesteps dentro del horizonte de pronóstico, la potencia generada en el instante ( $t + \Delta t \cdot i$ ) se calcula mediante la ponderación:

$$\hat{\mathcal{P}}(t + \Delta \cdot i) = \frac{I_T(t + \Delta t \cdot i)}{I_T(t)} \cdot \hat{\mathcal{P}}(t) \quad (2.11)$$

La irradiación ( $I_T$ ) es calculada siguiendo el modelo de cielo anisotrópico de Pérez, el cual se basa en un análisis más detallado de los tres componentes de la irradiación difusa: isotrópico, circumsolar y de horizonte [7]. Así, la irradiación total incidente sobre la instalación fotovoltaica está dada por la ecuación:

$$\begin{aligned} I_T &= I_{T,b} + I_{T,d,iso} + I_{T,d,cs} + I_{T,d,hz} + I_{T,refl} \\ I_T &= I_b R_b + I_d (1 - F_1) \left( \frac{1 + \cos \beta}{2} \right) + I_d F_1 \frac{a}{b} + I_d F_2 \sin \beta + I_g \rho_g \left( \frac{1 - \cos \beta}{2} \right) \end{aligned} \quad (2.12)$$

Donde  $R_b$  es un factor geométrico que corrige la irradiación directa para superficies inclinadas respecto del plano horizontal,  $F_1$  y  $F_2$  son coeficientes asociados a los componentes circumsolares y de horizonte de la irradiación (respectivamente),  $a$  y  $b$  son términos asociados al ángulo de incidencia del cono circumsolar, y  $\rho_g$  es un factor que representa la reflectividad del suelo.

## 2.4. Aprendizaje profundo

En los últimos años, técnicas relacionadas al Machine Learning (ML) han sido ampliamente aplicadas en una gran variedad de áreas e industrias para diversas aplicaciones. Entre estas técnicas se encuentra el uso de redes neuronales profundas mediante lo que se conoce como aprendizaje profundo o Deep Learning.

Una red neuronal profunda (DNN) consiste en un modelo matemático que se construye de la interconexión de múltiples capas de nodos o neuronas, organizadas de acuerdo a una arquitectura previamente definida. Estas conexiones, descritas por pesos numéricos, pueden ser ajustadas durante una fase de entrenamiento o aprendizaje de la red permitiendo el modelamiento y la computación de sistemas complejos y altamente no lineales. La principal ventaja del Deep Learning y las DNN radica en su capacidad de extraer y detectar características (features) en grandes volúmenes de datos automáticamente mediante entrenamiento tanto supervisado como no supervisado. Así, este tipo de redes neuronales profundas son capaces de decidir que transformaciones sobre los datos de un cierto dataset permiten obtener características útiles para, por ejemplo, clasificar y etiquetar datos con una mayor confiabilidad [9].

Por lo general, las arquitecturas de DNN se suelen componer por cuatro redes o arquitecturas fundamentales, existiendo también algunas variaciones e híbridos:

- Densely Connected Neural Networks.
- Convolutional Neural Networks (CNN).
- Recurrent Neural Networks (RNN).
- Recursive Neural Networks.

Desde de sus comienzos en 2006, el Deep Learning ha logrado superar la mayoría de los algoritmos convencionales de ML en cuanto a precisión (accuracy), requiriendo por lo demás de un mínimo esfuerzo manual en su entrenamiento. En particular, y de especial interés para este trabajo, ha demostrado ser una herramienta sumamente útil para aplicaciones de forecasting de irradiación solar, logrando modelos superiores a los desarrollados mediante métodos físicos numéricos (NWR) o de regresión tradicional, como lo son los de autoregresión (AR, ARMA, ARIMA).

Dado que estas técnicas tradicionales, al basarse en la aplicación de modelos físicos o en la simple regresión de datos históricos, no logran capturar por completo las dinámicas entre los múltiples factores y parámetros relacionados a la generación de energía fotovoltaica, la utilización de Deep Learning ha sentado un nuevo precedente en el estado del arte de esta línea de investigación.

### 2.4.1. Densely Connected Neural Networks

En este tipo de *Multilayer Feed-forward Neural Networks*, la red se compone de una capa de entrada (*input layer*), una o más capas ocultas (*hidden layers*), y una capa de salida (*output layer*). Cada capa se compone de uno o más nodos o neuronas artificiales como la

esquemática en la Figura 2.3. Cada neurona  $i$  posee una serie de pesos dados por el vector  $W_i$  con los cuales pondera los valores de entrada que recibe de los nodos de la capa anterior ( $A_i$ ). La salida (*output*) del nodo está dada por la ecuación vectorial:

$$a_i = f(W_i \cdot A_i + b) \quad (2.13)$$

En esta expresión,  $f$  corresponde a la función de activación que gobierna la salida de la neurona. Muchas de las transformaciones empleadas en redes neuronales (como por ejemplo, la función sigmoide y la tangente hiperbólica) activan la salida de la neurona dentro de un rango acotado ( $[0, 1]$  o  $[-1, 1]$ ) [9]. Por otro lado,  $b$  corresponde al sesgo (*bias*) asociado a la capa de la neurona. Este parámetro permite que nodos de la capa sigan siendo activados incluso si los valores de entrada  $A_i$  son nulos.

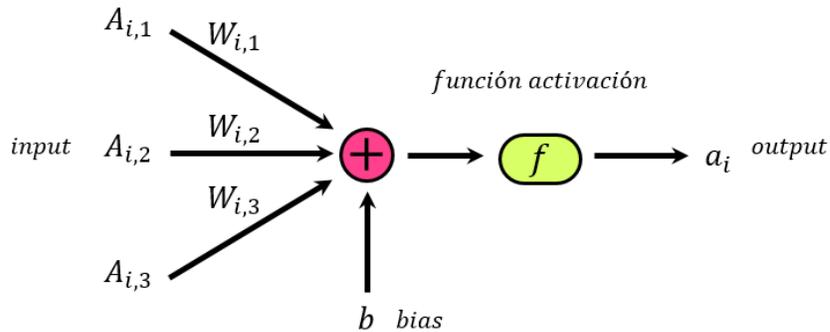


Figura 2.3: Esquema de una neurona artificial.

De este modo, en una red de capas Densely Connected (de ahora en adelante *Dense*) la información es procesada en una única dirección a medida que esta avanza desde la capa de entrada hasta la de salida.

## 2.4.2. Long Short-Term Memory RNN

Las redes LSTM consisten en una variación de las redes recurrentes, las cuales procesan secuencias o series temporales de datos de un modo en que no se altera la secuencialidad de la información. Su particularidad consiste en la incorporación de una célula interna de memoria y compuertas (*gates*) que añaden o eliminan información de esta célula, como se esquematiza en la Figura 2.4. Esta estructura permite que la información, a medida que avanza secuencialmente por la capa, sea retenida y transformada [9].

Así, cada nodo o unidad LSTM recibe de entrada información tanto de los nodos de la capa anterior ( $x_t$ ) como de las unidades correspondientes al *timestep* anterior en la secuencia ( $h_{t-1}$ ). Del mismo modo, la salida de las unidades LSTM es propagada hacia los nodos de la capa y del *timestep* siguiente.

Como se puede notar de la Figura 2.4, la activación de las unidades LSTM siguen una operación más compleja que las de una capa *Dense*. En primer lugar, la compuerta de entrada (*input gate*) pondera la información de entrada ( $x_t$  y  $h_t$ ) y mediante una activación sigmoide determina la cantidad de información que se añadirá a la célula de memoria ( $c_t$ ) con respecto

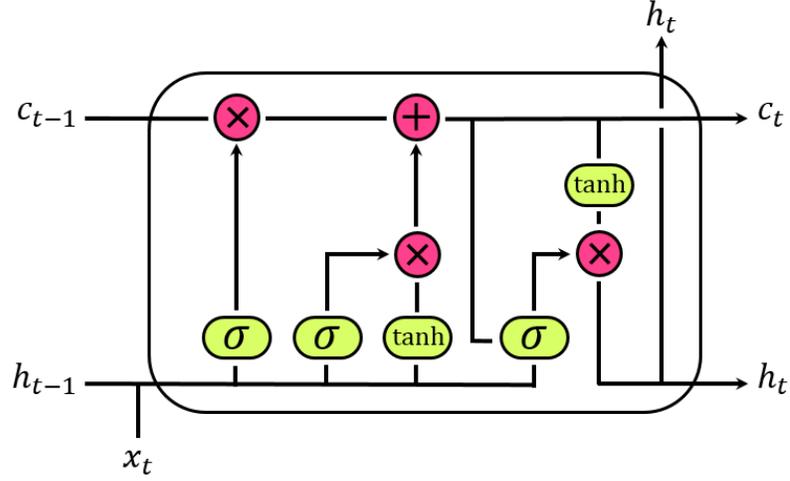


Figura 2.4: Esquema de una celda LSTM.

al estado anterior ( $c_{t-1}$ ) [10].

$$i_t = \sigma(U_i \cdot h_{t-1} + W_i \cdot x_t + b_i) \quad (2.14)$$

Análogamente, la compuerta de olvido (*forget gate*) pondera y determina la cantidad de información que se restará de la célula de memoria.

$$f_t = \sigma(U_f \cdot h_{t-1} + W_f \cdot x_t + b_f) \quad (2.15)$$

Así, la célula de memoria  $c_t$  es actualizada mediante la siguiente expresión:

$$c_t = f_t c_{t-1} + i_t \tanh(U_c \cdot h_{t-1} + W_c \cdot x_t + b_c) \quad (2.16)$$

Finalmente, la compuerta de salida (*output gate*) determina la amplificación o reducción de la información en la célula de memoria como salida de la unidad  $h_t$ :

$$o_t = \sigma(U_o \cdot h_{t-1} + W_o \cdot x_t + b_o) \quad (2.17)$$

$$h_t = o_t \tanh c_t \quad (2.18)$$

### 2.4.3. Just Another Network

Al igual que la unidad LSTM es una variación de una RNN, la unidad JANet es una modificación de una LSTM. En esta, las compuertas de entrada y salida son eliminadas de la celda, manteniendo únicamente la compuerta de olvido. Así, la activación de la salida  $h_t$  de la celda JANet está dada por:

$$\begin{aligned} f_t &= \sigma(U_f \cdot h_{t-1} + W_f \cdot x_t + b_f) \\ s_t &= \sigma(U_s \cdot h_{t-1} + W_s \cdot x_t + b_s - \beta) \\ c_t &= f_t c_{t-1} + (1 - s_t) \tanh(U_c \cdot h_{t-1} + W_c \cdot x_t + b_c) \\ h_t &= c_t \end{aligned} \quad (2.19)$$

Donde  $\beta$  es un parámetro que permite que en la célula se acumule un poco más de

información que la olvidada. Por lo general,  $\beta = 1$  provee los mejores resultados [11].

Producto de que con estas modificaciones el esquema estándar de inicialización de los pesos  $b_f$  provocaría que el gradiente de aprendizaje decaiga rápidamente, un nuevo método de inicialización es implementado. La **crono-inialización** inializa, valga la redundancia, el *bias* de la compuerta de olvido mediante:

$$b_f \sim \log(\mathcal{U}[1, T_{max} - 1]) \quad (2.20)$$

Donde  $T_{max}$  es el valor esperado de dependencia *long-term* en las secuencias procesadas. Así, las celdas JANet son capaces de retener información incluso después de varias entradas  $(U_f \cdot h_{t-1} + W_f \cdot x_t)$  nulas [11].

#### 2.4.4. Convolutional JANet

Las redes ConvJANet son una extensión de la arquitectura JANet, incorporando estructuras convolucionales en la entrada de la celda del mismo modo que las ConvLSTM. La intención de esta modificación es posibilitar la interpretación de patrones y características espacio-temporales en secuencias temporales de imágenes digitales. De este modo, tanto las entradas  $\mathcal{X}_t$ , como los estados ocultos  $\mathcal{C}_t$  y  $\mathcal{H}_t$  consisten en tensores 3D o 2D en el caso de imágenes de un solo canal [12].

Las ConvJANet actualizan sus células de memoria  $\mathcal{C}_t$  a partir de la información que reciben de la capa anterior ( $\mathcal{X}_t$ ) y del estado del *timestep* anterior en la secuencia ( $\mathcal{H}_{t-1}$ ,  $\mathcal{C}_{t-1}$ ) del mismo modo que las celdas JANet. No obstante, el producto punto es reemplazado por el operador de convolución (\*), y la multiplicación escalar por la multiplicación matricial Hadamard ( $\circ$ ).

$$\begin{aligned} f_t &= (U_f * \mathcal{H}_{t-1} + W_f * \mathcal{X}_t + b_f) \\ s_t &= (U_f * \mathcal{H}_{t-1} + W_f * \mathcal{X}_t + b_f - \beta) \\ \mathcal{C}_t &= f_t \circ \mathcal{C}_{t-1} + (1 - s_t) \circ \tanh(U_c * \mathcal{H}_{t-1} + W_c * \mathcal{X}_t + b_c) \\ \mathcal{H}_t &= \mathcal{C}_t \end{aligned} \quad (2.21)$$

Un esquema de la estructura interna de las redes ConvJANet se presenta en la Figura 2.5, en este caso los kernel  $U_f$  y  $W_f$  son tensores de 3x3. En el caso de la interpretación de objetos en movimiento, kernels de mayor dimensión debiesen ser capaces de capturar movimientos de mayor velocidad [12].

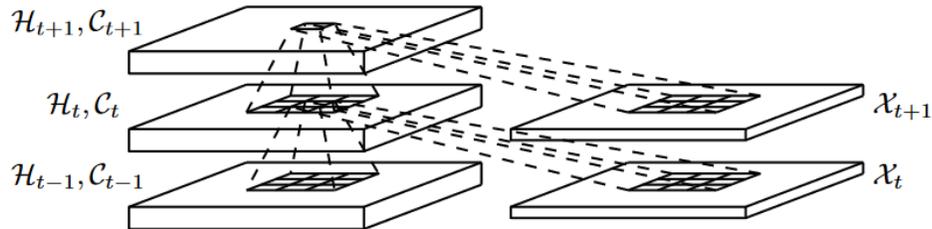


Figura 2.5: Esquema de una celda ConvJANet [12]

# Capítulo 3

## Revisión bibliográfica

El desarrollo de nuevas técnicas y herramientas para el pronóstico de energía solar ha sido un tema extensamente estudiado tanto en la industria energética como en la academia. En esta sección, se presenta una revisión de una serie de estudios relacionados a la aplicación de técnicas y arquitecturas de aprendizaje profundo para el pronóstico de bien, el recurso solar disponible en una determinada locación o bien, de la generación eléctrica de una planta fotovoltaica en particular. De este modo, esta selección de trabajos permite contextualizar de mejor manera los métodos y recursos empleados en el desarrollo de este tipo de modelos de pronóstico, así como también la diversidad de arquitecturas de redes neuronales empleadas con este propósito, cada cual con sus características y funcionalidades particulares.

### 3.1. Redes ANN para pronóstico fotovoltaico

Considerando la gran cantidad de variables, tanto operacionales como atmosféricas, que influyen en la producción energética de los sistema fotovoltaicos, la capacidad de modelamiento que introducen las redes neuronales (ANN) se adecua bastante bien como método de extracción de patrones y estimación de pronósticos mediante regresiones sobre las variables. De este modo, numerosos estudios han explorado el desempeño de arquitecturas ANN para el modelamiento y estimación de pronósticos solares, con la finalidad de desarrollar modelos cada vez más robustos y con una mayor confiabilidad y precisión.

Wang et. al. [2], en vez de trabajar directamente con los datos operacionales del sistema, este estudia la incorporación de parámetros estadísticos en los datos de entrada de la red, obtenidos mediante el preprocesamiento y transformación de los datos históricos, con el fin de facilitar la extracción de características relevantes para el pronóstico. El modelo propuesto por Wang (ANN-SFP) genera a partir de esta serie de parámetros correspondientes a un día  $n$ , la curva de irradiancia horaria del día siguiente ( $n + 1$ ). De acuerdo a los resultados obtenidos en la investigación, el modelo propuesto sigue de manera más efectiva las fluctuaciones y variaciones de la irradiancia solar registrada, logrando una reducción significativa en el error de pronóstico respecto a otros métodos convencionales. En este sentido, se destaca que si bien las redes ANN poseen una gran capacidad de extracción de patrones y características sobre los conjuntos de datos, un preprocesamiento sobre los datos de entrada puede mejorar el entrenamiento y posterior desempeño de la red.

Por otro lado, Chu et. al. [13] emplea redes Densely Connected para mejorar el desempeño de tres modelos estocásticos (Cloud Tracking Model, ARMA y KNN) para el pronóstico de potencia de una planta fotovoltaica. De este modo, los modelos ANNs propuestos utilizan las predicciones generadas por estos métodos estocásticos, en conjunto con los datos de generación de la media hora anterior, para estimar el pronóstico de generación de los próximos 15 min. Este procedimiento es denominado *reforecasting*, pues consiste en una corrección de las estimaciones iniciales mediante el modelo introducido. De los experimentos realizado, se observa que las redes ANNs mejoran significativamente el error de pronóstico en los tres métodos analizados, evidenciándose una reducción en el bias de la distribución del error y en la ocurrencia de errores de mayor magnitud. En este sentido, es posible complementar el modelamiento de arquitecturas de redes neuronales con información o parámetros adicionales obtenidos mediante procedimientos numéricos estocásticos y/o determinísticos.

## 3.2. Redes LSTM para pronóstico fotovoltaico

Dado el acelerado crecimiento que ha experimentado el campo del aprendizaje profundo en la última década, evidenciado tanto en el desarrollo de nuevas técnicas y arquitecturas como también en una mayor integración en numerosas áreas de investigación, estructuras como la Long-Short Term Memory (LSTM) han sido recientemente empleadas en la literatura en el desarrollo de nuevas técnicas de pronóstico solar [14, 5].

Como se describió en el capítulo anterior, las estructuras LSTM tienen la capacidad de modelar y reconocer patrones temporales a partir de secuencias de datos, gracias a sus celdas de memoria y la forma en que transmiten información entre sus unidades. De este modo, estas son capaces de procesar secuencias y registros de datos operacionales obtenidos mediante sensores u otro tipo instrumentación en la planta y así extraer información temporal que permita reducir el error de pronóstico.

Abdel-Nasser et. al. [6] presenta una implementación simple y directa de redes LSTM para el pronóstico en corto plazo de generación de electricidad de un sistema fotovoltaico a partir de la generación registrada en las horas previas. Para analizar el desempeño de este tipo de estructura y su dependencia con la secuencialidad de la información, cinco arquitecturas distintas respecto a la estructuración de los datos de entrada son evaluadas respecto a los mismos conjuntos de datos. El modelo resultante consiste en una red LSTM que procesa una ventana temporal de datos previos para el computo de la generación en la hora siguiente. Este modelo, al ser comparado con otros métodos de pronóstico tradicionales (Multiple Linear Regression, Bagged Regression Trees y Densely Connected Networks), este presentó métricas de error considerablemente menores [6].

Por otro lado, en una integración más compleja de estructuras LSTM, Lee et. al. [5] propone una implementación híbrida de redes convolucionales y LSTM para el procesamiento de series temporales de registros operacionales y con ello estimar la generación de energía del día siguiente. Si bien el procesamiento de la secuencialidad de la información mediante redes recurrentes es crucial para un mejor desempeño en el pronóstico de generación, Lee et. al. propone que la incorporación de capas CNN para el preprocesamiento de los datos puede

permitir la extracción de patrones que resulten de mayor utilidad a la hora de estimar el pronóstico de generación.

El modelo desarrollado por Lee et. al [5] consiste en dos redes CNN independientes (D-CNN), cada una con un distinto tamaño de filtro con el fin de extraer patrones y características de distinta escala a partir de una ventana temporal de las secuencias de datos operacionales registrados por los sensores del inversor del sistema. Cada una de estas redes preprocesa la información del sistema, para luego concatenar los vectores resultantes generando un solo vector latente (*feature vector*) el cual es posteriormente entregado a la red LSTM. De este modo, a medida que la ventana temporal va siendo desplazada a lo largo del registro de datos, la red D-CNN va generando una serie de vectores latentes. Finalmente, esta serie de vectores es procesada secuencialmente mediante una red LSTM cuya última capa latente es entregada a una capa Densely Connected para computar el pronóstico de generación.

Al implementar esta arquitectura sobre el registro de un sistema fotovoltaico los resultados mostraron que el modelo no solo supera los pronósticos obtenidos mediante métodos tradicionales, sino que también supera otras arquitecturas de aprendizaje profundo en el estado del arte de la literatura, aunque en menor medida. Este estudio revela que un preprocesamiento sobre las secuencias de datos del sistema contribuye a la extracción de patrones y características más relevantes para la estimación de generación fotovoltaica.

### 3.3. Imágenes satelitales para pronóstico solar

Si bien la utilización de datos y registros satelitales de irradiación solar en superficie (SSI) para la estimación de pronósticos solares ha sido extensamente estudiada tanto en la industria como en la academia, hasta el momento la mayoría de estos modelos consisten en métodos empíricos y numéricos basados en el modelamiento físico de la dinámica detrás de la transmisión de la irradiación a través de la atmósfera terrestre [3]. Algunos de estos métodos y modelos son el Heliosat-4, desarrollado por el MINES ParisTech y el DLR; el algoritmo Cloud Physical Properties (CPP), desarrollado por el KNMI; y el Physical Solar Model, utilizado por el NREL para la generación de sus bases de datos.

En este sentido, la integración de este tipo de datos con arquitecturas de aprendizaje profundo es escasa y no ha sido estudiada en profundidad en la literatura actual. No obstante, la calidad de información espacio-temporal que contienen las imágenes satelitales geoestacionarias presenta un gran potencial para el desarrollo de nuevos modelos basados en redes neuronales. Dado que los satélites geoestacionarios permanecen en todo momento sobre el mismo punto en la superficie terrestre, siguiendo la rotación del planeta, el registro periódico (5-15 min) que estos realizan genera finalmente una serie temporal de datos espaciales a partir de la cual es posible extraer información sobre la evolución de la nubosidad en una determinada zona [15]. De este modo, es posible complementar los datos registrados mediante instrumentación en terreno con información acerca de la vecindad del sitio.

Mazorra et. al. [4] experimenta con la incorporación de datos derivados de registros satelitales (Figura 3.1) en modelos de redes neuronales para el pronóstico de irradiación solar

(GHI) en el corto plazo en las Islas Canarias. Adicionalmente al registro temporal de datos obtenido mediante instrumentación solarimétrica en terreno, 30 valores, o más bien atributos, extraídos de la secuencia de imágenes satelitales son agregados a la base de datos. Esta serie de valores consisten en la intensidad de los pixeles que presentan una mayor correlación con los registros en terreno como se presenta en la Figura 3.2. De este modo, Mazorra et. al. propone extender la información disponible del sitio estudiado con información acerca de la nubosidad en el entorno próximo tanto espacial como temporalmente.

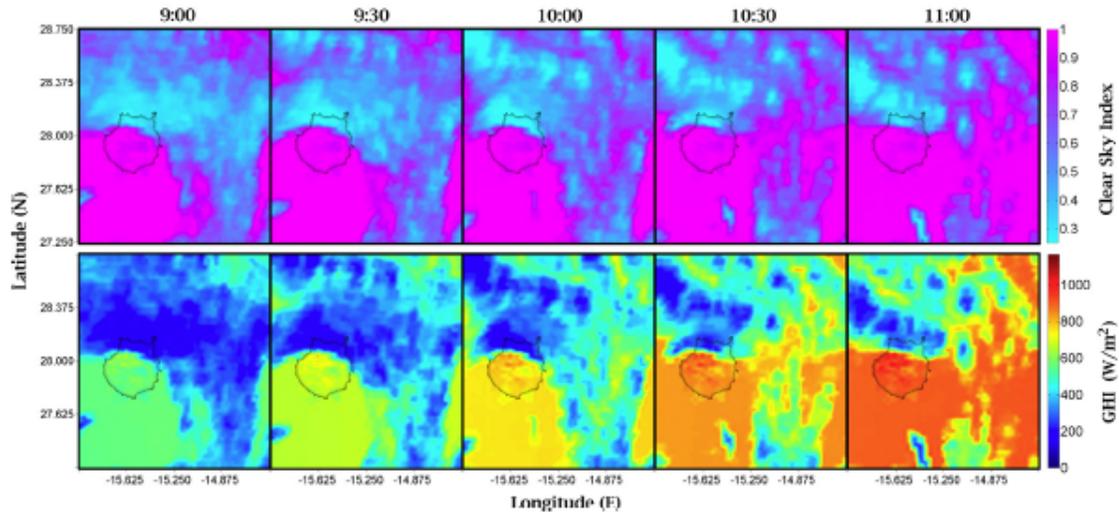


Figura 3.1: Registro temporal de imágenes satelitales sobre el área estudiada [4].

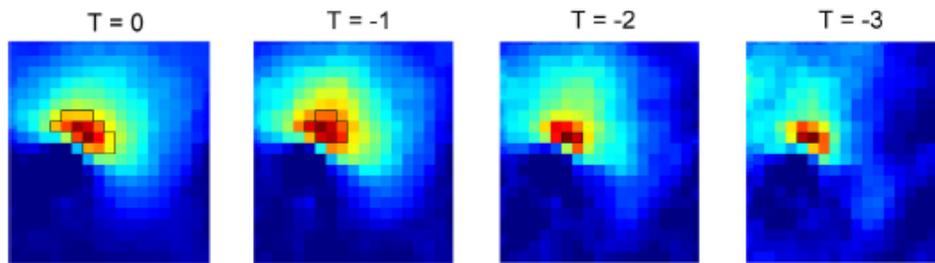


Figura 3.2: Selección de pixeles con mayor correlación dentro de la secuencia de imágenes satelitales con los datos registrados en terreno [4].

A partir de esta combinación de datos, una red neuronal Bayesiana (BNN) es entrenada y evaluada. De los resultados se desprende que la incorporación de los datos derivados del registro satelital permite reducir el error de pronóstico de la irradiación global en el corto plazo, no obstante, la selección de pixeles depende considerablemente de la estacionalidad de los datos y, por tanto, existe un factor temporal que debe ser considerando en la distribución de pixeles a incorporar en la base de datos [4]. De este modo, la implementación de arquitecturas ConvLSTM o ConvJANet presenta el potencial de procesar de manera más robusta la secuencialidad de la información contenida en el registro temporal de imágenes satelitales.

# Capítulo 4

## Metodología

### 4.1. Datos y recursos disponibles

En esta sección se presentan los datos y recursos utilizados en el desarrollo de este trabajo. Estos recursos incluyen desde mediciones y registros obtenidos mediante instrumentación en terreno, hasta imágenes satelitales geoestacionarias del espectro de radiación asociado a la nubosidad de la región.

#### 4.1.1. Sistema fotovoltaico

Esta investigación es realizada utilizando los datos registrados por el arreglo fotovoltaico de 124 módulos semitransparentes Nexpower Lucid Plus NB-130AP ubicado en la torre oriente del complejo Beauchef 851 de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile ( $33.45775^{\circ}\text{S}$ ,  $70.66466^{\circ}\text{W}$ ). Esta instalación, con una potencia nominal de 16.12 kW (CC), está compuesta por dos subarreglos dispuestos en dos aguas: ambos con una inclinación de  $16^{\circ}$  respecto al plano horizontal, pero orientados en direcciones Oriente-Poniente (acimut:  $-90^{\circ}$  y  $90^{\circ}$ , respectivamente) (etiqueta A en la Figura 4.1).

Cada subarreglo de la instalación posee las siguientes especificaciones técnicas:

- 14 strings de 3 módulos cada uno. Características eléctricas por string: tensión MPP 371.7 V, corriente MPP 1.05 A.
- 5 strings de 4 módulos cada uno. Características eléctricas por string: tensión MPP 495.6 V, corriente MPP 1.05 A.
- 1 inversor SMA Sunny Mini Central 5000 A-11. Características eléctricas: input 371.7 V (UMPP), output 220 Volt - 5.0 kW.
- 1 inversor SMA Sunny Boy 2500 HF. Características eléctricas: input 495.6 V (UMPP), output 220 Volt - 2.5 kW.

Los datos de la instalación son registrados y almacenados por un equipo de monitorización SunnyWebBox de la empresa SMA y corresponden a promedios, cada 15 minutos, de la potencia de salida de cada uno de los inversores. Adicionalmente, mediante un equipo de medición Sunny SensorBox, se registra la irradiación solar incidente sobre el arreglo y la



Figura 4.1: Imágen satelital del sistema fotovoltaico de 16.12 kW estudiado. La etiqueta A corresponde a la ubicación de instalación fotovoltaica, mientras que la etiqueta B corresponde a la ubicación de la estación solarimétrica.

temperatura en superficie de los paneles.

El periodo de tiempo analizado en este trabajo cubre desde el 27 de Agosto del 2018 hasta el 6 de Septiembre del 2019 (376 días). En la Figura 4.2 se presentan los registros de potencia generada por el sistema durante todo el periodo señalado, mientras que en la Figura 4.3 se presenta el registro de temperatura en superficie.

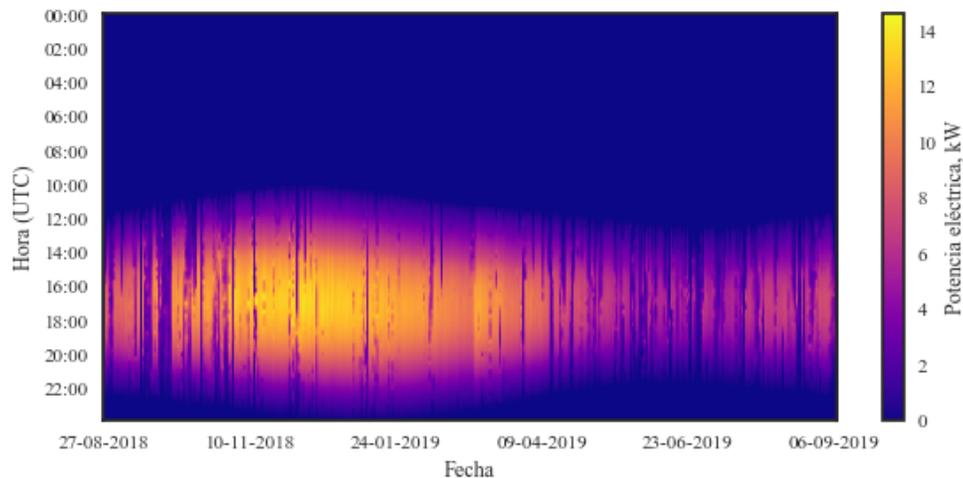


Figura 4.2: Gráfico de potencia generada por el sistema fotovoltaico (todos los inversores) durante el periodo analizado.

#### 4.1.2. Estación solarimétrica

Adicionalmente, se cuenta con mediciones cada 1 minuto de los siguientes componentes de la irradiación solar: irradiación global en el plano horizontal (GHI), irradiación difusa (DFI) e irradiación directa (DNI). Estos datos son registrados por la instrumentación de la estación solarimétrica de la Facultad de Ciencias Físicas y Matemáticas ubicada aproxima-

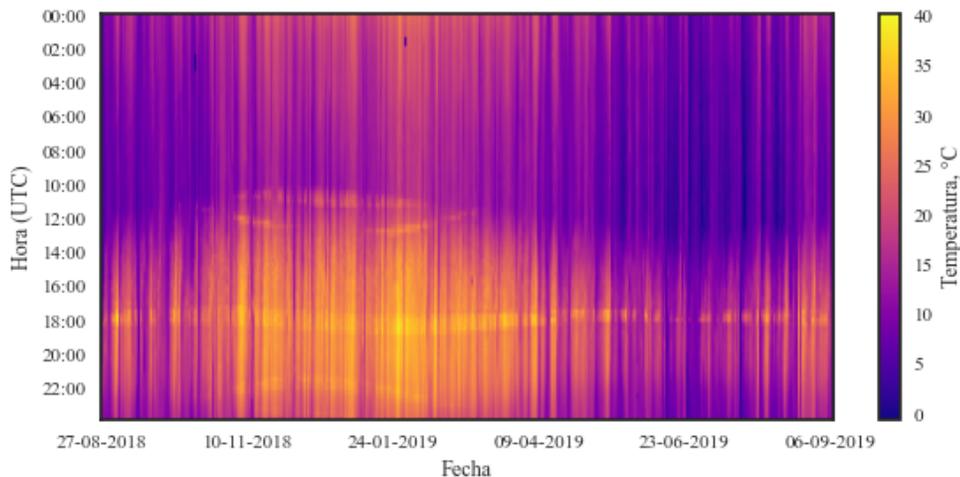


Figura 4.3: Gráfico de temperatura en superficie del sistema fotovoltaico durante el periodo analizado.

damente a 50 metros del sistema fotovoltaico, en la torre poniente del complejo Beauchef 851 (etiqueta B en la Figura 4.1). La estación cuenta con la siguiente instrumentación: Piranómetro Kipp&Zonnen CMP21, Pirgeómetro Kipp&Zonnen CGR4, Pirheliómetro Kipp&Zonnen CHP1, SunTracker Kipp&Zonnen SOLYS2 y un capturador de datos Campbell Scientific CR1000. A modo de ejemplo, en la Figura 4.4 se presenta el registro de la irradiación GHI durante el periodo analizado. Esta estación cuenta además con un registro de la temperatura ambiente mediante un Campbell Scientific HMP60 Temperature and Relative Humidity Probe y un capturador de datos Campbell Scientific CR6 Measurement and Control Center. El registro de temperatura ambiente se presenta en la Figura 4.5.

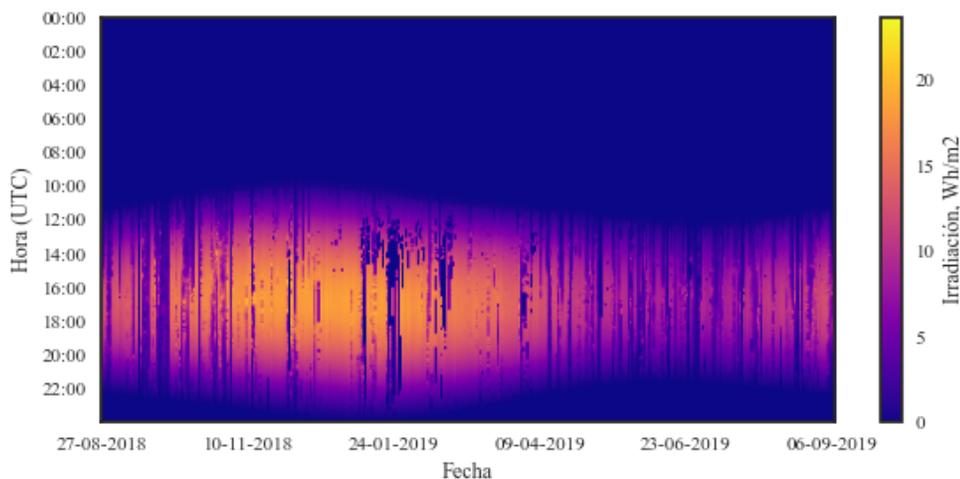


Figura 4.4: Gráfico de irradiación GHI ( $Wh/m^2$ ) registrado por la estación solarimétrica.

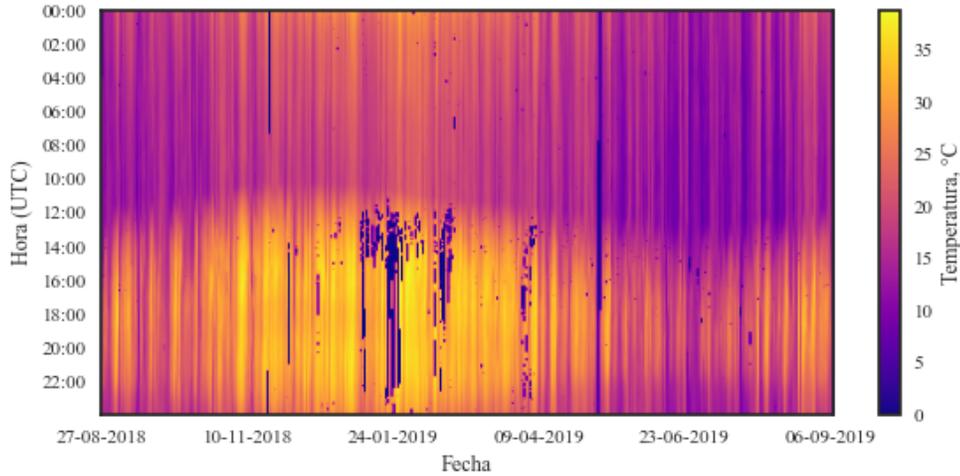


Figura 4.5: Gráfico de temperatura ambiente durante el periodo analizado.

### 4.1.3. Imágenes satelitales NOAA

Para incorporar a los modelos información sobre la nubosidad en el entorno espacial del sitio fotovoltaico se utiliza la base de datos de imágenes satelitales de la Administración Nacional Oceánica y Atmosférica (NOAA) que se encuentra disponible en un Amazon Web Storage (AWS). Los datos empleados corresponden al registro temporal del Advanced Baseline Imager (ABI) montado sobre el Geostationary Operational Environmental Satellite (GOES-16). Este satélite, que inició su operación en Diciembre del 2017, provee continuamente de información meteorológica de todo el disco planetario en sus 16 bandas espectrales. La resolución temporal de escaneo de las imágenes es de 15 minutos desde el inicio de su operación y 10 minutos desde que incorporó el modo 6 de escaneo en Abril del 2019 [15].

La banda utilizada en esta investigación corresponde a la cuarta banda espectral, denominada banda *Cirrus*, con una longitud de onda central de  $1.37 \mu m$ . Esta banda capta el espectro de radiación asociado fuertemente a la absorción producida por vapor de agua. La energía a  $1.37 \mu m$  es absorbida por el vapor de agua a medida que la radiación atraviesa la tropósfera. Así, esta banda logra captar delgadas nubes cirrus durante el día así como también polvo y nubes a baja altura [15]. La resolución espacial de este tipo de imágenes es de 2 km por pixel. La información de esta banda espectral se presenta en la Tabla 4.1.

Tabla 4.1: Resumen técnico correspondiente a la banda espectral Cirrus del ABI.

Longitud de onda	Banda	Tipo	Resolución
$1.37 \mu m$	4	<i>near-Infrared</i>	$2.0 km$

## 4.2. Pre-procesamiento de datos

### 4.2.1. Exploración de datos

Con la finalidad de analizar preliminarmente la correlación entre las distintas variables disponibles en las bases de datos, se realizan una serie de gráficos exploratorios de referencia. Adicionalmente, para incorporar información temporal en los gráficos, los puntos son coloreados de acuerdo a la hora del día correspondiente.

En la Figura 4.6 se presenta un gráfico de la correlación entre la fracción difusa  $I_d/I_g$  versus el índice de claridad  $I_g/I_o$ . La distribución de esta correlación permite visualizar el espectro de calidad de cielo registrado por la instrumentación de la estación. Así, se observa que los datos se concentran en dos cúmulos de baja y alta fracción difusa, donde en el rango 0.4 – 0.8 existe una mayor dispersión. Adicionalmente, en cada uno de estos cúmulos se distingue una segregación que depende significativamente de la hora del registro. En particular, existen distribuciones correspondientes a datos de la mañana (10:00 - 14:00 UTC) y del día/tarde (16:00 - 00:00 UTC). Por lo tanto, es claro que la fracción difusa no es dependiente únicamente del índice de claridad.

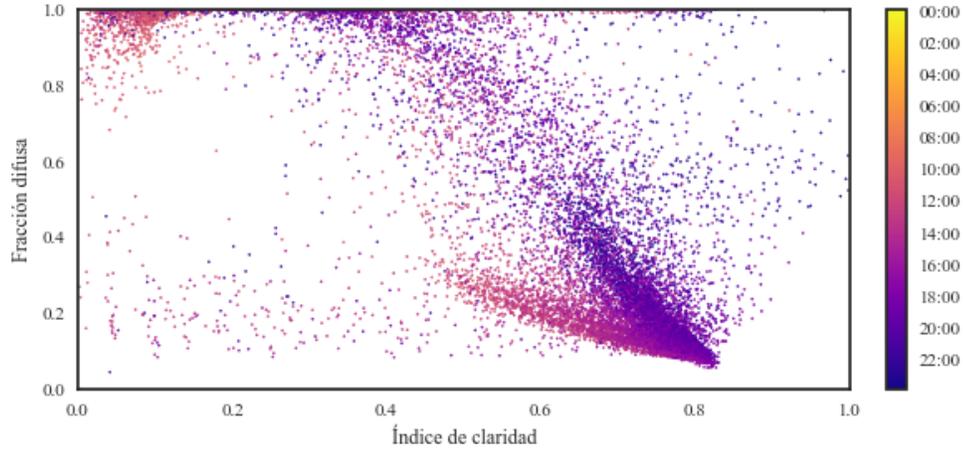


Figura 4.6: Gráfico de fracción difusa versus índice de claridad. Horas en UTC.

Para analizar, por otro lado, la correlación entre la potencia generada por la instalación fotovoltaica y los registros de irradiación solar se presentan las Figuras 4.8 y 4.7. La Figura 4.8 presenta la correlación entre la potencia generada por el arreglo fotovoltaico y la irradiación global en el plano horizontal ( $G_g$ ). Teóricamente, la relación entre ambas variables debiese ser lineal, no obstante, se observa que esta relación (si bien mantiene una evidente tendencia lineal) presenta una dispersión considerable. Este comportamiento se visualiza de mejor manera en la Figura 4.7, la cual consiste en el gráfico de *performance ratio* versus el índice de claridad. El *performance ratio* corresponde al cociente entre la potencia generada por el sistema en un determinado periodo y la que debiese generar de acuerdo a la potencia nominal de la instalación y el nivel de irradiación registrado. Así, el *performance ratio* indica la eficiencia fotovoltaica del sistema.

$$PR(t) = \frac{G_{STC}}{P_{STC}} \cdot \frac{P(t)}{G_g(t)} \quad (4.1)$$

Donde  $G_{STC}$  y  $P_{STC}$  corresponden a la irradiación y potencia del sistema en condiciones

estándares ( $G_{STC} = 1000W/m^2$ ,  $P_{STC} = 16.12kW$ ). En un sistema fotovoltaico ideal, el *performance ratio* debiese mantenerse constante a lo largo del rango del índice de claridad, no obstante, del gráfico se observa que existe una dispersión significativa. Incluso, se distinguen tendencias dependientes de la hora que se separan del cúmulo principal. Así, se visualiza que el pronóstico de la potencia generada no depende únicamente de la irradiación global, siendo en realidad un sistema bastante más complejo con más variables involucradas.

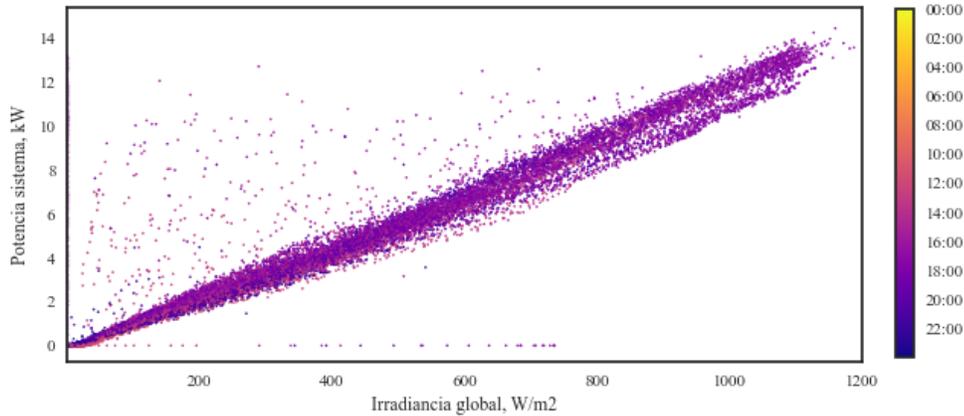


Figura 4.7: Gráfico de potencia generada versus irradiancia global en el plano horizontal. Horas en UTC.

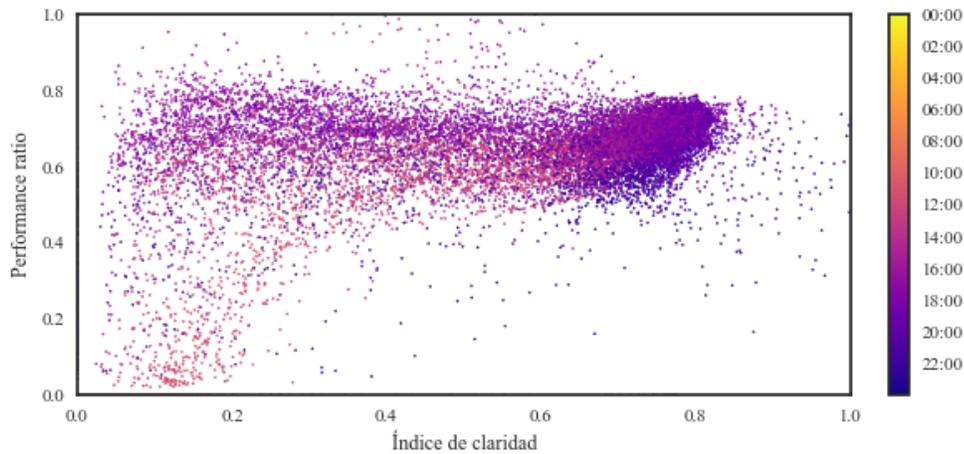


Figura 4.8: Gráfico de performance ratio versus índice de claridad. Horas en UTC.

### 4.3. Clustering para clasificación solar de días

Un enfoque para analizar e identificar patrones en los perfiles diarios de irradiación consiste en el uso de herramientas de *clustering*. Estas permiten segmentar y clasificar puntos dentro del espacio multivariable de los datos a partir de su cercanía o similitud con el resto de puntos [9]. Generar un método que permita clasificar consistentemente los tipos de días contenidos en el periodo analizado resulta de suma utilidad, pues permite realizar posteriormente una evaluación segmentada de los modelos desarrollados. Así, es posible obtener

resultados y métricas sobre cada uno de los clusters identificados, mejorando significativamente el análisis de desempeño de cada modelo [16].

Las variables y parámetros considerados en el clustering son los siguientes:

- **Índice de claridad:** Corresponde al cociente entre la irradiación global del día completo y la irradiación extraterrestre durante el mismo periodo,  $H_g/H_o$ .
- **Índice de nubosidad:** Podría considerarse el inverso de la fracción difusa. Se calcula a partir del cociente entre la irradiación difusa del día completo y la irradiación global durante el mismo periodo,  $1.0 - H_d/H_g$ .
- **Índice de suavidad:** Este parámetro cuantifica la suavidad de la curva de irradiación del día. Se calcula como el error cuadrado medio (MSE) entre la irradiación global del día y la convolución de la irradiación global con una gaussiana (perfil suavizado).

Se emplea el algoritmo de Agrupamiento Espacial Basado en Densidad de Aplicaciones con Ruido (DBSCAN) para realizar la clasificación de días sobre la serie de parámetros normalizados. Los parámetros seleccionados para el algoritmo son  $\epsilon = 0.09$  y  $min_{pts} = 9$ . Los resultados del *clustering* se presentan en las Figuras 4.9 y 4.10. En la primera se observa que la clasificación resulta en tres grupos, donde los índices de claridad y nubosidad son los más relevantes en la segregación. Por otro lado, en la segunda figura se distinguen los patrones o bien, los tipos de días, identificados. En la Tabla 4.2 se presenta un detalle de los clusters identificados por el algoritmo junto con el tiempo meteorológico asociado.

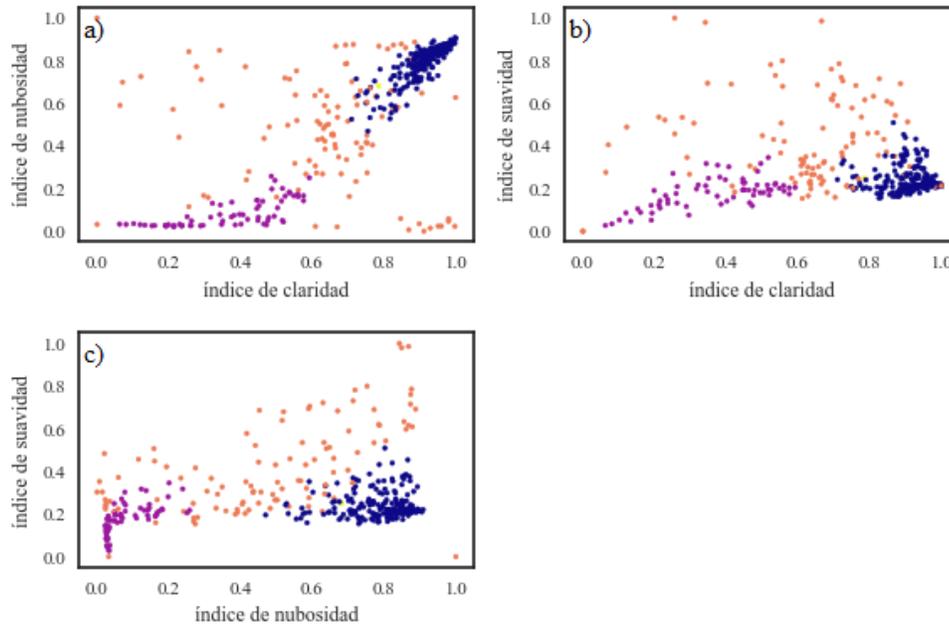


Figura 4.9: Gráfico del clustering realizado sobre los parámetros. (a) índice de nubosidad versus índice de claridad, (b) índice de suavidad versus índice de claridad y (c) índice de suavidad versus índice de nubosidad.

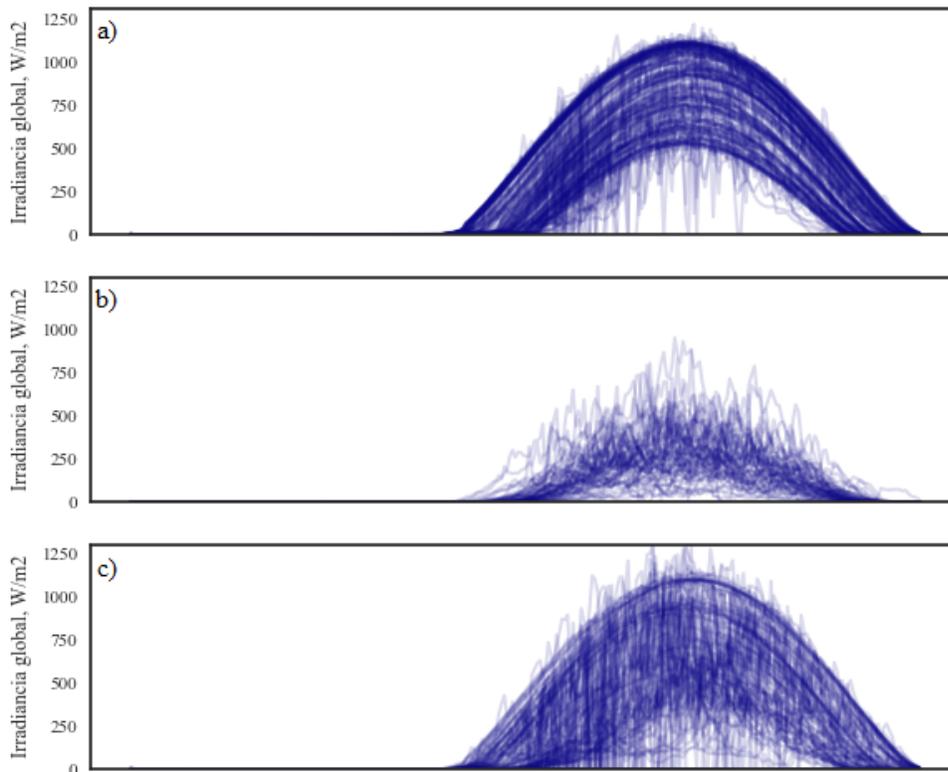


Figura 4.10: Gráfico de irradiancia global en los días correspondientes a cada cluster. (a) Cluster A, (b) Cluster B y (c) Cluster C.

Tabla 4.2: Características de los clusters identificados por el algoritmo DBSCAN.

Cluster	Tipo de día	Cantidad de días	Proporción
A	Soleados	205	54.5 %
B	Nublados	70	18.6 %
C	Nubosidad parcial	101	26.9 %

#### 4.4. Generación de conjuntos de datos de entrenamiento y validación

Para el desarrollo de las arquitecturas y modelos de redes neuronales, deben generarse dos conjuntos o sets de datos: uno de entrenamiento ( $X_{train}, Y_{train}$ ) con el cual son ajustados los pesos de la red, y otro de validación ( $X_{test}, Y_{test}$ ) con el cual se evalúa el desempeño y la generalización del modelo entrenado. Para que los resultados de la validación sean válidos y confiables, es importante que no exista ningún tipo de filtración de datos entre los dos sets (*data leak*), es decir, no pueden haber datos repetidos.

### 4.4.1. Datos del sistema y de irradiación

Para cumplir lo anterior, se implementa el siguiente algoritmo de generación de conjuntos de datos. En primer lugar, la base de datos es segmentada en periodos de 10 días continuos, utilizando una base horaria de 30 minutos y considerando todos los atributos (*features*) del registro. Luego, en cada uno de estos periodos, los primeros 8 días son procesados para el set de entrenamiento, mientras que los 2 días restantes son procesados para el set de validación (una muestra de esta segmentación se visualiza en la Figura 4.11).

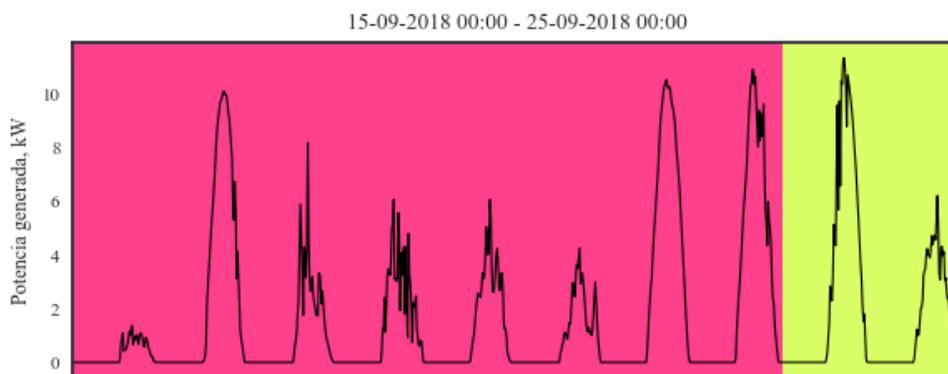


Figura 4.11: Muestra de procedimiento de segmentación de periodos para la generación de los conjuntos de datos de entrenamiento. Los datos sombreados de color morado son procesados para el set de entrenamiento, mientras que los sombreados de color amarillos son procesados para el set de validación. (Periodo entre el 15 y 25 de Septiembre del 2018).

Dado que los modelos LSTM/JANet reciben datos estructurados de la forma (*input timesteps, features*) y retornan los pronósticos de la forma (*output timesteps*), cada uno de los periodos producidos por la segmentación anteriormente descrita son procesados para generar estas ventanas de tiempo. Para esto, una ventana de tamaño (*input timesteps + output timesteps, features*) se va "deslizando" sobre el registro de datos (como se esquematiza en la Figura 4.12) para producir las ventanas de tiempo ( $x_i, y_i$ ) que se van agregando los conjuntos ( $X, Y$ ). La cantidad de *timesteps* en que se desliza la ventana está dado por el parámetro *overlap*.

En la Tabla 4.3 se presentan los parámetros seleccionados para la generación de los conjuntos de datos así como también los tamaños de los conjuntos resultantes. Por otro lado, en la Tabla 4.4 se presentan los porcentajes de tipos de días (tiempo meteorológico) contenidos en los conjuntos de entrenamiento y validación.

Tabla 4.3: Características de los conjuntos de datos generados.

Input	Output	Overlap	Features	Train samples	Test samples
12	6	1	13	13,542	2,923

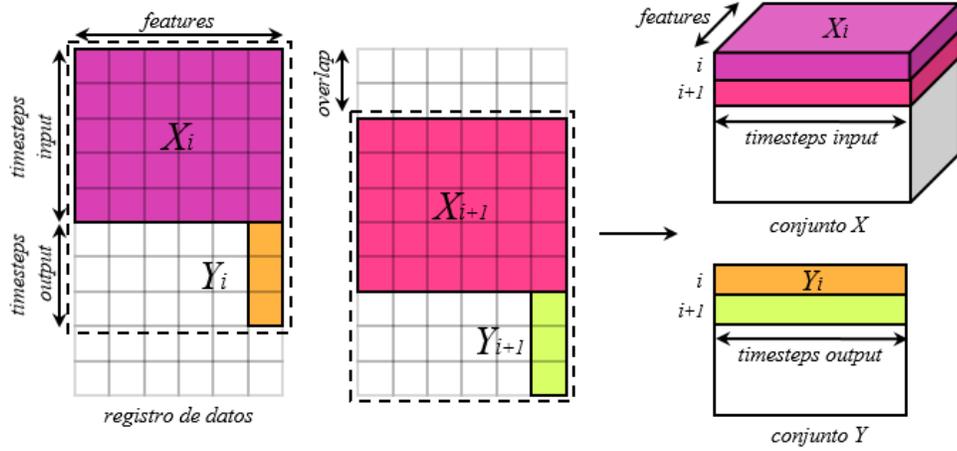


Figura 4.12: Muestra de procedimiento de generación de ventanas de datos para la creación de los conjuntos  $X$  e  $Y$ .

Tabla 4.4: Proporción de *samples* asociados a tiempos meteorológicos en los conjuntos de entrenamiento y validación.

Conjunto	Soleados	Nublados	Nubosidad parcial
Train	63.5 %	24.1 %	13.0 %
Test	53.1 %	17.7 %	29.2 %

Los 13 atributos (*features*) utilizados en el registro de datos y en el procesamiento de los conjuntos se describen en la Tabla 4.5. Una vez generados los sets  $(X_{train}, Y_{train})$  y  $(X_{test}, Y_{test})$ , los valores correspondientes a cada uno de los atributos son escalados mediante una normalización Min-Max entre 0 y 1 considerando únicamente los datos contenidos en los conjuntos de entrenamiento.

$$\begin{aligned}
 f_{min} &= \min(X_{train, feature}), & f_{max} &= \max(X_{train, feature}) \\
 \bar{X}_{feature} &= \frac{X_{feature} - f_{min}}{f_{max} - f_{min}}
 \end{aligned}
 \tag{4.2}$$

#### 4.4.2. Datos satelitales del Advanced Baseline Imager

Las arquitecturas ConvLSTM/ConvJANet requieren datos estructurados de la forma (*input timestamps, width, height, channels*) para ser entrenadas. Para generar los conjuntos de entrenamiento y validación de este tipo de datos, en primer lugar la banda 04 de la base de datos de la NOAA es procesada para extraer una ventana de 48x48px centrada en las coordenadas del sitio estudiado (33.45775°S, 70.66466°W). Estas imágenes son ordenadas en un registro temporal de la forma (*number images, width x height*).

Tabla 4.5: Descripción de los atributos (*features*) utilizados en la generación de los conjuntos de entrenamiento y validación.

Atributo	Descripción
Potencia MC	Potencia eléctrica generada por el sistema completo.
Potencia SB 2500-HF(1)	Potencia eléctrica de salida del inversor que recibe de 5 strings orientados al oriente.
Potencia SB 2500-HF(2)	Potencia eléctrica de salida del inversor que recibe de 5 strings orientados al poniente.
Potencia SMC 5000-A(1)	Potencia eléctrica de salida del inversor que recibe de 14 strings orientados al oriente.
Potencia SMC 5000-A(2)	Potencia eléctrica de salida del inversor que recibe de 14 strings orientados al poniente.
Irradiación Global	Irradiación global en el plano horizontal registrada en la estación solarimétrica.
Irradiación Difusa	Irradiación difusa registrada en la estación solarimétrica.
Irradiación Directa	Irradiación directa registrada en la estación solarimétrica.
Temperatura Módulos	Temperatura en superficie de los módulos fotovoltaicos, registrados por el Sunny SensorBox.
Temperatura Ambiente	Temperatura ambiente registrada en la estación solarimétrica.
Día del Año	Número de día del año (1-365) correspondiente al <i>timestamp</i> del dato.
Minuto del Día	Número de minuto del día (1-1440) correspondiente al <i>timestamp</i> del dato.
Días Limpieza	Cantidad de días transcurridos desde la última limpieza de los módulos fotovoltaicos.

Siguiendo un procedimiento similar al utilizado para los datos del sistema, una ventana de tamaño (*input timesteps* + *output timesteps*, *width* x *height*) se va deslizando sobre el registro temporal para producir secuencias de imágenes estructuradas de la forma (*input timestamps*, *width*, *height*, 1) las cuales se van agregando a los conjuntos  $X_{img,train}/X_{img,test}$  correspondientes (Figura 4.13).

Una vez generado los conjuntos ( $X_{img,train}$ ,  $X_{img,test}$ ) los valores son escalados mediante una normalización Mix-Max considerando los niveles de intensidad dentro del set  $X_{img,train}$ . En la Figura 4.14 se muestra, a modo de ejemplo, el resultado de una de las secuencias procesadas.

## 4.5. Configuración de arquitecturas

En esta sección se describen las arquitecturas de redes neuronales propuestas. En estas, estructuras LSTM, JANet y ConvJANet son utilizadas para la extracción de patrones y

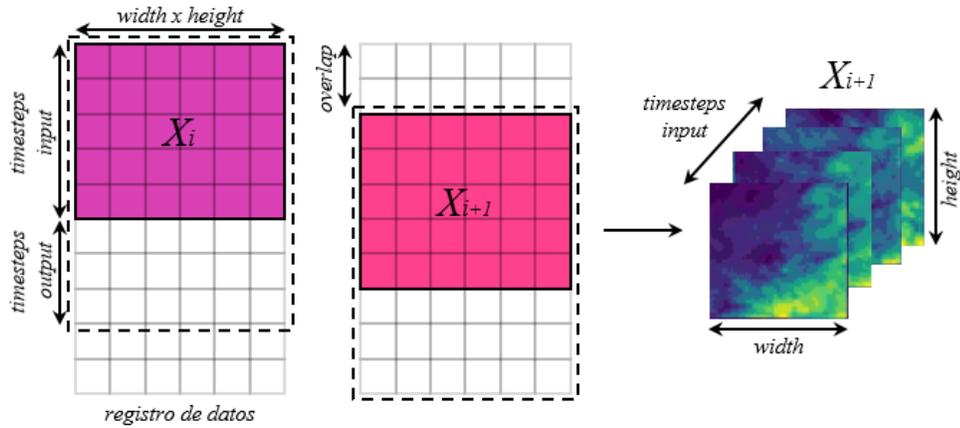


Figura 4.13: Muestra de procedimiento de generación de ventanas de datos para la creación de los conjuntos  $X$  de secuencias de imágenes.

características que permitan estimar la potencia de salida del sistema fotovoltaico en el horizonte de pronóstico.

La organización y cantidad de nodos de las capas, las funciones de activación y el resto de hiper-parámetros relevantes en la estructuración de los modelos de pronósticos fueron seleccionados luego de numerosas iteraciones de entrenamiento, siguiendo como criterio la minimización del RMSE de pronóstico en el conjunto de validación ( $X_{test}$ ,  $X_{train}$ ). Así, las arquitecturas detalladas a continuación corresponden a los modelos que obtuvieron los mejores resultados en el conjunto de alternativas evaluadas.

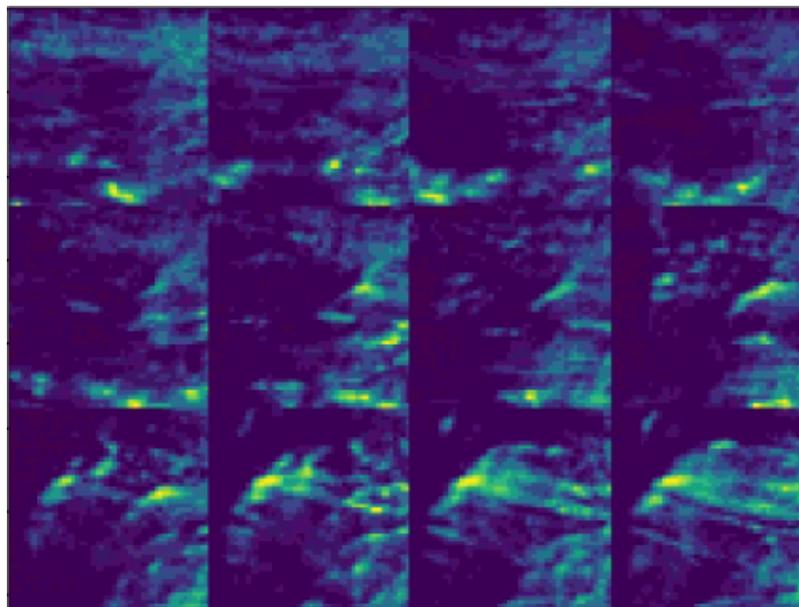


Figura 4.14: Muestra de una secuencia de imágenes satelitales procesadas. Periodo entre las 14:00 y 20:00 del 2 de Diciembre del 2018.

### 4.5.1. Modelos LSTM y JANet

En la Figura 4.15 se esquematiza el flujo de cómputo que siguen estos modelos. Como se observa, estos modelos reciben la información de operación de las 6 horas previas como una colección de las series temporales de los 13 atributos/variables del sistema (Tabla 4.5) y operan de la siguiente manera:

- Las series temporales son procesadas por las capas LSTM/JANet, que procesan individualmente cada atributo, para extraer patrones o características temporalmente relevantes de las secuencias.
- Luego, una serie de capas *Dense* post-procesan esta información para posteriormente combinar la información de todos los atributos mediante un *Flatten* del tensor.
- Otra serie de capas *Dense* continúan procesando la información conjunta de los atributos para finalmente estimar la potencia generada por el sistema durante las próximas 3 horas mediante una capa *Dense* con una función de activación lineal.

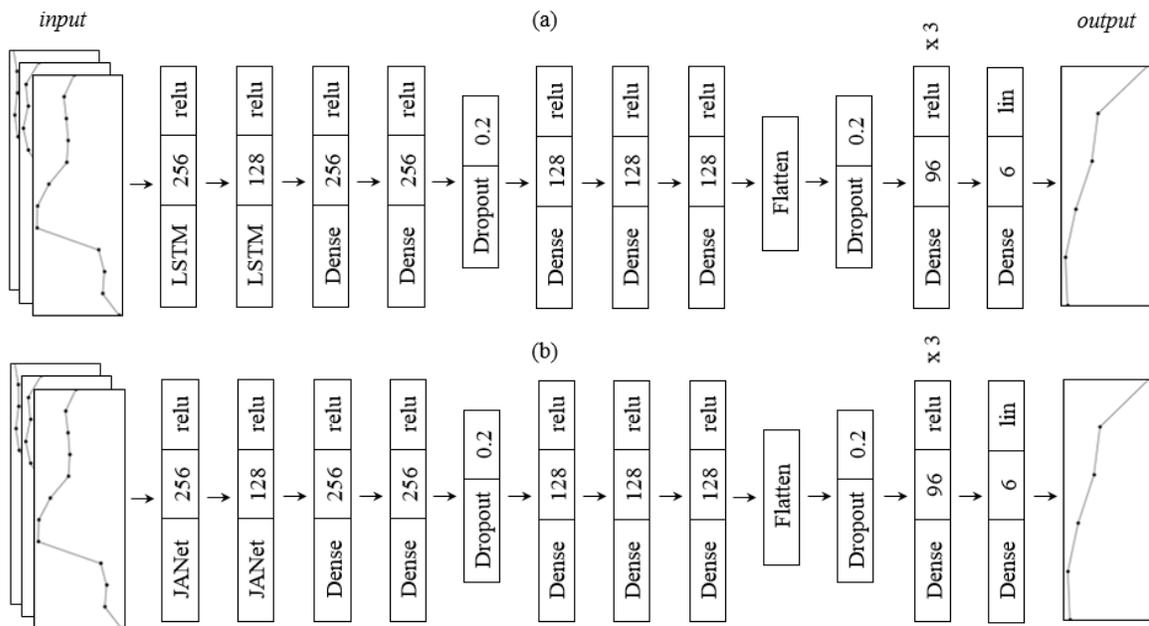


Figura 4.15: Esquema de las arquitecturas empleadas para los modelos LSTM y JANet.

### 4.5.2. Modelo JANet-ConvJANet

En la Figura 4.16 se esquematiza la arquitectura propuesta para este modelo de pronóstico. A diferencia de los modelos anteriormente descritos, aparte de recibir los datos operativos de las 6 horas previas, este modelo recibe la secuencia de imágenes satelitales correspondientes a las mismas 6 horas previas (12 imágenes). Esto se logra, como se puede ver en el esquema, dividiendo inicialmente la arquitectura en dos sub-modelos. Así, el modelo opera como se describe a continuación:

- La secuencia de imágenes es procesada inicialmente por una capa ConvJANet que extrae características desde la resolución original de la secuencia. Posteriormente, se aplica un MaxPooling3D para reducir el tamaño de las imágenes a 24x24px sin alterar la secuencialidad de la información.
- Luego, una serie de capas ConvJANet/MaxPooling3D consecutivas continúan procesando las secuencias latentes hasta que una última capa ConvJANet con una función de activación *sigmoide* retorna una única imagen latente de 12x12px.
- La información de esta imagen latente es post-procesada mediante un *Flatten* y una capa *Dense* sigmoide. Esta última, es luego concatenada con la salida del sub-modelo que procesa los datos operativos del sistema. Se puede observar que este sub-modelo posee una estructura similar a del modelo JANet descrito en la sección anterior.
- Una serie de capas *Dense* procesa y combina la información extraída de los datos operacionales y de la secuencia de imágenes para finalmente estimar las próximas 3 horas de generación mediante una capa *Dense* con una función de activación lineal.

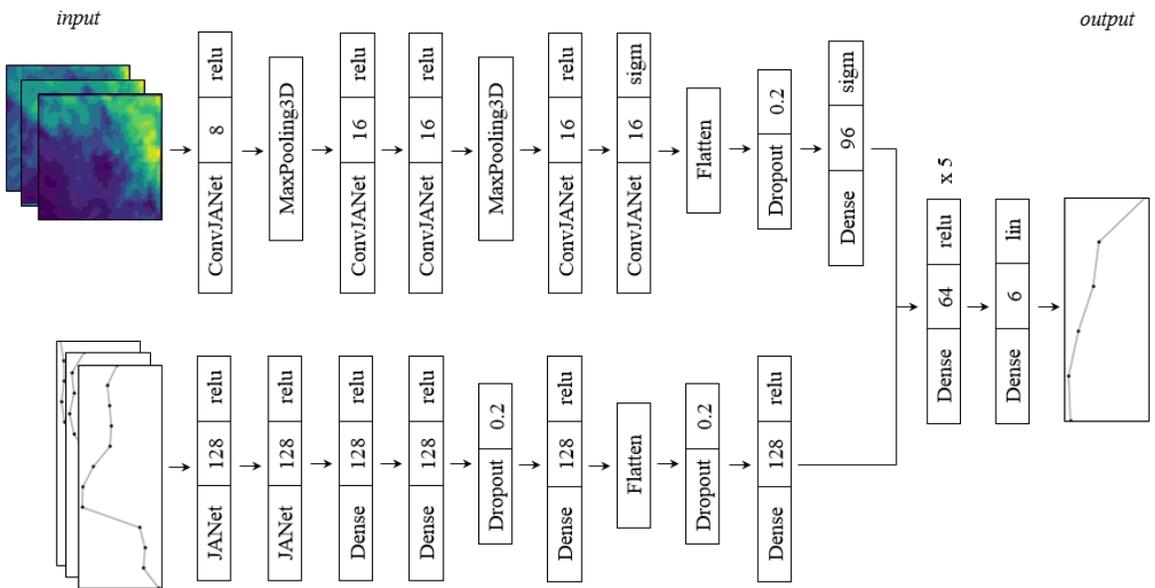


Figura 4.16: Esquema de la arquitectura empleada para el modelo JANet-ConvJANet.

## 4.6. Entrenamiento

Todos los modelos propuestos son entrenados y validados utilizando los conjuntos  $(X_{train}, Y_{train})$  y  $(X_{test}, Y_{test})$  anteriormente descritos. Por otro lado, la optimización de la función de pérdida MSE o error cuadrado medio, es llevada a cabo mediante el método ADAM con tasa de aprendizaje inicial de  $10^{-3}$ . Durante el entrenamiento, la tasa de aprendizaje es reducida, hasta un mínimo de  $10^{-6}$ , cada vez que se detecta una convergencia en la función de pérdida.

Los modelos resultantes son aquellos que alcanzaron el menor error MSE sobre el conjunto de validación durante 200 épocas de entrenamiento.

## 4.7. Métricas de evaluación

El desempeño de todos los modelos desarrollados es evaluado mediante seis métricas estadísticas que cuantifican la distribución del error entre los valores de potencia pronosticados  $\hat{\mathcal{P}}(t)$  y los reales registrados  $\mathcal{P}(t)$ . Con el fin de presentar los resultados de un modo que sea comparable con otras investigaciones tanto los valores reales como los estimados son normalizados utilizando un escalador MinMax en base a los registros de potencia presentes en el conjunto de datos  $X_{train}$ . Así, los valores  $\mathcal{P}(t)$  y  $\hat{\mathcal{P}}(t)$  están acotados al intervalo  $[0, 1]$ .

El Mean Bias Error (MBE) es el promedio del error de pronóstico y representa la tendencia del modelo a subestimar o sobrestimar los valores de potencia futuros.

$$MBE = \frac{1}{n} \sum_{i=1}^n (\hat{\mathcal{P}}(t) - \mathcal{P}(t)) \quad (4.3)$$

El Mean Absolute Error (MAE) mide la desviación promedio de los valores pronosticados respecto a los valores registrados al igual que el Root Mean Squared Error (RMSE). No obstante, este último es más sensible a *outliers* que presenten errores significativamente mayores al resto de los datos.

$$MAE = \frac{1}{n} \sum_{i=1}^n (|\hat{\mathcal{P}}(t) - \mathcal{P}(t)|) \quad (4.4)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{\mathcal{P}}(t) - \mathcal{P}(t))^2} \quad (4.5)$$

El Forecast Skill (FS %) representa la reducción o incremento en el error de pronóstico del modelo respecto a un método de referencia. En este caso, se utiliza el método de persistencia. Así, si el modelo tiene un desempeño equivalente al método de persistencia, el FS es 0.

$$FS = 1.0 - \frac{RMSE}{RMSE_{pers}} \quad (4.6)$$

La Standard Deviation (STD) cuantifica la dispersión de la distribución del error de pronóstico con respecto al error medio en el conjunto de datos.

$$STD = \sqrt{\frac{1}{n} \sum_{i=1}^n [(\hat{\mathcal{P}}(t) - \mathcal{P}(t)) - MBE]^2} \quad (4.7)$$

Por último, la Skewness (SKW) cuantifica el sesgo o *bias* en el pronóstico del modelo a partir de la distribución del error. Un Skewness negativo representa una tendencia a la subestimación en el pronóstico, mientras que uno positivo indica una sobrestimación.

$$SKW = \frac{MBE - Moda[\hat{\mathcal{P}}(t) - \mathcal{P}(t)]}{STD} \quad (4.8)$$

# Capítulo 5

## Resultados y discusión

Los tres modelos de pronóstico propuestos: LSTM, JANet y JANet-ConvJANet, son evaluados individualmente sobre cada uno de los *clusters* identificados con el fin de analizar su desempeño frente a distintas condiciones meteorológicas. Así, es posible analizar aisladamente la distribución del error de pronóstico de los modelos en días soleados, nublados y parcialmente nublados. Además, para disponer de una referencia base con la que comparar la capacidad de pronóstico de cada modelo, en la Tabla 5.1 se presentan las métricas de evaluación obtenidas con los métodos estocásticos detallados en la sección 2.3 al ser entrenados con los mismos conjuntos  $(X_{train}, Y_{train})$  utilizados para los modelos de redes neuronales.

Tabla 5.1: Métricas estadísticas obtenidas en la evaluación de desempeño de los métodos de estocásticos de referencia.

Método	Cluster	MBE	MAE	RMSE	FS %	STD	SKW
Persistencia	Soleado	-0.0318	0.0360	0.0785	0.000 %	0.0717	-2.7842
	Nublado	-0.0093	0.0253	0.0563	0.000 %	0.0555	-1.6189
	Nublado parcial	-0.0458	0.0630	0.1500	0.000 %	0.1429	-3.0946
AR	Soleado	-0.0276	0.0587	0.0994	-26.745 %	0.0955	-1.4877
	Nublado	-0.0073	0.0457	0.0815	-44.932 %	0.0812	0.5365
	Nublado parcial	-0.0248	0.0762	0.1326	11.615 %	0.1303	-0.0953
KNN	Soleado	-0.0004	0.0268	0.0539	31.360 %	0.0538	-1.9195
	Nublado	0.0232	0.0377	0.0669	-18.942 %	0.0628	1.3646
	Nublado parcial	0.0017	0.0429	0.0833	44.468 %	0.0833	-0.7101

Adicionalmente, para una mejor comprensión de la capacidad de pronóstico de los modelos, los gráficos se presentan para cada intervalo dentro del horizonte de pronóstico de 3 horas. De este modo, es posible estudiar la evolución de la distribución del error a medida que se incrementa el intervalo de pronóstico  $\Delta t = \{30 \text{ min}, 45 \text{ min}, \dots, 180 \text{ min}\}$ .

## 5.1. Análisis modelos LSTM y JANet

De los experimentos realizados con las arquitecturas que operan sobre los datos operacionales y meteorológicos del sistema, se obtiene, como se muestra en las Tablas 5.2 y 5.3, que el modelo JANet presenta un desempeño de pronóstico superior al del modelo LSTM. En particular, en todos los grupos de evaluación analizados, el modelo JANet presenta errores entre un 4.13 % y un 7.42 % menores que su contraparte.

Tabla 5.2: Métricas estadísticas obtenidas en la evaluación de desempeño del modelo LSTM.

Cluster	MBE	MAE	RMSE	FS %	STD	SKW
Soleado	-0.0039	0.0127	0.0266	66.131 %	0.0263	-0.6550
Nublado	0.0134	0.0255	0.0544	3.282 %	0.0527	1.4827
Nublado parcial	0.0026	0.0291	0.0647	56.881 %	0.0646	1.3119

Tabla 5.3: Métricas estadísticas obtenidas en la evaluación de desempeño del modelo JANet.

Cluster	MBE	MAE	RMSE	FS %	STD	SKW
Soleado	-0.0044	0.0123	0.0255	67.546 %	0.0251	-0.7321
Nublado	0.0122	0.0242	0.0516	8.352 %	0.0501	1.3443
Nublado parcial	0.0017	0.0267	0.0599	60.108 %	0.0598	1.1538

Esta diferencia en las métricas estadísticas puede explicarse por la cantidad de parámetros que deben ser entrenados en cada una de estas arquitecturas, ya que si bien las celdas LSTM y JANet procesan las secuencias temporales de forma análoga, esta última utiliza solo una compuerta de procesamiento, la *forget gate*. En este sentido, al manejar una menor cantidad de pesos ponderadores ( $U$  y  $W$ ) se reduce considerablemente la dimensionalidad del espacio de búsqueda en la optimización de la función de pérdida y, por tanto, el problema del desvanecimiento del gradiente durante el entrenamiento.

Ahora, si bien la mejora en el desempeño del modelo JANet, con respecto al LSTM, es significativo, se puede observar que, con respecto a sus distribuciones del error (métricas STD y SKW), ambas arquitecturas presentan un comportamiento similar en el pronóstico de la potencia del sistema, con la diferencia de que la JANet tiene una menor dispersión en el error de pronóstico. De este modo, dado que los análisis de desempeño de ambos modelos sobre los tres grupos de evaluación serían equivalentes, a continuación solo se presenta el análisis del modelo JANet.

### 5.1.1. Desempeño en días soleados

Como se muestra en la Tabla 5.3 el modelo JANet se desempeña considerablemente mejor en este tipo de días con un error MAE de 0.0123 y un RMSE de 0.0255 que, debido a que los datos se encuentran normalizados entre 0 y 1, equivaldría a un error de pronóstico de 1.23 %. Dado que los días pertenecientes a este *cluster* se caracterizan, como se puede observar en las Figuras 4.9 y 4.10, por perfiles de irradiación de baja variabilidad (bajo índice de suavidad),

resulta esperable que el modelo presente un mejor desempeño en estos días en comparación al resto de los grupos. Del mismo modo, como en los días soleados las condiciones atmosféricas son relativamente constantes, la distribución del error de pronóstico, como se puede ver en las Figuras 5.1 y 5.2 se mantiene prácticamente igual en cada uno de los intervalos dentro del horizonte de pronóstico.

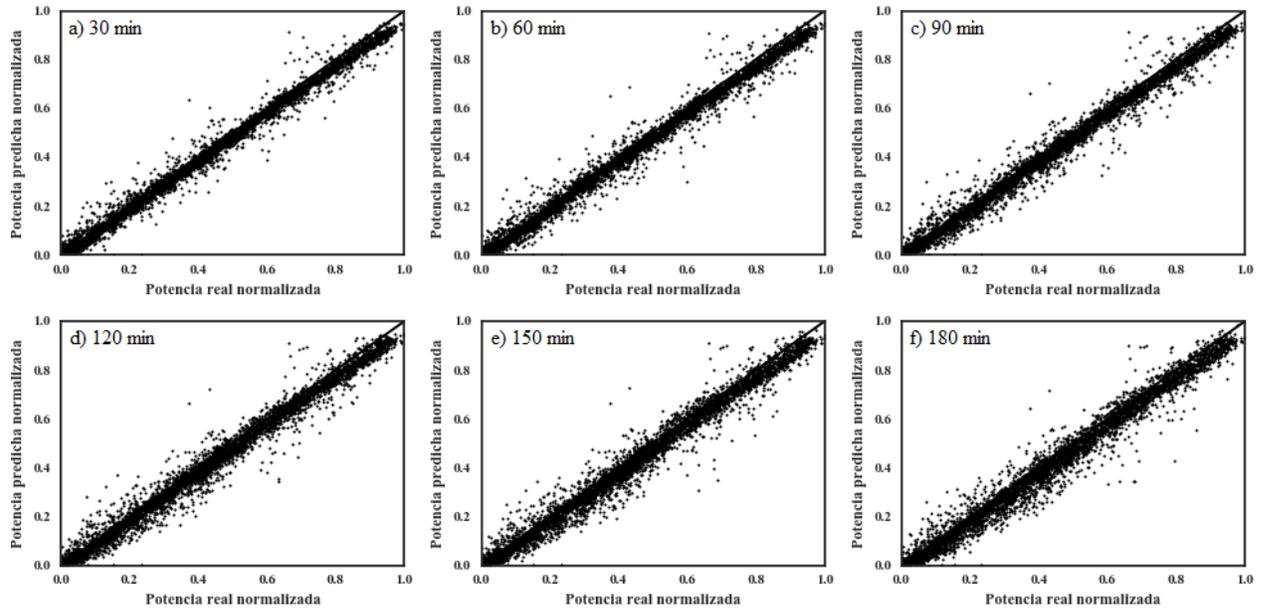


Figura 5.1: Dispersión de los pronósticos del modelo JANet con respecto a los valores reales para cada uno de los intervalos del horizonte de pronóstico en días soleados.

Mientras los métodos estocásticos de referencia (Modelo de persistencia, AR y KNN) en general muestran una tendencia a subestimar la potencia a medida que el intervalo de tiempo aumenta, con distribuciones que presentan además dispersiones cada vez mayores (colas más extendidas en la Figura 5.2); el modelo JANet, si bien también presenta una ligera mayor dispersión en el pronóstico de los últimos intervalos, no evidencia una tendencia clara de subestimación o sobrestimación en la potencia. Esto último se comprueba observando las métricas MBE y SKW en la Tabla 5.3 que aunque indican una leve tendencia a la subestimación, principalmente cuando se pronostican valores altos de potencia (Figura 5.1), esta es insignificante.

Por otro lado, con respecto a la capacidad de pronóstico de este modelo frente a distintas condiciones de variabilidad en la irradiación en este tipo de días, como se puede observar en la Figura 5.3, el error de pronóstico de la arquitectura JANet, al igual que el de los modelos de referencia, no muestra una correlación evidente con respecto a la variabilidad atmosférica. Este comportamiento tiene sentido pues, dado que los días soleados presentan una baja variabilidad en las condiciones meteorológicas, la mayor cantidad de puntos se concentran en el rango 0.00-0.05, por lo que el resto de puntos no son capaces de mostrar una tendencia clara.

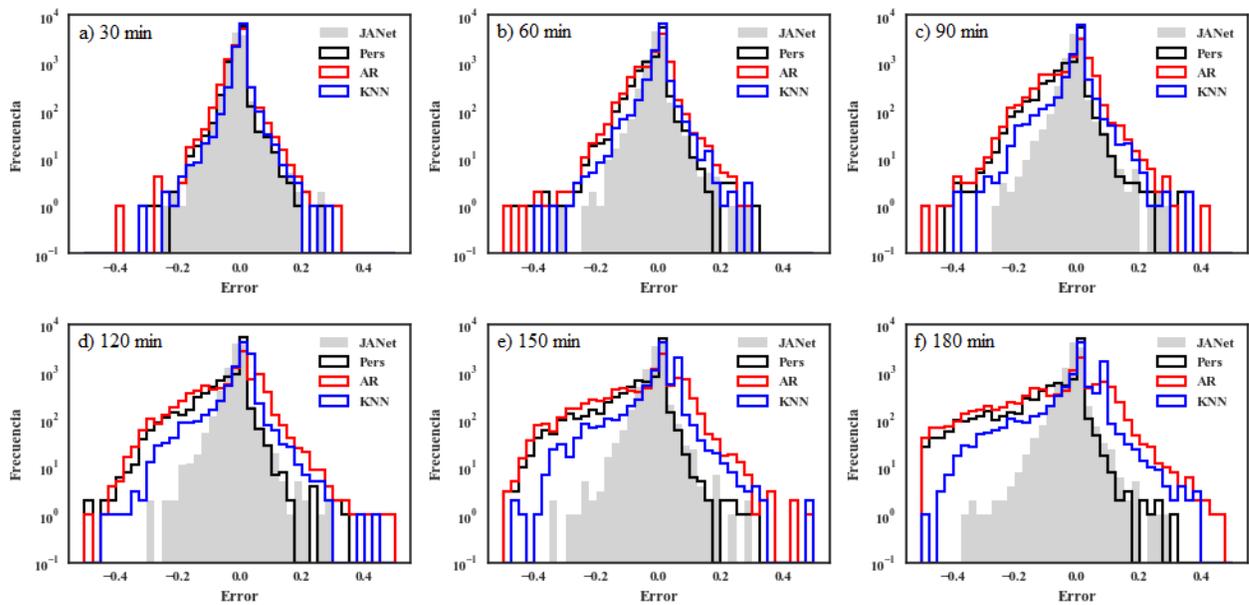


Figura 5.2: Distribución del error de pronóstico del modelo JANet para cada uno de los intervalos del horizonte de pronóstico en días soleados.

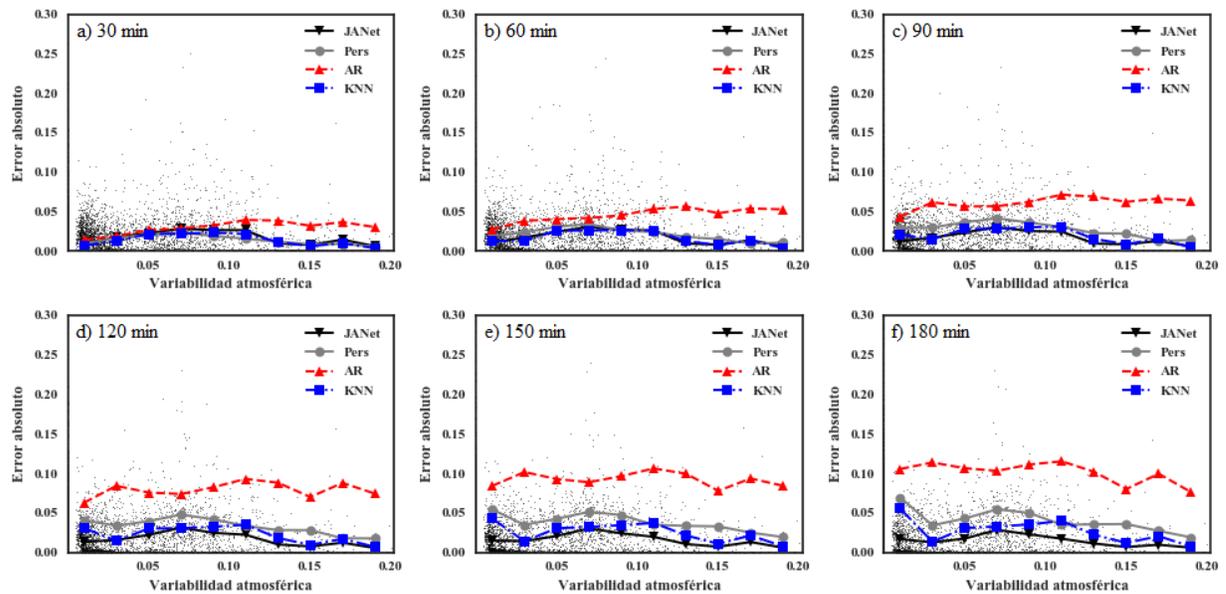


Figura 5.3: Error absoluto promedio del modelo JANet versus la variabilidad atmosférica para cada uno de los intervalos del horizonte de pronóstico en días soleados.

### 5.1.2. Desempeño en días nublados

Si bien este grupo de días, al igual que los días soleados, presentan un índice de suavidad bajo (Figura 4.9) y por tanto curvas de irradiación con una relativa baja variabilidad, de las métricas presentadas en la Tabla 5.3 y de los gráficos de las Figuras 5.4 y 5.5, es evidente que la capacidad de pronóstico del modelo JANet es significativamente menor en este caso. No obstante, aún con un error MAE de 0.0242 y un RMSE 2.02 veces mayor que el obtenido sobre los días soleados, estos resultados son considerablemente mejores que los obtenidos con los métodos de referencia AR y KNN (Forecast Skills de -44.9 % y -18.9 %, respectivamente), aunque no tanto con respecto al modelo de persistencia, alcanzando un Forecast Skill de solo un 8.352 %.

La razón de esta disminución en el desempeño de pronóstico puede deberse a que el modelo, como recibe únicamente datos operativos del sistema, no es capaz de distinguir si una secuencia temporal con una baja irradiación solar corresponde a un día nublado o a uno parcialmente nublado. Se puede notar en la Figura 4.10 que hay días clasificados como parcialmente nublados que contienen intervalos muy similares a los de un día nublado. Así, como el modelo JANet trata de estimar la secuencia de potencia más probable a partir de la información que recibe, es posible que en ocasiones interprete erróneamente que un perfil de nubosidad parcial sea más probable que uno de alta nubosidad, aunque este último sea el caso.

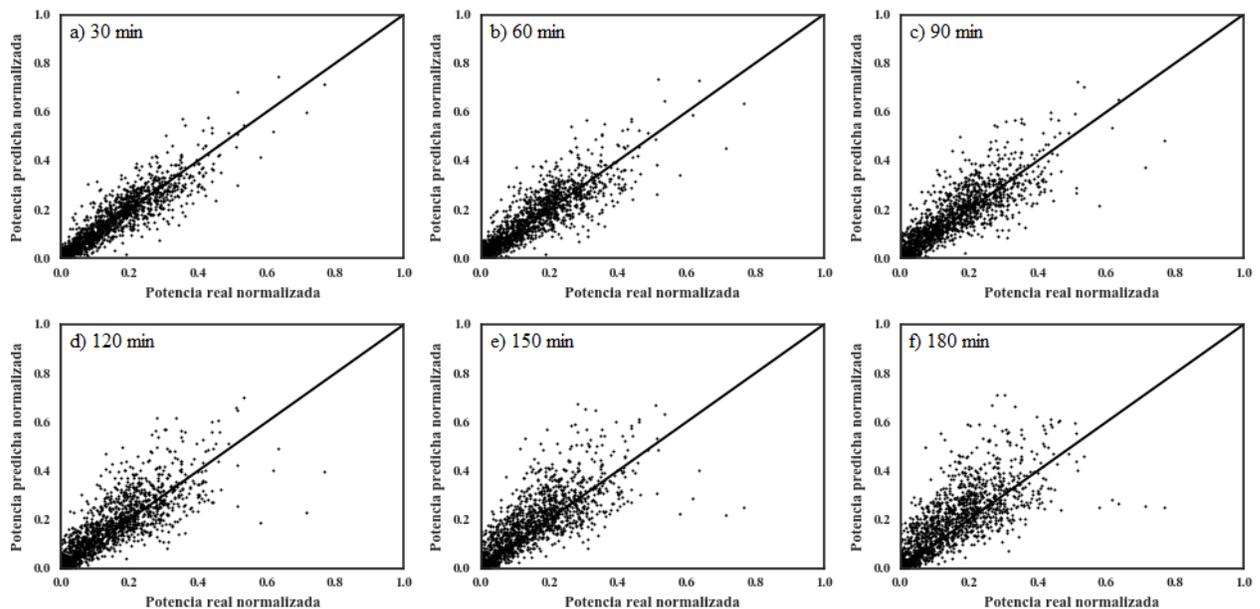


Figura 5.4: Dispersión de los pronósticos del modelo JANet con respecto a los valores reales para cada uno de los intervalos del horizonte de pronóstico en días nublados.

Esta última interpretación también podría explicar el comportamiento que se observa en los gráficos de las Figuras 5.4 y 5.5, donde se puede notar que la dispersión del error incrementa considerablemente a medida que el intervalo de pronóstico aumenta, ya que el modelo en ocasiones tendería a pronosticar la potencia que se tendría en un día de nubosidad parcial.

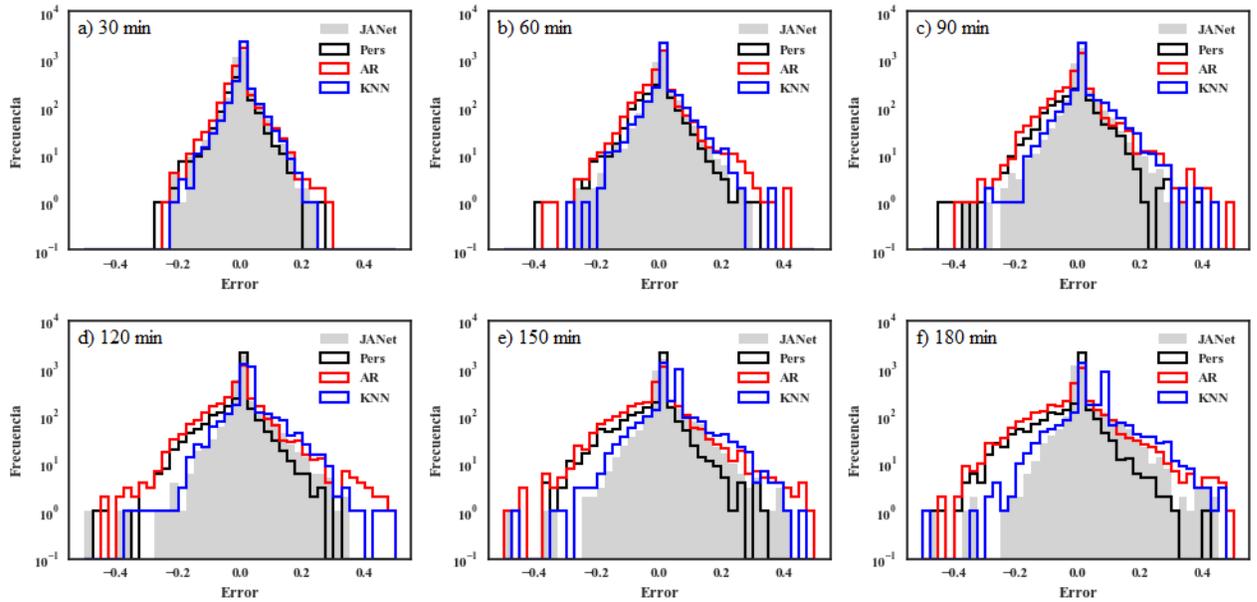


Figura 5.5: Distribución del error de pronóstico del modelo JANet para cada uno de los intervalos del horizonte de pronóstico en días nublados.

Así, en este caso, y a diferencia de lo observado en días soleados, la tendencia del modelo es a sobrestimar la potencia del sistema, como lo indica el Skewness de la distribución, con un valor de 1.344. Se puede observar que este mismo comportamiento se observa, y en mayor medida, en los métodos estocásticos utilizados como referencia.

Ahora, con respecto a la relación entre el error de pronóstico y la variabilidad meteorológica, se puede observar en la Figura 5.6 que en todos los intervalos de pronóstico existe un leve pero claro aumento del error promedio a medida que la variabilidad es mayor (en el rango 0.00 - 0.10). Luego, en el rango 0.100 - 0.160, esta relación decae lo cual se puede explicar nuevamente por la baja concentración de puntos en este rango. Este comportamiento, en general, resulta esperable considerando que una mayor variabilidad meteorológica dificulta significativamente la estimación de la potencia en el futuro.

### 5.1.3. Desempeño en días parcialmente nublados

De las Figuras 5.7 y 5.8 se desprende que los resultados de este conjunto de datos exhiben un comportamiento muy similar al observado en el caso anterior de días nublados. Con un MAE de 0.0267 y un RMSE de 0.0599, este *cluster* presenta los peores resultados del modelo JANet, siendo un 10.3 % más altos que los observados en los días nublados. Aún así, frente a este grupo de datos, el modelo presenta un Forecast Skill 7.2 veces más alto que en los días nublados. Esto significa que si bien el modelo JANet tiene un desempeño similar al de los días nublados, este es considerablemente mejor que cualquiera de los modelos estocásticos empleados, como se puede verificar en la Tabla 5.1.

Se puede ver en los gráficos de las Figuras 5.7 y 5.8 que la distribución del error, al igual que en los casos anteriores, tiende a una mayor dispersión a medida que el intervalo de tiempo

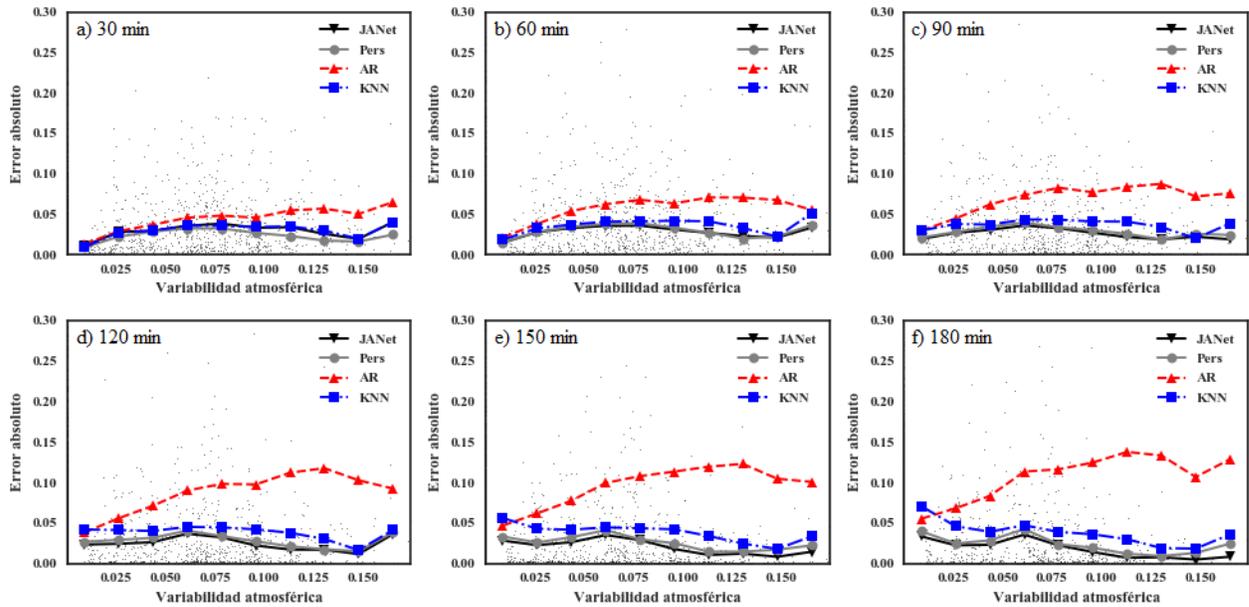


Figura 5.6: Error absoluto promedio del modelo JANet versus la variabilidad atmosférica para cada uno de los intervalos del horizonte de pronóstico en días nublados.

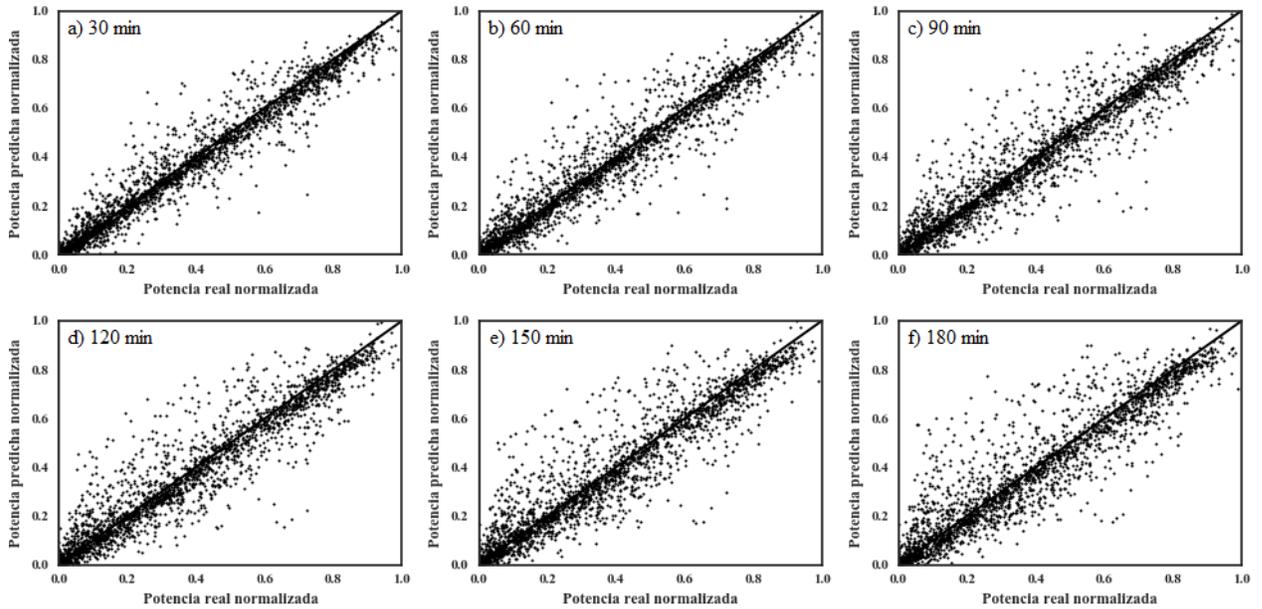


Figura 5.7: Dispersión de los pronósticos del modelo JANet con respecto a los valores reales para cada uno de los intervalos del horizonte de pronóstico en días parcialmente nublados.

en el pronóstico aumenta. No obstante, vale la pena notar que en cada uno de los tiempos de pronóstico, la dispersión de los datos es relativamente uniforme a lo largo de todo el rango de potencia (Figura 5.7), a diferencia de por ejemplo, lo observado en los días nublados. En cuanto al *bias* del modelo al pronosticar en este contexto, un Skewness de 1.1538 refleja que el modelo JANet presenta una tendencia a sobrestimar la potencia del sistema, contraria a la que presentan los métodos estocásticos, los cuales tienden a subestimar la potencia del sistema a medida que aumenta el intervalo de pronóstico, como se puede ver en el gráfico de distribución.

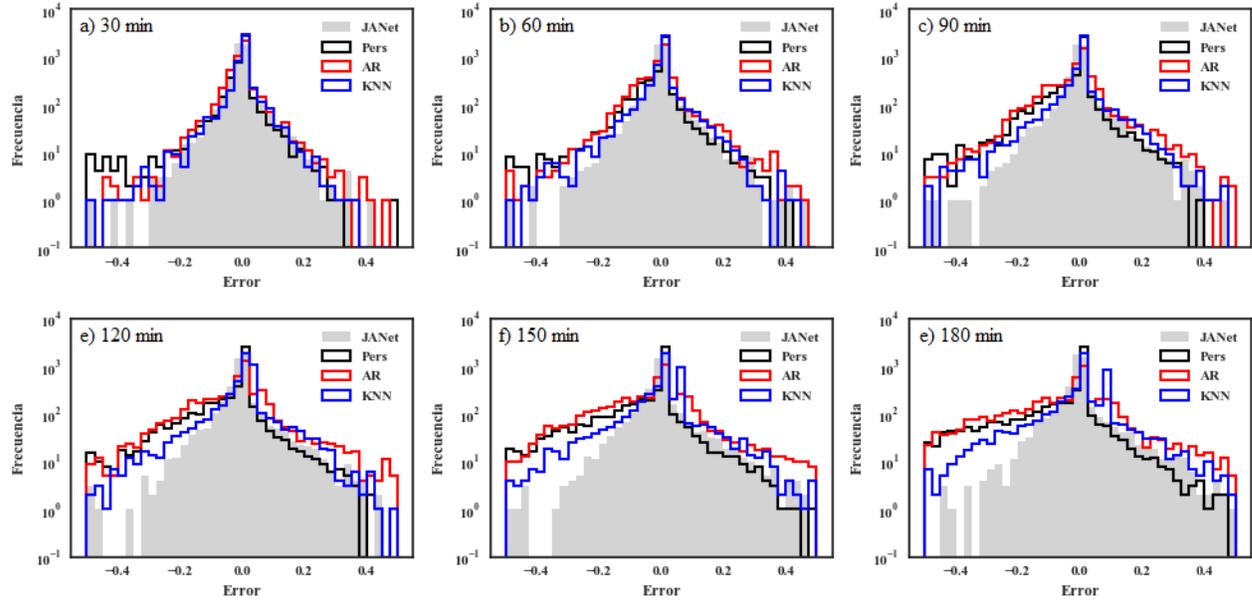


Figura 5.8: Distribución del error de pronóstico del modelo JANet para cada uno de los intervalos del horizonte de pronóstico en días parcialmente nublados.

Por último, se puede notar de la Figura 5.9 que en este caso, si bien el error de pronóstico también aumenta a medida que la variabilidad meteorológica es mayor, a diferencia de lo que se había visto anteriormente en los días soleados y nublados, este incremento es considerablemente menor que el de los métodos estocásticos. En particular, el modelo JANet consigue mantener un error relativamente uniforme a lo largo de todo el rango de variabilidad atmosférica del gráfico.

## 5.2. Análisis modelo JANet-ConvJANet

Como se detalló en la sección anterior, el modelo de pronóstico JANet-ConvJANet propuesto, a diferencia de los modelos anteriores, tiene la particularidad de recibir e interpretar la información espacio-temporal de la secuencia de imágenes satelitales de la nubosidad en el entorno del sitio estudiado en un radio de 48 km. Si bien estas imágenes no son de una gran resolución, como se puede ver en la Tabla 5.4, esta información adicional incrementa significativamente el desempeño del modelo con respecto a los casos anteriores.

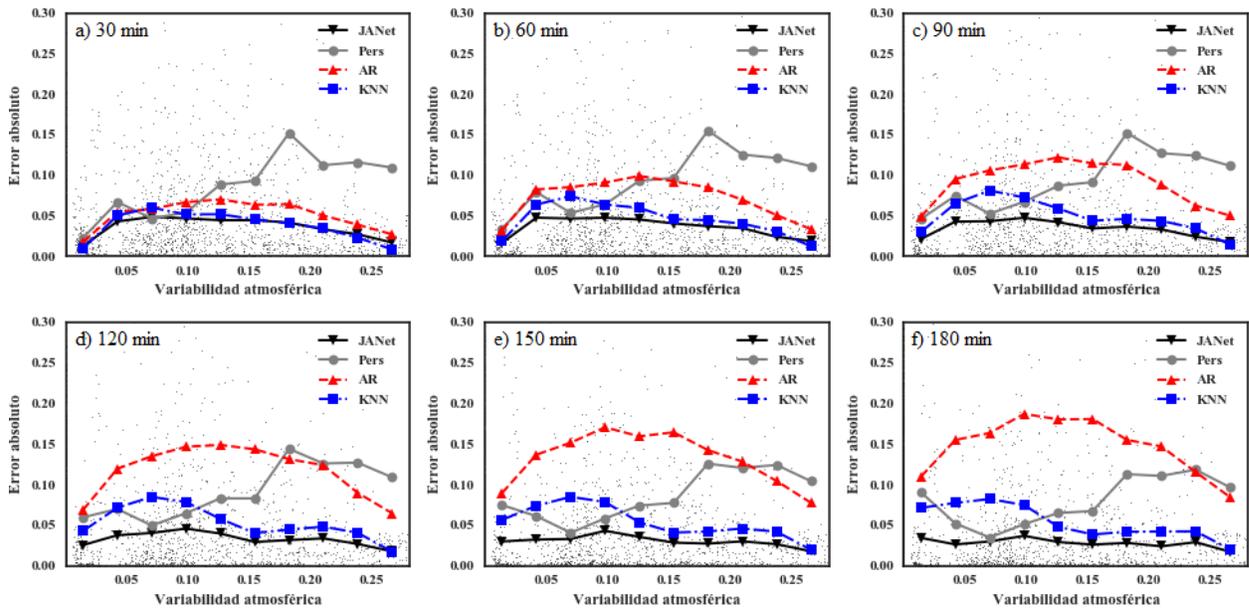


Figura 5.9: Error absoluto promedio del modelo JANet versus la variabilidad atmosférica para cada uno de los intervalos del horizonte de pronóstico en días parcialmente nublados.

En general, con esta arquitectura el error de pronóstico es reducido, en menor o en mayor medida, en todos los *clusters* de evaluación. No obstante, es en los días nublados donde la reducción resulta considerablemente mayor, alcanzando un Forecast Skill 2.0 veces mayor y una reducción del 9.3% en el RMSE respecto a los resultados obtenidos con el modelo JANet. Por lo tanto, en primera instancia, las secuencias de imágenes satelitales si logran ser interpretadas por la estructura ConvJANet como información adicional sobre las condiciones previas al horizonte de pronóstico. No obstante, este punto será analizado en mayor profundidad en las siguientes secciones.

Tabla 5.4: Métricas estadísticas obtenidas en la evaluación de desempeño del modelo JANet-ConvJANet.

Cluster	MBE	MAE	RMSE	FS %	STD	SKW
Soleado	-0.0046	0.0113	0.0245	68.787 %	0.0241	-0.8826
Nublado	0.0089	0.0214	0.0468	16.754 %	0.0460	1.2108
Nublado parcial	0.0005	0.0238	0.0554	63.091 %	0.0554	0.9882

### 5.2.1. Desempeño en días soleados

Observando, tanto las métricas entregadas en la Tabla 5.4 como los gráficos de las Figuras 5.10, 5.11 y 5.12, es evidente que el desempeño del modelo JANet-ConvJANet no se aleja mucho del visto con el modelo JANet. Si bien este modelo logra reducir en un 3.9% el error RMSE y aumentar el Forecast Skill a un 68.8%, el incremento del *bias* en la distribución del error de pronóstico (métricas MBE y SKW) no permite concluir sobre si efectivamente se

logra una mejora en la capacidad de pronóstico en estos días.

De todas maneras, esta similaridad entre los resultados de los modelos JANet y JANet-ConvJANet es esperable pues en este tipo de días las imágenes satelitales no aportan mucha información adicional a la que ya contienen los datos operacionales y meteorológicos, salvo quizás indicar la ausencia de nubosidad en el cielo en las primeras horas de la secuencia de entrada. De este modo, el análisis de la distribución del error es análogo al realizado para el modelo JANet por lo que no se profundizará en este.

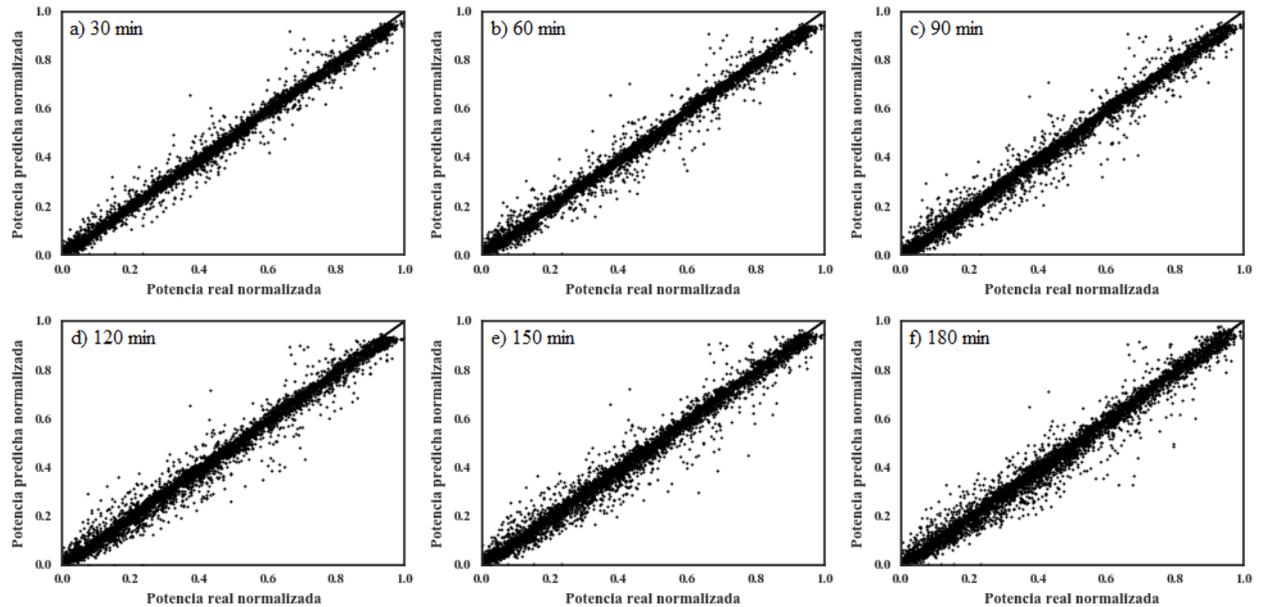


Figura 5.10: Dispersión de los pronósticos del modelo JANet-ConvJANet con respecto a los valores reales para cada uno de los intervalos del horizonte de pronóstico en días soleados.

### 5.2.2. Desempeño en días nublados

Como se mencionó anteriormente, es en este tipo de días donde se comprueba la capacidad de la estructura ConvJANet de interpretar la información contenida en las secuencias de imágenes satelitales y lograr una representación que le permita a la red tener una noción sobre la nubosidad en el entorno del arreglo fotovoltaico.

Se puede ver de las Figuras 5.13 y 5.14 que la dispersión de los pronósticos resulta menor a la observada con el modelo JANet en este mismo *cluster*. Por otro lado, la sobrestimación de la potencia también es reducida siendo mucho menos aparante en los primeros intervalos de pronóstico. En términos estadísticos, se evidencia una reducción de la desviación estándar del error de pronóstico en un 8.2%, mientras que el Skewness de la distribución disminuye a 1.21.

Ahora bien, es evidente que la reducción observada en el error de pronóstico no es lo suficiente como para concluir que el modelo es capaz de interpretar y modelar la dinámica de la nubosidad en el entorno del sitio estudiado. Es posible que, producto de la semejanza entre la

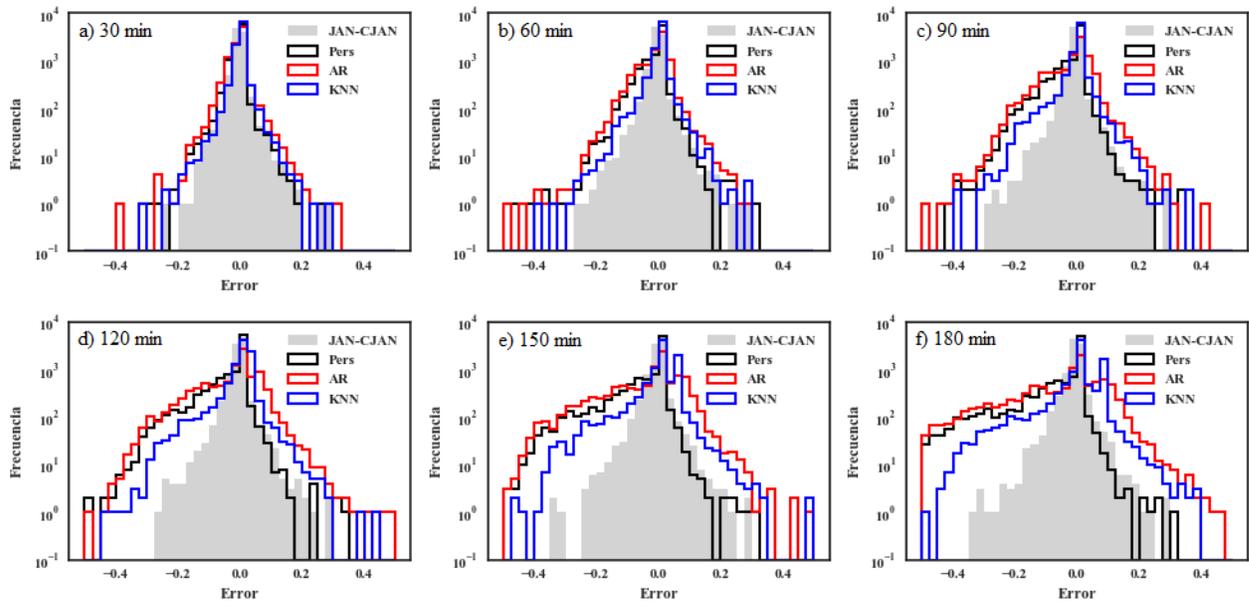


Figura 5.11: Distribución del error de pronóstico del modelo JANet-ConvJANet para cada uno de los intervalos del horizonte de pronóstico en días soleados.

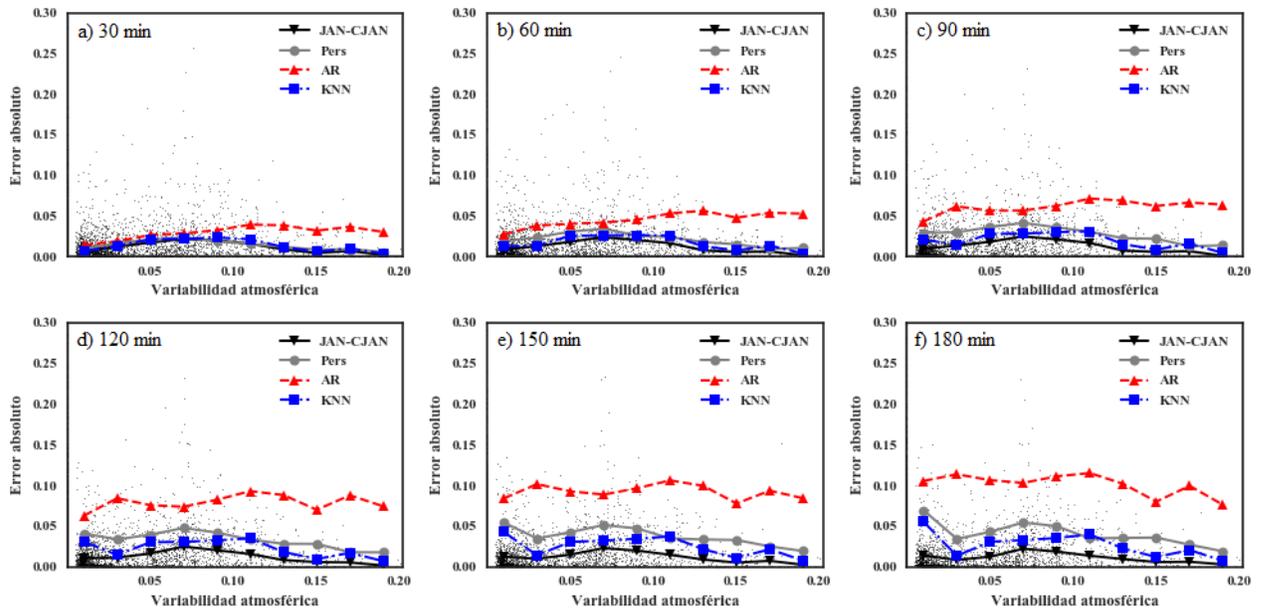


Figura 5.12: Error absoluto promedio en el modelo JANet-ConvJANet versus la variabilidad atmosférica para cada uno de los intervalos del horizonte de pronóstico en días soleados.

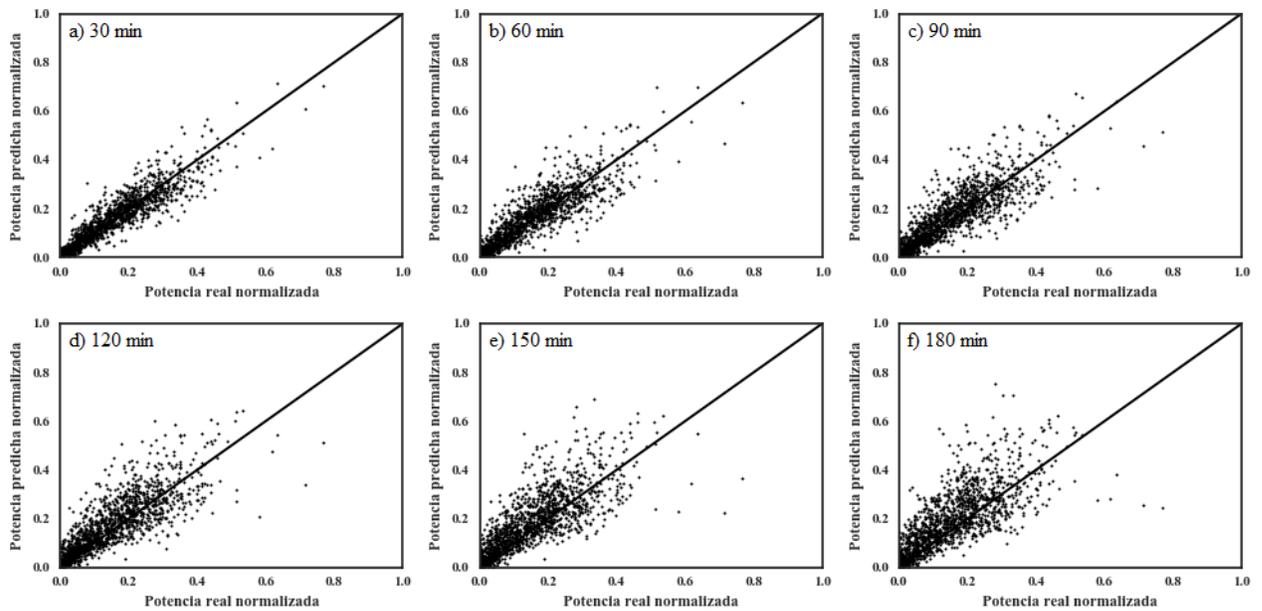


Figura 5.13: Dispersión de los pronósticos del modelo JANet-ConvJANet con respecto a los valores reales para cada uno de los intervalos del horizonte de pronóstico en días nublados.

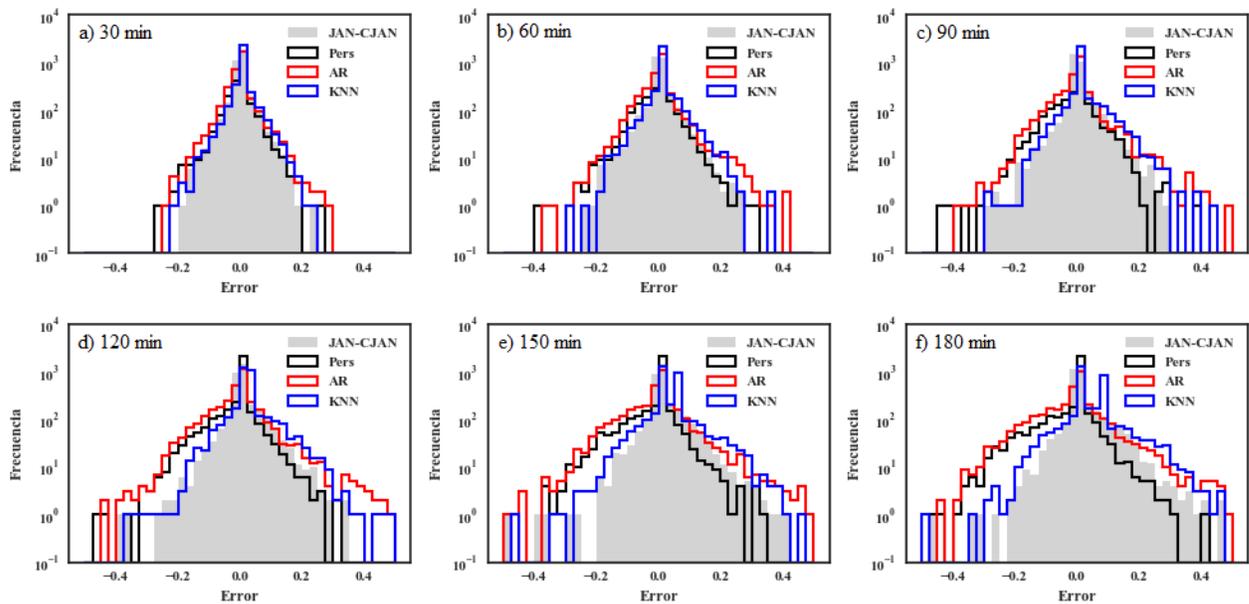


Figura 5.14: Distribución del error de pronóstico del modelo JANet-ConvJANet para cada uno de los intervalos del horizonte de pronóstico en días nublados.

nubosidad que presentan algunos días nublados y otros parcialmente nublados (en términos de las imágenes satelitales), la red JANet-ConvJANet presente dificultades similares a las descritas para el modelo JANet. De este modo, es probable que la estructura ConvJANet sea capaz de aportar información relevante principalmente en los casos en que las secuencias presenten una alta nubosidad en la zona.

Con respecto al gráfico del error absoluto presentado en la Figura 5.15, no se distinguen mayores diferencias con respecto al observado para el caso del modelo ConvJANet. Por lo tanto, tampoco se profundizará en el análisis de este.

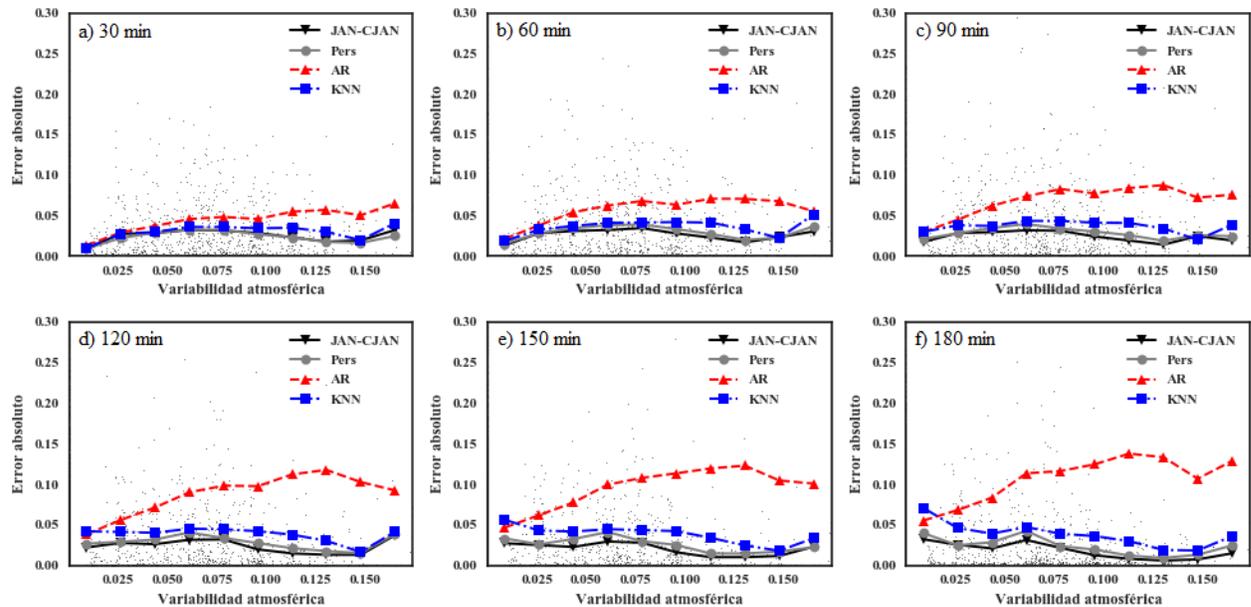


Figura 5.15: Error absoluto promedio del modelo JANet-ConvJANet versus la variabilidad atmosférica para cada uno de los intervalos del horizonte de pronóstico en días nublados.

### 5.2.3. Desempeño en días parcialmente nublados

Análogamente a lo observado en el caso de los días nublados, en este grupo de datos también se evidencia una mejora en la capacidad de pronóstico del modelo JANet-ConvJANet. Esta mejora se ve reflejada en la reducción del RMSE en un 7.5 %, que si bien no es mayor a la observada en los días nublados, no deja de ser significativa.

En cuanto a la distribución del error de pronóstico, en las Figuras 5.16 y 5.17 se distingue que, en comparación al modelo JANet, el modelo reduce la sobrestimación de la potencia como también se desprende de su Skewness (0.99). Por otro lado, de acuerdo a la desviación estándar, este modelo reduce la dispersión de los resultados, no obstante, esta mejora no se desprende directamente de los gráficos.

El hecho de que las mejoras en la capacidad de pronóstico de este modelo hayan sido

menores en este *cluster* en comparación al de días nublados, puede deberse a que las características de cielo de este tipo de días suelen ser más variables y con un comportamiento más dinámico que el de días de alta nubosidad. Así, es posible que la estructura ConvJANet de la arquitectura tenga más dificultades de extraer patrones y características útiles de los datos.

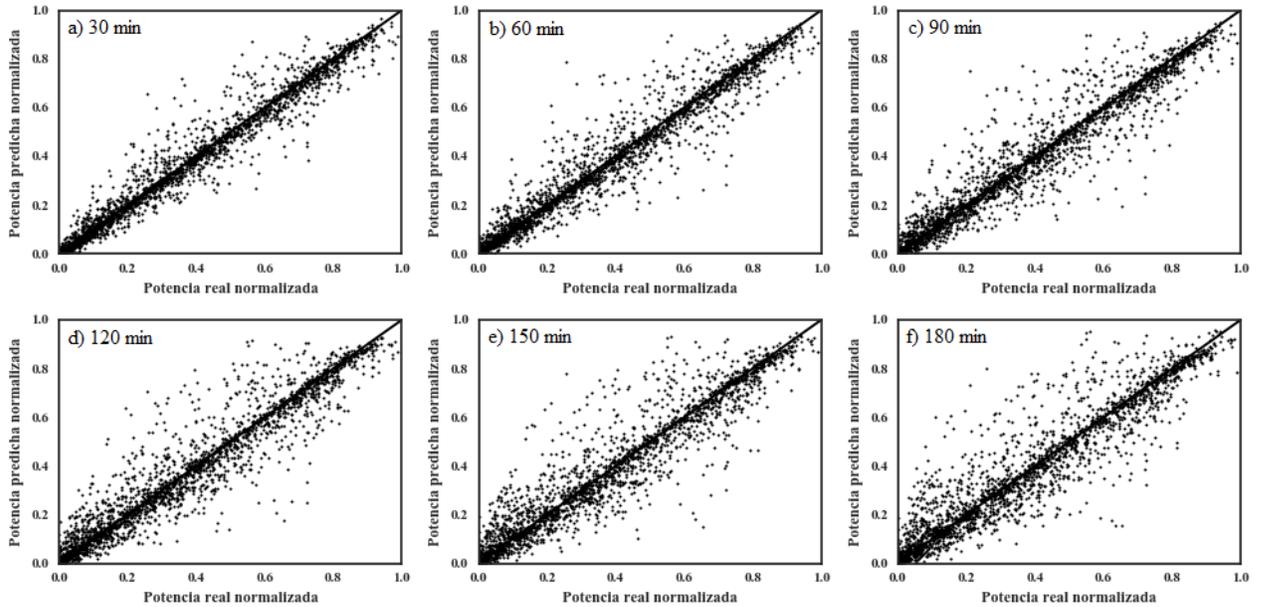


Figura 5.16: Dispersión de los pronósticos del modelo JANet-ConvJANet con respecto a los valores reales para cada uno de los intervalos del horizonte de pronóstico en días parcialmente nublados.

Finalmente, con respecto a la Figura 5.18, nuevamente el resultado es similar al obtenido con el modelo JANet. Así, la arquitectura JANet-ConvJANet también logra mantener un error promedio relativamente constante frente a distintos contextos de variabilidad atmosférica. Vale la pena agregar que, en este sentido, el modelo resulta muy superior a los métodos de referencia utilizados.

### 5.3. Resumen

Con el fin de facilitar la comparación entre todos los métodos y modelos analizados en esta investigación, en la Tabla 5.5 se presentan las principales métricas obtenidas para cada modelo. Así, tal como se observó en la sección anterior, resulta claro que el modelo JANet-ConvJANet es el que presenta el mejor desempeño de pronóstico de entre todas las arquitecturas propuestas.

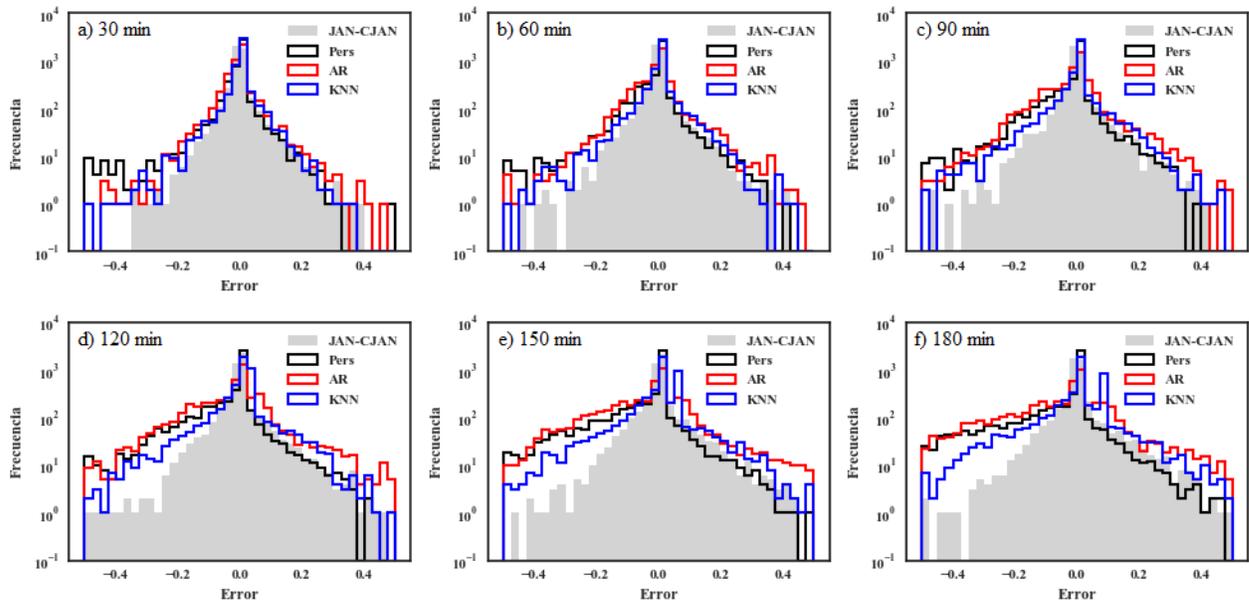


Figura 5.17: Distribución del error de pronóstico del modelo JANet-ConvJANet para cada uno de los intervalos del horizonte de pronóstico en días parcialmente nublados.

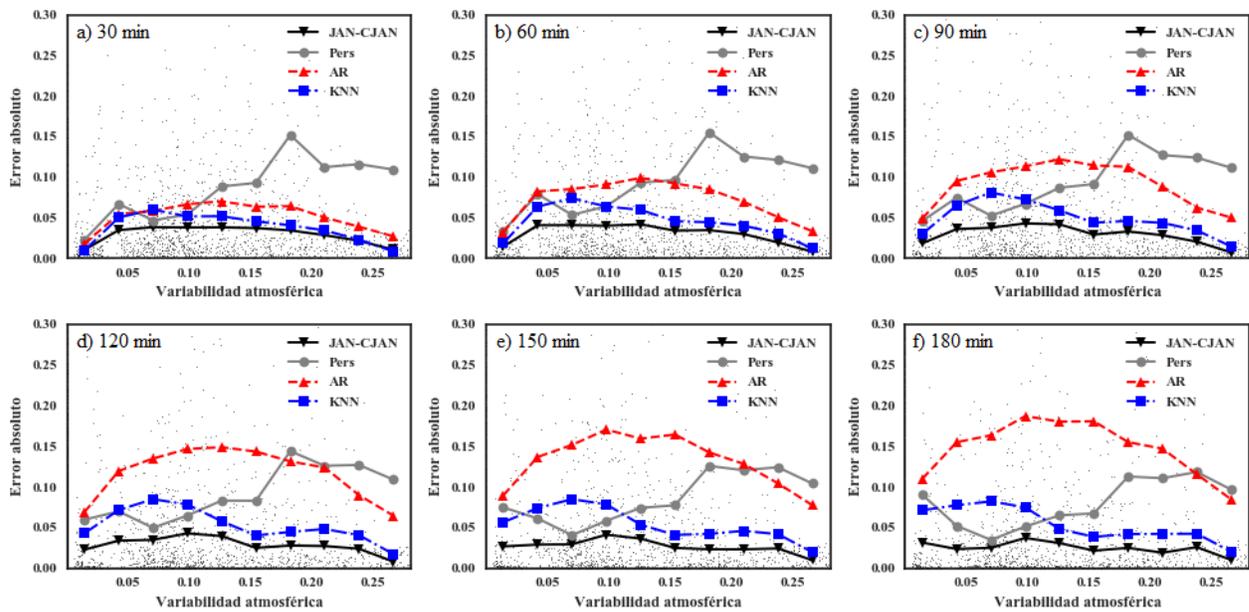


Figura 5.18: Error absoluto promedio del modelo JANet-ConvJANet versus la variabilidad atmosférica para cada uno de los intervalos del horizonte de pronóstico en días parcialmente nublados.

Tabla 5.5: Métricas estadísticas de evaluación para cada uno de los métodos empleados

Método	Soleados			Nublados			Nublados parciales		
	MAE	RMSE	FS%	MAE	RMSE	FS%	MAE	RMSE	FS%
LSTM	0.0127	0.0266	66.131	0.0255	0.0544	3.282	0.0291	0.0647	56.881
JANet	0.0144	0.0279	64.454	0.0230	0.0501	10.910	0.0270	0.0593	60.491
JANet-ConvJANet	<b>0.0113</b>	<b>0.0245</b>	<b>68.787</b>	<b>0.0214</b>	<b>0.0468</b>	<b>16.754</b>	<b>0.0238</b>	<b>0.0554</b>	<b>63.091</b>
Persistencia	0.0360	0.0785	0.000	0.0253	0.0563	0.000	0.0630	0.1500	0.000
AR	0.0587	0.0994	-26.745	0.0457	0.0815	-44.932	0.0762	0.1326	11.615
KNN	0.0268	0.0539	31.360	0.0377	0.0669	-18.942	0.0429	0.0833	44.468

# Capítulo 6

## Conclusiones

En cuanto a la capacidad de pronóstico de los modelos evaluados frente a los distintos contextos meteorológicos analizados, en general, se obtiene que las tres arquitecturas propuestas: LSTM, JANet y JANet-ConvJANet, logran los mejores resultados sobre el conjunto de días soleados, presentando un desempeño muy superior al de los métodos estocásticos. Dado que este tipo de días se caracterizan por curvas de irradiación suaves y de poca variabilidad en cuanto a condiciones atmosféricas, los modelos son capaces de estimar sin mayores problemas el perfil de irradiación en el horizonte de pronóstico.

Por otro lado, con respecto a los días nublados y de nubosidad parcial, si bien la diferencia en el desempeño de los modelos en este conjunto de días no es del todo considerable y en ambos casos las métricas son considerablemente superiores a las obtenidas mediante métodos estocásticos, los resultados evidencian un mejor desempeño de pronóstico en el conjunto de días nublados en los tres casos estudiados. De este modo, aunque ninguna de las arquitecturas analizadas presenta la capacidad de estimar cambios abruptos en el horizonte de pronóstico, dado que los días nublados presentan una relativa menor variabilidad atmosférica, el error de la distribución de pronóstico resulta menor en este caso. En este sentido, la mayor dificultad que presentan los modelos consiste en la capacidad de predecir correctamente el grado de nubosidad que se tendrá en las horas futuras a partir de la información recibida de las horas previas. Así, en ocasiones los modelos no son capaces de discernir a partir de los datos de entrada si el horizonte de pronóstico próximo corresponde a un intervalo de nubosidad parcial o completa, provocando subestimaciones/sobrestimaciones erróneas en el pronóstico. Esto se evidencia en mayor medida en las arquitecturas LSTM y JANet, pues estas operan principalmente con los datos registrados en el sitio.

Ahora bien, entre todas las arquitecturas evaluadas, la JANet-ConvJANet es la que presenta las mejores métricas estadísticas con respecto a la distribución del error de pronóstico, superando al resto de los modelos y métodos estocásticos en cada uno de los distintos contextos meteorológicos estudiados. En este sentido, dado que estructuralmente la única diferencia entre el modelo JANet y el JANet-ConvJANet es la incorporación de la red ConvJANet que procesa las imágenes satelitales, se comprueba la capacidad de esta última para extraer e interpretar los patrones de la nubosidad en el entorno del sitio estudiado a partir de las secuencias de imágenes entregadas.

Del mismo modo, y consecuentemente, se verifica que las secuencias de imágenes sate-

litas sí proveen de información adicional relevante a la red, posibilitando así las mejoras evidenciadas en los pronósticos de generación. En particular, es en los días nublados y de nubosidad parcial en donde se distinguen la mayores mejoras en la distribución del error de pronóstico, pues es precisamente en estos casos donde poseer información acerca de la evolución de la nubosidad en la vecindad resulta de mayor utilidad para el modelo. Así, este último es capaz de combinar los patrones y representaciones latentes de las series temporales de los datos operacionales con la información adicional proveniente de la estructura ConvJANet para ajustar las estimaciones de generación en el horizonte de pronóstico de mejor manera.

Con respecto al comportamiento que exhiben las distribuciones del error de pronóstico en los distintos modelos analizados, se observa que todos presentan un aumento en la dispersión de las estimaciones, con respecto a los valores reales de potencia, a medida que el intervalo temporal de pronóstico aumenta. En este aspecto, ninguno de los modelos propuestos es capaz de reducir o controlar efectivamente este tipo de incertidumbre, aunque sí se evidencia que en el conjunto de días soleados este comportamiento es notoriamente menor. Por otro lado, considerando todos los modelos de redes neuronales evaluados, se observa que el *bias* o Skewness de la distribución del error de pronóstico no es igual en todos los casos. En particular, se evidencia que en el conjunto de datos de días soleados existe una ligera tendencia a subestimar la potencia, mientras que en los días nublados y de nubosidad parcial se tienden a sobrestimar los pronósticos de potencia generada por el sistema. Como se mencionó anteriormente, dado que ninguno de las arquitecturas propuestas es capaz de estimar cambios abruptos en el horizonte de pronóstico, estas subestimaciones/sobrestimaciones resultan producto de la diferencia entre las curvas estimadas por los modelos y el perfil real de generación registrado que, dependiendo del caso, presentará más o menos fluctuaciones.

Por otro lado, dentro de este último aspecto, cabe destacar que, al comparar los modelos propuestos con las distribuciones del error de pronóstico obtenidas mediante los métodos estocásticos tradicionales, se evidencia que las arquitecturas mantienen distribuciones considerablemente menos dispersas y con un sesgo bastante más reducido en los distintos conjuntos de datos observados. De esta manera, si bien los modelos no logran reducir del todo la incertidumbre en el horizonte de pronóstico, estos presentan una mejora importante respecto a los métodos tradicionales analizados.

Finalmente, al analizar la correlación entre el error promedio de pronóstico y la variabilidad atmosférica, se observa que las tres arquitecturas, incluso bajo los distintos tipos de días analizados, presentan un error promedio que se mantiene relativamente constante a lo largo de todo el rango de variabilidad atmosférica, evidenciándose algunos leves incrementos a medida que la variabilidad aumenta, principalmente en los conjuntos de datos de días nublados y de nubosidad parcial. Este comportamiento se contrasta bastante con lo observado en los modelos estocásticos también evaluados, donde tanto el método de persistencia como el de autoregresión presentan un significativo incremento en el error a medida que la variabilidad aumenta. Dada la gran cantidad de parámetros y variables que son procesadas e interpretadas por las redes, a diferencia de los métodos estocásticos que operan con solo una o dos variables, los modelos propuestos son capaces de controlar de mejor manera la incertidumbre producida por la variabilidad atmosférica en los datos de entrada. De esta manera, mantienen un error promedio relativamente bajo y constante dentro de todo el rango de variabilidad.

# Bibliografía

- [1] IEA: International Energy Agency. *Renewables 2017: Analysis and Forecasts to 2022*. OECD, 2017.
- [2] F. Wang, Z. Mi, S. Su, and H. Zhao. Short-term solar irradiance forecasting model based on artificial neural network using statistical feature parameters. *Energies*, 5(5):1355–1370, 2012.
- [3] G. Kariniotakis. *Renewable Energy Forecasting: From Models to Applications*. Woodhead Publishing, 2017.
- [4] L. Mazorra, B. Pereira, M. David F. Díaz, and P. Lauret. Use of satellite data to improve solar radiation forecasting with bayesian artificial neural networks. *Solar Energy*, 122:1309–1324, 2015.
- [5] J. Park J. Kim Y. Kim W. Lee, K. Kim. Forecasting solar power using long-short term memory and convolutional neural networks. *IEEE Access*, 6:73068 – 73080, 2018.
- [6] M. Abdel-Nasser and K. Mahmoud. Accurate photovoltaic power forecasting models using deep lstm-rnn. *Neural Computing and Applications*, 31:2727–2740, 2019.
- [7] J. Duffie and W. Beckman. *Solar Engineering of Thermal Processes*. Wiley Sons, 4th edition, 2013.
- [8] C. Coimbra and H. Pedro. Assessment of forecasting techniques for solar power production with no exogenous inputs. *Solar Energy*, 86(7):2017–2028, 2012.
- [9] J. Patterson and A. Gibson. *Deep Learning: A practitioner’s Approach*. O’Reilly, 1st edition edition, 2017.
- [10] Z. C. Lipton, J. Berkowitz, and C. Elkan. A critical review of recurrent neural networks for sequence learning, 2015.
- [11] Jos van der Westhuizen and Joan Lasenby. The unreasonable effectiveness of the forget gate, 2018.
- [12] X. Shi, Z. Chen, H. Wang, D. Yeung, W. Wong, and W. Woo. Convolutional lstm network: A machine learning approach for precipitation nowcasting, 2015.
- [13] Y. Chu, B. Urquhart, S. Gohari, H. Pedro, J. Kleissl, and C. Coimbra. Short-term reforecasting of power output from a 48 mwe solar pv plant. *Solar Energy*, 112:68–77, 2015.
- [14] S. Srivastava and S. Lessmann. A comparative study of lstm neural networks in forecasting day-ahead global horizontal irradiance with satellite data. *Solar Energy*, 162:232–247, 2018.

- [15] T. J. Schmit, S. S. Lindstrom, J. J. Gerth, and M. M. Gunshor. Applications of the 16 spectral bands on the advanced baseline imager (abi). *Journal of Operational Meteor*, 6(4):33–46, 2018.
- [16] P. Govender, M. J. Brooks, and A. P. Matthews. Cluster analysis for classification and forecasting of solar irradiance in durban, south africa. *Journal of Energy in Southern Africa*, 29(2):51–62, 2018.