

# PRONÓSTICO DE LA IRRADIACIÓN SOLAR HORARIA EN URUGUAY Y SUR DE BRASIL UTILIZANDO REDES NEURONALES RECURRENTE

**Marcelo Coletto** – marcelo.coletto@utec.edu.uy

Laboratorio Solar, ITR Centro-Sur, UTEC, Uruguay

**Rodrigo Alonso-Suárez** – rodrigoa@fing.edu.uy

Laboratorio de Energía Solar, Facultad de Ingeniería, UDELAR, Uruguay

**Resumen.** El aumento de la penetración de la energía solar fotovoltaica (PV) en las redes eléctricas requiere el uso de técnicas de pronóstico de la irradiación solar tendientes a un despacho eficiente de la generación, optimizando costos y valorizando mejor los excedentes para exportación. En este trabajo evaluamos el desempeño del pronóstico de la irradiación solar a partir del empleo de Redes Neuronales Recurrentes (RNN) del tipo Long Short Term Memory (LSTM), para la región del Uruguay y Rio Grande do Sul (RS). Se aborda el pronóstico horario de irradiación solar hasta 6 horas en adelante. Se evalúa el desempeño de las RNN con distintas entradas, utilizando el índice de cielo claro ( $k_c$ ) en el primer modelo y agregando datos de reflectancia planetaria por satélite (ST) para el segundo modelo. Se encontró, para la predicción de 1 a 6 horas, un rRMSD promedio en la región que varía entre 20,1 % y 39,2 % para el modelo con  $k_c$  y entre 18,6 % y 39,0 % para el modelo con  $k_c$  y ST. Se observa que incluir la información satelital mejora el desempeño del pronóstico sólo hasta 3 horas en adelante. Esta limitación está asociada al tamaño de la celda utilizada para el promediado espacial de la reflectancia satelital. El desempeño de ambos modelos supera el de la mejor persistencia inteligente, demostrando que se trata una técnica útil para la predicción de irradiación solar en la región.

**Palabras clave:** pronóstico solar, RNN-LSTM, imágenes de satélite.

## 1. INTRODUCCIÓN

Uruguay destaca en el mundo por su transición hacia una matriz energética con alto contenido de energías renovables (IRENA, 2019). En la actualidad, la matriz de generación eléctrica es cubierta en un 97 % por estas energías, con sólo un 3 % de respaldo de generación térmica por combustibles fósiles. La energía eólica y solar fotovoltaica (PV) presentan niveles importantes de penetración, representando respectivamente un 34 % y 3 % de la generación anual<sup>1</sup>. En particular, la capacidad instalada de energía solar PV creció en los años 2015-2018 de 65 MW a 230 MW, lo que corresponde a la fecha a un 5 % de la capacidad instalada del país (ADME, 2018). La planificación óptima de la expansión de la red eléctrica da como resultado que la capacidad de generación solar se incremente en forma significativa en los próximos 5 años (Cornalino and Chaer, 2018). Una alta penetración del recurso solar complejiza el proceso de despacho de la energía debido a las fluctuaciones provocadas por la meteorología local. Una forma de mitigar el impacto de la variabilidad del recurso es el uso de herramientas de pronóstico de la irradiación solar. Un pronóstico de menor incertidumbre permite una planificación del despacho más eficiente, reduciendo costos y garantizando una mejor comercialización de los excedentes.

Pronosticar la irradiación solar a nivel de superficie es un problema de alta complejidad. Esta complejidad está relacionada a la dificultad que existe para anticipar el movimiento, formación y deformación de las nubes (Coimbra et al., 2013). Se han desarrollado distintos modelos para la predicción de irradiación solar, los cuales se pueden dividir en cuatro grupos principales (Diagne et al., 2013): modelos numéricos de atmósfera (NWP), pronóstico utilizando imágenes de satélites geoestacionarios, pronóstico utilizando cámaras de cielo ubicadas en Tierra y pronóstico por técnicas estadísticas o de aprendizaje automático a partir de distintos datos de entrada. Este último grupo de técnicas, conocidas como de análisis de series temporales, son las que tienen relación con el presente trabajo. Dentro de las técnicas de análisis de series temporales para pronóstico de la irradiación solar destacan el modelado de procesos auto-regresivos (ARMA) (Box and Jenkins, 1970) y las redes neuronales artificiales (RNA) (Mellit and Kalogirou, 2008). Estos trabajos demuestran que el desempeño

<sup>1</sup>Información extraída del sitio web del Administrador del Mercado Eléctrico (ADME, Uruguay): <https://adme.com.uy/>

de los modelos depende, además del horizonte de pronóstico, de la elección de la arquitectura del modelo y de los parámetros de configuración.

El presente trabajo se enfoca en el pronóstico de la irradiación solar horaria hasta 6 horas en adelante empleando un modelo de redes neuronales recurrentes (RNN, Recurrent Neural Networks) del tipo Long Short Term Memory (LSTM). Se evalúa el desempeño de la técnica para la región de Uruguay y el sur de Brasil, utilizando como entrada las propias medidas registradas en Tierra e información satelital. El documento está organizado de la siguiente manera. La Sección 2 presenta una descripción de la base de datos utilizada. La Sección 3 describe la metodología implementada, el procesamiento de los datos para su ajuste y validación y las métricas de desempeño. La Sección 4 discute los resultados y la Sección 5 resume las principales conclusiones.

## 2. DATOS

Se utilizaron datos horarios de irradiación global horizontal (GHI) registrados en cinco sitios distribuidos en el territorio de Uruguay y un sitio en Rio Grande do Sul (Brasil). Las estaciones ubicadas en Uruguay son parte de la Red de Medida Continua de Irradiancia Solar (RMCIS) que administra el Laboratorio de Energía Solar de la UdelaR (LES, <http://les.edu.uy>). La estación ubicada en el sur de Brasil integra el Sistema de Organização Nacional de Dados Ambientais (SONDA, <http://http://sonda.ccst.inpe.br>), la cual es administrada por el Instituto Nacional de Pesquisas Espaciais (INPE). La ubicación de estos sitios se muestra en la Figura 1 y se detalla en la Tabla 1. La Tabla 1 indica también el periodo temporal de datos considerado para cada estación, siendo todos periodos mayores a 2 años completos. La región bajo estudio tiene un clima templado sub-tropical con características climáticas homogéneas (Pampa Húmeda) y su recurso solar presenta una variabilidad inter-anual intermedia (Alonso-Suárez, 2017). Las características climáticas son semejantes a las encontradas en la cuenca del Río Mississippi, en algunas zonas de la costa oeste de EEUU y en varios estados de la zona centro-este de dicho país (Gueymard and Wilcox, 2011).

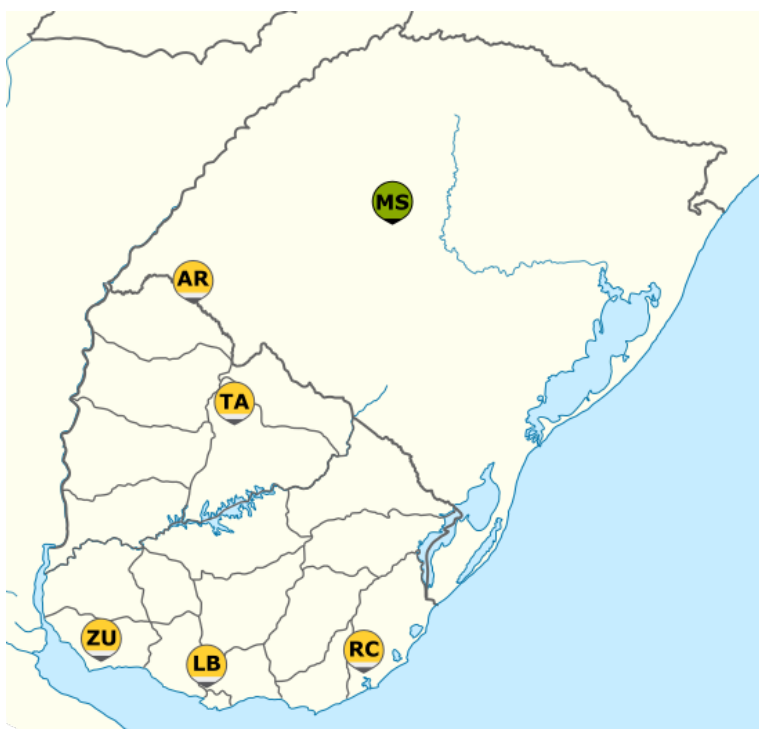


Figura 1- Ubicación de las estaciones de medida.

La RMCIS está equipada con piranómetros Kipp & Zonen de primera clase o superior para la medida de GHI, que son calibrados cada dos años contra un estándar secundario que el LES mantiene con trazabilidad al patrón primario mundial en el World Radiation Center (WRC - Davos, Suiza). Las estaciones están ubicadas en predios rurales o semi-urbanos, y sus piranómetros reciben mantenimiento de limpieza de cúpulas en base mensual. La

Tabla 1- Detalle de la ubicación de las estaciones de medida.

Estación	Cód	Lat	Lon	Alt	Período
Artigas	AR	-30.40	-56.51	136	01/2015 - 12/2017
Tacuarembó	TA	-33.71	-55.83	140	01/2016 - 12/2017
Colonia	ZU	-34.34	-57.69	81	01/2016 - 12/2017
Canelones	LB	-34.67	-56.34	32	01/2015 - 12/2017
Rocha	RC	-34.49	-56.17	24	01/2015 - 12/2017
São Martinho da Serra	MS	-29.44	-53.82	489	01/2005 - 12/2015

red SONDA cuenta con 20 estaciones distribuidas en el territorio brasileño, las cuales son clasificadas en tres grupos según los equipos instalados en cada estación (Solarimétricas, Anemométricas o ambos equipamientos). De esta red, se utilizaron los datos de la estación de São Martinho da Serra (MS), ubicada en el centro de Rio Grande do Sul. Esta estación cuenta con un tracker de alta precisión para la medida de las tres componentes de la radiación solar (global, directa y difusa) y equipos con calidad de estándar secundario. La estación integra la Red de Estaciones de Superficie para Medición de Radiación – “Baseline Surface Radiation Network” (BSRN) bajo el código SMS y fue seleccionada para este trabajo por la calidad de sus medidas y porque se encuentra cercana al territorio de Uruguay, a unos 280 km de la estación Artigas (AR).

Además de los datos de GHI registrados en Tierra, se exploró la utilización de las imágenes del canal visible del satélite geostacionario GOES-East como entrada a las RNN. Se utilizó la información de albedo terrestre ( $\rho$ ) promediado en una celda espacial de  $10 \text{ min} \times 10 \text{ min}$  en latitud y longitud, que equivale aproximadamente en la región a unos  $16 \text{ km} \times 18 \text{ km}$ . El promedio espacial contiene información de la nubosidad en los alrededores del sitio, no sólo sobre el punto específico, brindando una indicación indirecta de la nubosidad futura al corto plazo. Las imágenes GOES-East para la región fueron registradas por el satélite con cadencia irregular hasta el año 2017 inclusive (satélite GOES12 y GOES13), usualmente disponibles cada media hora, pero con periodos horarios y tri-horarios. La información satelital con cadencia horaria se generó vía interpolación lineal de la serie temporal irregular de satélite, obteniendo una serie satelital en concurrencia temporal con las integrales horarias de GHI. No se interpolaron huecos de satélite de más de dos horas consecutivas.

### 3. METODOLOGÍA

En esta Sección se presenta el procesamiento de los datos para actuar como entrada a los modelos de predicción, las RNN implementadas y las métricas de desempeño a utilizar. Se indica la separación de los datos para entrenamiento y validación de las RNN, y las configuraciones y parámetros utilizados para su ejecución.

#### 3.1. Procesamiento de datos

Los datos de cada estación fueron divididos en 50 % para entrenamiento y 50 % para validación. Es importante mencionar que las horas nocturnas son filtradas de los datos, dejando sólo los datos que cumplen  $\cos \theta_z > 0.1$ . La GHI no se utiliza directamente como entrada. Previamente, se normaliza por la irradiación global horaria de cielo despejado, obteniéndose el índice de cielo claro ( $k_c$ ). Esta magnitud no presenta las variaciones geométricas observadas en la serie de GHI debido al movimiento aparente del Sol. En cambio, sus fluctuaciones representan solamente las variaciones debido a la nubosidad. Para el cálculo del índice de claridad se utilizaron las estimaciones de la irradiación horaria de cielo despejado del modelo McClear (Lefèvre et al., 2013), públicamente disponible en el CAMS (Copernicus Atmosphere Monitoring Service, <http://www.soda-pro.com/>). Desde el punto de vista estadístico, la serie de  $k_c$  es más cercana a ser una señal estocástica estacionaria que la serie de GHI, y por tanto es más adecuada como entrada para las RNN. La señal  $\rho$  de albedo terrestre ya se encuentra normalizada por las variaciones geométricas de la posición del Sol y por tanto cumple también con esta característica.

### 3.2. Pronostico por modelo RNN-LSTM

Las RNN tienen una arquitectura tal que la información previamente computada en los estados de las neuronas se retro-alimenta en la siguiente iteración. Este tipo de red no fue históricamente de uso extendido debido a su alto costo computacional, pero actualmente es uno de los tipos de redes más utilizados para reconocer y modelar secuencias (Buduma and Locascio, 2017). Las redes del tipo Long-Short Term Memory (LSTM), propuesta por Hochreiter and Schmidhuber (1997), es un tipo de red recurrente. La principal diferencia entre el modelo LSTM y el RNN tradicional es el uso de celdas de memoria y compuertas de entrada y olvido. Esta estructura en forma de compuertas permite retener la información a través de varios pasos temporales. Así, el modelo LSTM es capaz de superar el problema de la fuga por gradiente (Vanishing Gradient) que se produce en la mayoría de las redes recurrentes tradicionales. En la literatura distintos trabajos hacen el uso de redes LSTM para predecir la radiación solar en distintos horizontes de pronóstico (Abdel-Nasser and Mahmoud, 2019; Gensler et al., 2016; Qing and Niu, 2018; Srivastava and Lessmann, 2018).

En este trabajo utilizamos las redes RNN-LSTM para predecir el índice de claridad futuro,  $\hat{k}_c(t + \Delta t)$ , para horizontes de pronóstico  $\Delta t$  de 1 a 6 horas. Las entradas son la propia serie de índice de claridad considerando el valor en tiempo ( $t$ ) y sus valores pasados ( $t - i$ ) con  $i$  hasta 6 horas, e información satelital a tiempo ( $t$ ) promediada espacialmente. La elección de los parámetros de configuración de las RNN fueron determinados en forma heurísticas. El espacio de parámetros y configuraciones es de alta dimensionalidad, por lo que la búsqueda en una grilla de parámetros es costosa computacionalmente. No obstante, en nuestra experiencia, a partir de cierto punto el desempeño no es afectado críticamente por un valor u otro. Como lo que buscamos es cuantificar la diferencia de desempeño al utilizar dos conjuntos de entradas distintos, la elección óptima de la red con alta precisión no es crítica. Las configuraciones utilizadas para la ejecución de los modelos se presentan en la Tabla 2. La configuración de la red RNN-LSTM a utilizar es la misma en ambos casos y la diferencia entre ambos modelos empleados está sólo en el número de entradas utilizadas para la predicción, además del entrenamiento de sus parámetros. Se entrena una RNN distinta para cada sitio y para cada horizonte de pronóstico.

Tabla 2- Configuración de los parámetros de la redes LSTM.

Parámetros	
Neuronas	16
Capas ocultas	1
Función de activación	Rectified Linear Unit (ReLU)
Optimizador	Adam (Kingma and Ba, 2014)
Métrica de desempeño	Mean Absolut Error (MAE)
Tamaño del lote	128
Numero de épocas	30

### 3.3. Métricas de desempeño

La evaluación de desempeño se realizó utilizando las métricas comunes en el área: el desvío promedio (MBD), el desvío cuadrático medio (RMSD) y la habilidad de pronostico (FS, Forecasting Skill). Este último indicador compara el RMSD del pronóstico con el de la persistencia, resultando en una métrica de utilidad que permite intercomparar el desempeño de la predicción ante diferentes condiciones de variabilidad del recurso (Coimbra et al., 2013). La comparación entre la predicción y la medida se realiza a nivel de GHI, por lo que el pronóstico del índice de cielo claro debe ser convertido a GHI utilizando la estimación de cielo claro.

El MBD y RMSD se definen según,

$$\text{MBD} = \frac{1}{N} \sum_{i=1}^N (\hat{I}_h - I_h), \quad (1)$$

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{I}_h - I_h)^2}, \quad (2)$$

donde  $N$  es la cantidad de horas,  $\hat{I}_h$  es la irradiación pronosticada y  $I_h$  es la irradiación medida en tierra, ambas expresadas en  $\text{Wh/m}^2$ . Los valores porcentuales de estas cantidades, rMBD y rRMSD, se calculan relativos a la media de las medidas,  $\bar{I}_h$ . Sólo se utilizan los datos donde la medida de referencia y la predicción del modelo estén ambas disponibles.

La habilidad de pronóstico, FS, se define como,

$$\text{FS} = 1 - \frac{\text{RMSD}_m}{\text{RMSD}_p}, \quad (3)$$

a partir del RMSD del modelo ( $\text{RMSD}_m$ ) y el de la persistencia ( $\text{RMSD}_p$ ). Expresaremos el FS en términos porcentuales, multiplicando a la expresión de la Ec. (3) por 100. Este indicador es  $\text{FS} = 100\%$  si el pronóstico es perfecto y es  $\text{FS} \leq 0\%$  si su desempeño es igual o peor que el de la persistencia. Como la persistencia es un procedimiento simple para establecer una predicción, su desempeño se utiliza como nivel de referencia para técnicas de pronóstico de mayor complejidad.

La definición de FS supone definir un mecanismo para la persistencia del recurso. Conceptualmente, la persistencia es establecer que las condiciones meteorológicas, en particular las asociadas a la nubosidad, permanecen constantes en el tiempo para los pasos temporales posteriores. Para irradiación solar la persistencia normalmente se implementa a partir del índice de cielo claro, estableciendo que  $k_c(t + \Delta t) = k_c(t)$ . Es posible también establecer la persistencia como promedio de muestras pasadas, de la forma,

$$k_c(t + \Delta t) = \frac{\sum_{j=1}^n k_c(t - (j - 1))}{n}, \quad (4)$$

donde  $n$  es la cantidad de pasos temporales hacia el pasado que son considerados. Este procedimiento se lo conoce usualmente como persistencia 'inteligente' (Smart Persistence), que denominaremos PSTM. Para  $n = 1$  se recupera la persistencia común. Cuanto mayor es el valor de  $n$ , más parecido es el valor del  $k_c$  persistido a su valor climatológico ( $n \gg 1$ ), nivel referencia para el desempeño que se utiliza con frecuencia en el campo de la meteorología. En este trabajo, buscamos utilizar la 'mejor' PSTM como referencia de desempeño, utilizando el mejor valor de  $n \in [1, 6]$  para cada horizonte  $\Delta t$ . Este análisis se reporta en la Sección 4.1.

## 4. RESULTADOS

En esta Sección se presenta y compara el desempeño del procedimiento PSTM y los modelos RNN-LSTM para la predicción horaria de la irradiación solar hasta 6 horas en adelante.

### 4.1. Persistencia

Se implementó el procedimiento PSTM para cada estación por separado, variando el valor de  $n$  de 1 a 6 horas hacia atrás. Por simplicidad, se presenta en la Figura 2 la curva de desempeño rRMSD promedio sobre todas las estaciones. En gris se identifican las curvas de cada PSMT para los distintos valores de  $n$  (i.e. PSTM2 refiere a PSTM con  $n = 2$ ) y en rojo se identifica la mejor curva de desempeño. Se observa que, con granularidad horaria, la PSTM1 es la de mejor desempeño hasta 4 horas en adelante, comportamiento que tiene un punto de quiebre a partir de las 5 horas. Para el horizonte de 6 horas, el mejor desempeño se tiene con la PSTM6, el mayor valor de  $n$  considerado en este trabajo. En suma, los menores valores de  $n$  son mejores para los primeros horizontes de pronóstico y los mayores valores de  $n$  para los últimos horizontes de pronóstico. La curva identificada en rojo (mejor persistencia) es la que se utilizará en lo siguiente para el cálculo del FS.

### 4.2. Desempeño de los modelos en la región

El ajuste de modelos y evaluación de desempeño se realiza para cada estación y horizonte de pronóstico por separado. En el Apéndice se presentan los resultados completos para cada estación y horizonte (Tablas 4

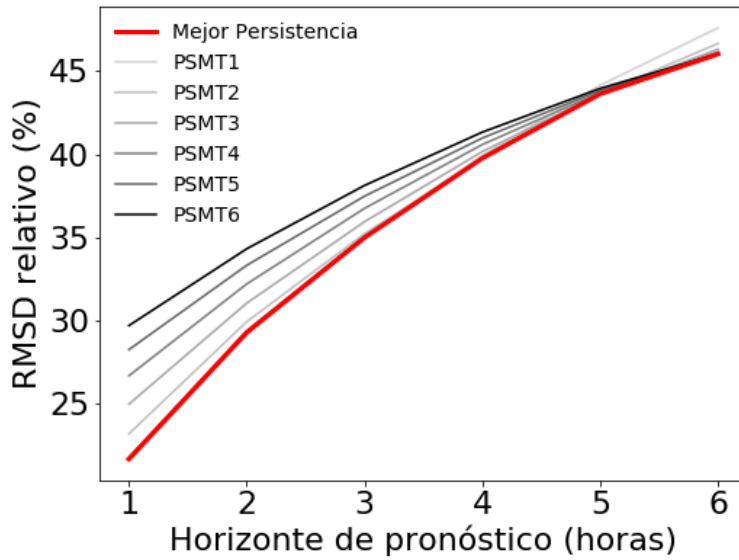


Figura 2- Desempeño de la persistencia inteligente para la región.

y 5). Similar a lo anterior, aquí se discute el desempeño promedio en la región (promedio inter-estaciones), resumido en la Figura 3 y la Tabla 3. Se presenta el modelo RNN-LSTM utilizando 6 valores pasados de  $k_c$  y el modelo RNN-LSTM utilizando esta información más la adición de la información satelital  $\rho$  en tiempo ( $t$ ). La diferencia de desempeño entre ambos modelos se observa sólo para los menores horizontes de pronóstico (de 1 a 3 horas en adelante). Si bien la ganancia de incorporar información satelital en la predicción en estos horizontes es moderada en rRMSD (Figura 3a), es más observable en FS (Figura 3b). Para los mayores horizontes de pronóstico (de 4 a 6 horas) el efecto de agregar la información satelital es despreciable en el desempeño de la técnica RNN-LSTM. El incremento de rRMSD con el horizonte de pronóstico es significativo, variando entre 20.1 % y 39.2 % para el modelo con  $k_c$  y entre 18.6 % y 39.0 % para el modelo con  $k_c$  y  $\rho$ , entre 1 y 6 horas respectivamente. En ambos casos, la incertidumbre de los dos modelos se encuentran por debajo de la mejor PSTM para la región, lo que representa una ganancia de desempeño respecto al nivel de referencia. Esta ganancia (FS) varía entre 20 % y 15 % a 1 hora, para el modelo con y sin satélite respectivamente, y decrece en forma monótona hasta una ganancia de 9 % a 6 horas, para ambos modelos.

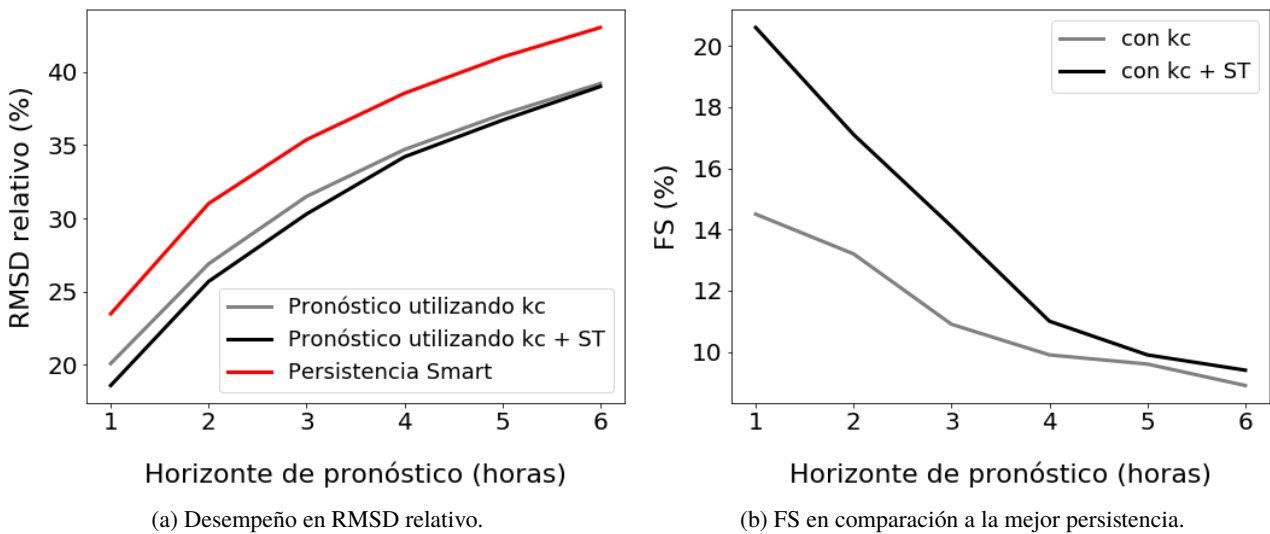


Figura 3: Comparación de desempeño entre los distintos modelos de pronóstico (distintas entradas).

Tabla 3: Indicadores de desempeño promedio para la región del pronóstico por RNN-LSTM. El rMBD y rRMSD están normalizados respecto a la media de las medidas de 449 Wh/m<sup>2</sup>.

(a) RNN-LSTM sólo con entrada de $k_c$ .						(b) RNN-LSTM con entrada de $k_c$ y $\rho$ (ST).					
$\Delta t$ (horas)	MBD		RMSD		FS	$\Delta t$ (horas)	MBD		RMSD		FS
	(Wh/m <sup>2</sup> )	(%)	(Wh/m <sup>2</sup> )	(%)	(%)		(Wh/m <sup>2</sup> )	(%)	(Wh/m <sup>2</sup> )	(%)	(%)
1	1,7	0,4	90,2	20,1	14,5	1	5,81	1,3	83,7	18,6	20,6
2	8,1	1,8	120,9	26,9	13,2	2	11,5	2,6	115,8	25,7	17,1
3	11,5	2,6	141,6	31,5	10,9	3	5,2	1,2	136,8	30,3	14,1
4	22,6	5,0	155,9	34,7	9,9	4	24,2	5,4	153,9	34,2	11,0
5	25,2	5,6	166,6	37,1	9,6	5	21,6	4,8	165,3	36,7	9,9
6	27,1	6,0	176,1	39,2	8,9	6	31,8	7,1	175,1	39,0	9,4

El la Figura 4 se presentan los diagramas de dispersión entre la medida de GHI y su pronóstico por el modelo que incluye información satelital para la estación de São Martinho da Serra (estación con más cantidad de datos). Los 3 gráficos son para distintos horizontes de pronóstico. Se puede apreciar que con el aumento del horizonte de pronóstico hay un una mayor dispersión, con más puntos alejados de la diagonal  $x = y$  (línea de pronóstico perfecto). En la predicción a 3 y 5 horas hacia adelante, se puede notar una mayor cantidad de puntos en la parte superior de la diagonal, representando que las horas fueron pronosticadas, en promedio, con mayor sobrestimación. Esto es consistente con los indicadores de MBD positivos observados en la Tabla 3, que aumentan con el horizonte de pronóstico.

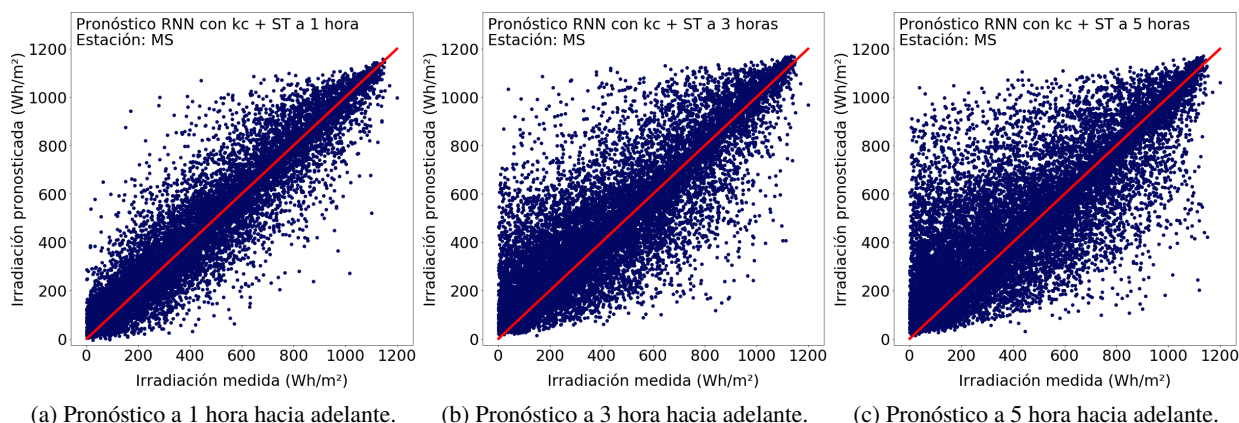


Figura 4: Diagramas de dispersión entre la medida (eje x) y el pronóstico (eje y) para el modelo que incluye información satelital ( $k_c$  y  $\rho$ ). Datos de la estación de São Martinho da Serra.

## 5. CONCLUSIONES

Se presentaron los resultados obtenidos para el pronóstico de GHI en la región (Uruguay y Rio Grande do Sul) utilizando la técnica de redes neuronales recurrentes del tipo LSTM con distintas variables de entrada. Se presentó y comparó el desempeño de estos modelos al utilizar únicamente la serie de índice de cielo claro y esta serie con la adición de la reflectancia planetaria por satélite. Ambos modelos, ya sea que incluyan o no la información satelital, superan el desempeño del procedimiento de persistencia para la región. En la comparación entre los modelos, se observa que la información satelital introduce una ganancia moderada para las primeras 3 horas de pronóstico. Para los horizontes de pronóstico mayores a 4 horas, la introducción de esta información resulta en una ganancia muy pequeña, que no compensa el aumento de complejidad debido a la necesidad de contar operacionalmente con las imágenes de satélite para realizar el pronóstico. La utilidad limitada de la

información satelital, sólo hasta 3 horas en adelante, esta relacionada con el tamaño de la celda utilizada para el promediado espacial. Este tamaño es el óptimo en la región para la estimación horaria de la irradiación solar por satélite, y no está adaptado para el pronóstico a diferentes horizontes temporales. En ese sentido, es razonable que la información satelital promediada de este manera reporte mayor utilidad en los primeros horizontes de pronóstico horario.

El trabajo a futuro incluye el análisis de desempeño incluyendo otras variables de entrada, como variables geométricas (el coseno del ángulo cenital o el ángulo horario) o información satelital con mayor representatividad espacial. Asimismo, es necesario explorar en forma más intensiva el espacio de parámetros y configuraciones que presentan las redes RNN-LSTM. Es necesario seguir explorando la correcta combinación de modelos, configuraciones y datos de entrada para mejorar la predicción de irradiación solar. Las RNN del tipo LSTM son una buena herramienta a considerar para este objetivo, y el impacto de las distintas variables de entrada no es necesariamente el mismo en esta técnica que en otras.

## APÉNDICE

Esta Sección presenta la evaluación de desempeño por estación, para los modelos RNN-LSTM utilizando la serie de  $k_c$  y la serie de  $k_c$  y  $\rho$  como entradas, en la Tablas 4 y 5 respectivamente.

Tabla 4: Indicadores de desempeño utilizando el modelo con  $k_c$  como única entrada.

(a) Artigas						(b) Tacuarembó					
$\Delta t$ (hrs)	MBD		RMSD		FS (%)	$\Delta t$ (hrs)	MBD		RMSD		FS (%)
	(Wh/m <sup>2</sup> )	(%)	(Wh/m <sup>2</sup> )	(%)			(Wh/m <sup>2</sup> )	(%)	(Wh/m <sup>2</sup> )	(%)	
1	-2,3	≈0	88,2	19,0	3,9	1	-5,9	-1,3	92,3	20,3	1,6
2	14,1	3,0	120,1	25,9	5,3	2	-1,6	≈0	121,2	26,7	7,0
3	11,5	2,5	140,0	30,1	8,9	3	7,5	1,7	139,5	30,8	11,2
4	8,4	1,8	153,7	33,1	11,7	4	9,2	2,0	158,8	35,0	7,8
5	33,0	7,1	169,5	36,4	10,9	5	17,2	3,8	164,5	36,3	9,7
6	26,0	5,6	174,7	37,5	13,9	6	14,3	3,2	171,0	37,7	9,7

(c) Colonia						(d) Canelones					
$\Delta t$ (hrs)	MBD		RMSD		FS (%)	$\Delta t$ (hrs)	MBD		RMSD		FS (%)
	(Wh/m <sup>2</sup> )	(%)	(Wh/m <sup>2</sup> )	(%)			(Wh/m <sup>2</sup> )	(%)	(Wh/m <sup>2</sup> )	(%)	
1	2,2	≈0	85,3	19,1	2,5	1	0,7	≈0	87,6	19,5	2,3
2	14,9	1,1	113,9	25,5	4,7	2	4,8	3,0	117,4	26,2	6,0
3	17,0	3,7	135,8	30,4	7,4	3	16,6	2,5	136,9	30,5	8,0
4	8,3	1,0	149,3	33,5	12,5	4	4,3	1,8	151,5	33,8	11,0
5	30,0	6,5	166,1	37,2	12,4	5	29,0	7,1	164,7	36,7	11,3
6	37,7	8,2	177,6	39,8	12,9	6	36,9	5,6	171,8	38,3	12,9

(e) Rocha						(f) São Martinho da Serra					
$\Delta t$ (hrs)	MBD		RMSD		FS (%)	$\Delta t$ (hrs)	MBD		RMSD		FS (%)
	(Wh/m <sup>2</sup> )	(%)	(Wh/m <sup>2</sup> )	(%)			(Wh/m <sup>2</sup> )	(%)	(Wh/m <sup>2</sup> )	(%)	
1	≈0	≈0	92,7	21,4	2,2	1	8,2	1,8	95,2	21,1	2,4
2	9,6	2,2	125,5	28,9	5,0	2	24,5	5,4	128,7	28,5	2,5
3	8,4	1,9	144,4	33,3	8,3	3	14,5	3,2	147,6	32,7	6,5
4	25,2	5,8	162,1	37,4	7,5	4	17,7	3,9	161,5	35,8	10,0
5	13,3	3,1	165,8	38,2	12,6	5	33,0	7,3	175,3	38,9	11,0
6	24,0	5,5	172,1	39,7	14,3	6	27,2	6,0	181,5	40,2	12,6



Tabla 5: Indicadores de desempeño utilizando el modelo con  $k_c$  y ST ( $\rho$ ) como entradas.

(a) Artigas						(b) Tacuarembó					
$\Delta t$ (hrs)	MBD		RMSD		FS (%)	$\Delta t$ (hrs)	MBD		RMSD		FS (%)
	(Wh/m <sup>2</sup> )	(%)	(Wh/m <sup>2</sup> )	(%)			(Wh/m <sup>2</sup> )	(%)			
1	12,2	2,6	80,7	18,2	8,1	1	2,5	0,6	84,4	18,6	10,0
2	13,9	3,0	112,2	24,6	9,9	2	-7,1	-1,6	115,1	25,4	11,6
3	14,5	3,1	138,5	29,4	11,0	3	-2,4	-0,5	136,7	30,2	13,0
4	29,6	6,4	158,2	32,8	12,4	4	13,1	2,9	154,6	34,1	10,3
5	18,8	4,1	171,4	35,2	13,9	5	19,0	4,2	166,8	36,8	8,4
6	34,9	7,5	185,8	37,6	13,8	6	21,1	4,7	172,6	38,1	8,9

(c) Colonia						(d) Canelones					
$\Delta t$ (hrs)	MBD		RMSD		FS (%)	$\Delta t$ (hrs)	MBD		RMSD		FS (%)
	(Wh/m <sup>2</sup> )	(%)	(Wh/m <sup>2</sup> )	(%)			(Wh/m <sup>2</sup> )	(%)			
1	2,9	0,7	77,9	17,5	11,1	1	0,6	≈0	80,2	17,9	10,6
2	13,0	2,9	107,1	24,0	10,3	2	20,4	4,5	114,6	25,5	8,2
3	0,4	≈0	128,2	28,7	12,6	3	-16,3	-3,6	132,5	29,5	11,0
4	21,2	4,8	147,8	33,1	13,4	4	25,2	5,6	150,5	33,5	11,6
5	25,1	5,6	162,2	36,3	14,4	5	12,5	2,8	161,3	36,0	13,1
6	31,4	7,0	174,9	39,2	14,2	6	33,0	7,4	171,8	38,3	12,9

(e) Rocha						(f) São Martinho da Serra					
$\Delta t$ (hrs)	MBD		RMSD		FS (%)	$\Delta t$ (hrs)	MBD		RMSD		FS (%)
	(Wh/m <sup>2</sup> )	(%)	(Wh/m <sup>2</sup> )	(%)			(Wh/m <sup>2</sup> )	(%)			
1	4,6	1,1	85,0	19,6	10,3	1	12,4	2,7	90,0	20,0	7,5
2	7,5	1,7	120,3	27,7	8,9	2	21,1	4,7	122,6	27,2	7,1
3	11,4	2,6	141,3	32,6	10,2	3	23,9	5,3	142,8	31,6	9,5
4	22,2	5,1	156,4	36,0	10,7	4	34,1	7,6	161,4	35,8	10,0
5	26,3	6,1	165,7	38,2	12,7	5	27,8	6,2	171,7	38,1	12,8
6	33,6	7,8	172,5	39,8	14,1	6	36,5	8,1	183,5	40,7	11,7

## AGRADECIMIENTOS

Los autores agradecen el apoyo financiero del Proyecto ANII-FSE-2016-131799 de la Agencia Nacional de Investigación e Innovación (Uruguay). Se agradece también la disponibilidad de datos medidos de la RMCIS del Laboratorio de Energía Solar de la Udelar (Uruguay) y de la red SONDA del Instituto Nacional de Pesquisas Espaciais (INPE, Brasil). Finalmente, agradecemos la colaboración del Dr. M. Delbracio en las discusiones científicas sobre los modelos RNN-LSTM y su implementación.

## REFERENCIAS

- Abdel-Nasser, M. and Mahmoud, K. (2019). Accurate photovoltaic power forecasting models using deep LSTM-RNN. *Neural Computing and Applications*, 31(7):2727–2740.
- ADME (2018). Informe Anual 2018. Technical report, Administración del Mercado Eléctrico, Uruguay.
- Alonso-Suárez, R. (2017). *Estimación del recurso solar en Uruguay mediante imágenes satelitales*. PhD thesis, Facultad de Ingeniería, Universidad de la República. PhD Thesis.
- Box, G. and Jenkins, G. (1970). *Time Series Analysis Forecasting And Control*, volume 3. Holden-Day.

- Buduma, N. and Locascio, N. (2017). *Fundamentals of deep learning: Designing next-generation machine intelligence algorithms*. .<sup>o</sup>Reilly Media, Inc."
- Coimbra, C. F., Kleissl, J., and Marquez, R. (2013). Chapter 8 - overview of solar-forecasting methods and a metric for accuracy evaluation. In Kleissl, J., editor, *Solar Energy Forecasting and Resource Assessment*, pages 171–194. Academic Press, Boston.
- Cornalino, E. and Chaer, R. (2018). Actualización de estudios de complementariedad eólico solar. Technical report, Facultad de Ingeniería, UdelaR, Montevideo, Uruguay.
- Diagne, M., David, M., Lauret, P., Boland, J., and Schmutz, N. (2013). Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. *Renewable and Sustainable Energy Reviews*, 27:65–76.
- Gensler, A., Henze, J., Sick, B., and Raabe, N. (2016). Deep learning for solar power forecasting—an approach using autoencoder and LSTM neural networks. In *2016 IEEE international conference on systems, man, and cybernetics (SMC)*, pages 002858–002865. IEEE.
- Gueymard, C. A. and Wilcox, S. M. (2011). Assessment of spatial and temporal variability in the US solar resource from radiometric measurements and predictions from models using ground-based or satellite data. *Solar Energy*, 85(5):1068–1084.
- Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.*, 9(8):1735–1780.
- IRENA (2019). Renewable energy statistics 2019. Technical report, The International Renewable Energy Agency, Abu Dhabi.
- Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Lefèvre, M., Oumbe, A., Blanc, P., Espinar, B., Qu, Z., Wald, L., Homscheidt, M. S., and Arola, A. (2013). McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions. *Atmospheric Measurement Techniques, European Geosciences Union*, 6:2403–2418.
- Mellit, A. and Kalogirou, S. A. (2008). Artificial intelligence techniques for photovoltaic applications: A review. *Progress in energy and combustion science*, 34(5):574–632.
- Qing, X. and Niu, Y. (2018). Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy*, 148:461–468.
- Srivastava, S. and Lessmann, S. (2018). A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data. *Solar Energy*, 162:232–247.

## HOURLY SOLAR IRRADIATION FORECAST IN URUGUAY AND SOUTHERN BRAZIL USING RECURRENT NEURAL NETWORKS

**Abstract.** Increasing the photovoltaic solar energy (PV) contribution in electricity grids requires the use of solar irradiation forecast techniques, seeking for an efficient dispatch, cost optimization and valorization of surplus energy for export. The aim of this work is to evaluate the performance of solar irradiation forecast for the region of Uruguay and Rio Grande do Sul (RS) using Long Short-Term Memory (LSTM) Recurrent Neural Networks (RNN). Hourly solar irradiation forecast up to 6 hours ahead is assessed. The performance is evaluated using different inputs for the RNN-LSTM. In the first model, the clear sky index ( $k_c$ ) time-series is used as only input. In the second model, the satellite planetary reflectance (ST) is aggregated. A site-average rRMSD between 20.1 % to 39.2 % is found for the  $k_c$  model and between 18.6 % to 39.0 % is found for the  $k_c$  and ST model, from 1 to 6 hours ahead respectively. It is observed that satellite information only improves the forecast up to 3 hours ahead. This limitation is associated with the cell size that is used for spatial smoothing of the satellite reflectance. The results obtained for both models outperform the best smart persistence, showing that the technique is useful for solar energy forecast in the region.

**Keywords:** solar forecast, RNN-LSTM, satellite images.