

# Metodología para control de calidad de datos de radiación fotosintéticamente activa

Lucas J. Burgos<sup>1,2</sup>, María J. Denegri<sup>1,2</sup>

E mail: lucasjburgos@gmail.com

<sup>1</sup>Departamento de Tecnología, Universidad Nacional de Luján (UNLu)

<sup>2</sup>Grupo de Estudios de la Radiación Solar (GERSolar, INEDES UNLu-CONICET)

**Palabras clave:** medición de PAR en superficie, combinación de filtros analíticos, umbrales

## 1) INTRODUCCIÓN

La radiación fotosintéticamente activa (PAR) es importante para las comunidades de plantas y el modelado del crecimiento de la vegetación debido a su relación con el proceso de fotosíntesis. Es una componente necesaria para estimar y comprender, por ejemplo, la productividad primaria bruta, la eficiencia en el uso de la radiación y el intercambio neto de dióxido de carbono en los ecosistemas (Zhang y otros, 2000; Aguiar y otros, 2012; Nwokolo y Amadi, 2018). Sin embargo, no existe a nivel mundial una red que haya establecido un método rutinario para la medición de PAR y el control de calidad de los datos generados. Estas medidas solo se toman en algunas redes experimentales cuyo objetivo es medir los intercambios de dióxido de carbono, vapor de agua y energía entre la biosfera y la atmósfera.

La calidad de una serie de datos es resultado de las condiciones en que se midió, los instrumentos utilizados, su calibración periódica con trazabilidad a un estándar y, sobre todo, el mantenimiento periódico que recibieron. Por lo tanto, el control de calidad busca asegurar la consistencia de la medición, tanto interna entre sus componentes, como con otras series de medida cercanas y coetáneas. Este proceso es un paso indispensable antes de confiar en los registros para extraer conclusiones significativas. El objetivo de este trabajo es desarrollar una metodología de control de calidad de datos de PAR que permita detectar posibles anomalías producto de las diversas fuentes de error existentes en la cadena de obtención de los datos.

## 2) METODOLOGÍA Y RESULTADOS

La metodología de control de calidad de datos de PAR elaborada en este trabajo constituye el resultado del mismo, por lo tanto, “metodología” y “resultados” se desarrollarán en conjunto. Considerando la necesidad de controlar datos de PAR medida en tierra y, pertenecientes a una serie continua, se propone caracterizar cada uno de ellos mediante aplicación de un conjunto de algoritmos, a los que llamaremos filtros (F). Estos imponen ciertas condiciones que el dato debe cumplir para que sea etiquetado como cumplidor (1) o no (0) de tales requisitos. Los filtros son independientes entre sí, sin embargo, las posibles combinaciones entre ellos permiten caracterizar cada dato en función de sí mismo y de su entorno en términos temporales. A continuación se desarrollan siete filtros analíticos de control de calidad de datos de PAR medida en tierra en el plano horizontal, que pertenezcan a series no menores al mes ni superiores al año, con frecuencia de medición ( $\Delta t$ ) de entre 1 y 10 minutos. Llamaremos “Q” al dato de PAR en mol/m<sup>2</sup>, y “QFn” a aquel etiquetado por el filtro número n, agregando el sufijo “(1)” o “(0)” según sea etiquetado de forma positiva o negativa respectivamente. Los parámetros de filtrado pueden ser ajustados según las exigencias de control deseadas, la  $\Delta t$  y la localidad.

**Filtro 1 Mínima altura solar:** impone que el dato se corresponda con una altura solar “ $\alpha_s$ ” superior a la mínima preestablecida (recomendable 7°), y tiene por objeto eliminar muestras con error instrumental que pueden producirse durante el amanecer y el anochecer.

**Filtro 2 Límites físicamente posibles por BSRN:** propuestos por Long y Shi (2008), que configurados para PAR mediante eficiencia fotosintética ( $Q/H = 1,97$  mol/MJ; Denegri, 2014), identifican datos cuyos valores ( $I_Q$ ) son físicamente posibles pero altamente improbables. El factor de conversión ( $F_{CEF}$ ) adapta los límites energéticos originales a límites cuánticos para PAR. “ $I_a$ ” es la constante solar ajustada por la distancia Sol-Tierra y  $\theta_z$  es el ángulo cenital.

$$-2 \frac{W}{m^2} \cdot F_{CEF} < I_Q < \left[ 1,2 \cdot I_a \cdot (\cos \theta_z)^{1,2} + 50 \frac{W}{m^2} \right] \cdot F_{CEF} \quad (1) \quad F_{CEF} = 0,0036 \frac{MJ}{m^2} \cdot \frac{Q}{H} \quad (2)$$

**Filtro 3 Límites ESRA CSM:** el modelo de cielo claro de ESRA (Rigollier y otros, 2000), configurado para PAR mediante  $F_{CEF}$ , altitud del sitio y un índice de turbidez de Linke ( $T_L$ ) de 1,8 para una atmósfera excepcionalmente limpia (Laguarda y otros, 2021), establece el límite superior por debajo del cual debe encontrarse el valor del dato ( $I_Q$ ).

$$0 \frac{mol}{m^2} < I_Q < I_H^{ESRA\ CSM_{T_L=1,8}} \cdot F_{CEF} \quad (3)$$

**Filtro 4 Máximo índice de claridad:** impone un valor máximo al índice de claridad modificado por Pérez y otros ( $K_{tp}$ ; 1990), que corrige la limitante asociada a la dependencia entre el índice de claridad usual ( $K_t$ ) y la altura solar. Para pequeñas  $\Delta t$ , por exploración visual, se propone un máximo  $K_{tpPAR}$  de 0,89, quedando definido el filtro como  $K_{tpPAR} < 0,89$ , donde:

$$K_{tpPAR} = \frac{K_{tPAR}}{1,031 \cdot \exp[-1,4 / (0,9 + 9,4/m)] + 0,1} \quad (4) \quad K_{tPAR} = \frac{I_Q}{I_a \cdot F_{CEF} \cdot \cos\theta_Z} \quad (5)$$

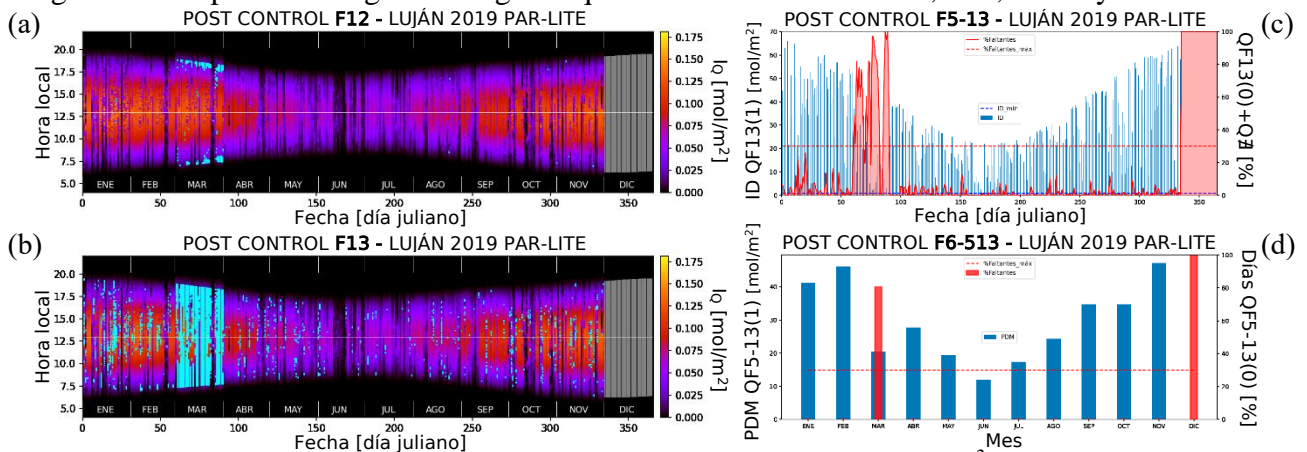
Los siguientes filtros etiquetan el total de datos pertenecientes a un día, una década (diez días) o un mes, según se cumplan o no las condiciones impuestas a los respectivos conjuntos de datos. Deben ser ejecutados en combinación con el Filtro 1 (F1) y opcionalmente con los tres restantes, de forma individual (F1n, siendo n 2, 3 o 4) o en conjunto (F1n, siendo n 24 o 34).

**Filtro 5-1n Mínima integridad y valor de la integral diaria:** calculada la integral diaria (IDF1n) como la sumatoria del valor de los datos  $QF1n(1)$  contenidos en cada día (cuya duración horaria fue acotada por el F1) debe ser superior a 1 mol/m<sup>2</sup>d, mínimo valor de integral diaria altamente improbable. A su vez, la proporción de datos  $QF1n(0)$  e inexistentes ( $Q\bar{A}$ ), no debe superar la tolerancia máxima admitida (recomendable  $\leq 30\%$ ) en tal período de tiempo.

**Filtro 6-51n Mínima integridad del promedio diario mensual:** el promedio diario mensual (PDMF5-1n) es calculado como la media de las integrales diarias de los días del mes cuyos datos son  $QF5-1n(1)$ . Por lo tanto, la proporción de días cuyos datos son  $QF5-1n(0)$ , no debe superar la tolerancia máxima admitida (recomendable  $\leq 30\%$ ) en tal período de tiempo.

**Filtro 7-51n Mínima integridad del promedio diario decádico:** el promedio diario decádico (PDDF5-1n) es calculado como la media de las integrales diarias de los días de la década cuyos datos son  $QF5-1n(1)$ . Por lo tanto, la proporción de días cuyos datos son  $QF5-1n(0)$ , no debe superar la tolerancia máxima admitida (recomendable  $\leq 30\%$ ) en tal período de tiempo.

**Resultado de la aplicación de esta metodología:** desarrollado el software de control de calidad en lenguaje Python, que integra los filtros descritos, se somete a control una serie anual de datos de PAR medida por el GERSolar durante el año 2019 en Luján, Bs.As., Argentina, con el sensor PAR-Lite de Kipp & Zonen, cuya  $\Delta t$  es de 1 minuto. Del reporte generado se presentan algunas imágenes que hacen referencia al F12, F13, F5-13 y F6-513.



**Figura 1:** (a y b) Diagrama de hora local vs fecha coloreado por  $I_Q$  en mol/m<sup>2</sup>. Puntos grises indican  $Q\bar{A}$ , mientras que puntos cian indican  $QF12(0)$  y  $QF13(0)$ . (c) En azul, integrales diarias en mol/m<sup>2</sup> a partir de  $QF13(1)$  y, en rojo, proporción de  $QF13(0)+Q\bar{A}$  vs fecha. Líneas punteadas azul y roja representan el mínimo valor de integral diaria admitido y la tolerancia porcentual máxima de  $QF13(0)+Q\bar{A}$  respectivamente. (d) En azul, promedios diarios mensuales en mol/m<sup>2</sup> a partir de

*integrales diarias de días con QF5-13(1) y, en rojo, proporción de días con QF5-13(0) vs mes. Línea punteada roja representa la tolerancia porcentual máxima de días con QF5-13(0).*

Suponiendo la necesidad de extraer datos para un fin en particular, como se observa en la figura 1 a y b, el F12 tiende a detectar pocos datos y suficientemente anormales, mientras que F13, F14 y F134 detectan muchos datos e insuficientemente anormales, muchas veces asociados a sobreirradiancias normales producidas por cúmulos (puntos cian aislados). Dicho esto, en caso de evaluarse una serie defectuosa, los datos QF12(1) sobreestimarían la PAR total incidente, mientras que los QF13(1), QF14(1) y QF134(1) la subestimarían. Resultando insuficiente la implementación aislada de estos filtros. Los F5-1n, F6-1n y F7-1n (siendo n 3, 4 o 34), permiten identificar datos anormales en relación a días, décadas y meses específicos del año (como el mes de marzo en la figura 1 c y d) que F12 no hubiera detectado. Por lo tanto, a fin de evitar la sobre y subestimación de PAR mencionada anteriormente, se puede combinar F12 con F5-1n, F6-1n y/o F7-1n. En el ejemplo de la figura 1, esa posibilidad de extracción de datos para un fin en particular es, elegir los QF12(1) evitando aquellos QF5-13(0) y QF6-513(0), siendo la serie resultante conformada por el conjunto:  $QF12(1) \cap QF5-13(1) \cap QF6-513(1)$ .

A su vez, dado el ajuste cercano del límite superior de F13 al valor de los datos medidos, este resulta útil para detectar corrimientos horarios, es decir, desfases entre la hora de medición asignada al dato y la hora real a la que fue medido. Situación frecuente que puede ser corregida automáticamente mediante convoluciones de décadas o meses buscando, por corrimiento, la menor diferencia entre la cantidad de QF13(0) antes y después del mediodía solar. Los F13, F14 y F134 resultan útiles para detectar patrones anormales en el conjunto de datos de PAR.

### 3) CONCLUSIONES

- Las combinaciones entre los distintos filtros brindan una importante oportunidad para automatizar el control de calidad de grandes volúmenes de datos, generar productos confiables, representativos de la realidad, y elevar los estándares de calidad de las bases de datos de PAR destinadas a aplicaciones de vanguardia como las de Machine Learning e Inteligencia Artificial.
- El próximo paso es pensar, por ejemplo, en un F8-1n que permita identificar datos integrantes de promedios diarios mensuales o decádicos cuyos valores sean satisfactoriamente inferiores a umbrales máximos pero anormalmente bajos para el momento del año en cuestión. Podría ser en función de un umbral mínimo porcentual de datos QF13(0) en cada período de tiempo.
- Se continuará trabajando en un sistema de validación cruzada con datos de radiación solar global en el plano horizontal (H) que permita sumar más información al control de calidad de datos PAR coetáneos de datos de H en un mismo sitio mediante  $Q/H$  y relación  $H_{PAR}/H$ .

### REFERENCIAS

- Aguiar, L.J., Fischer, G.R., Ladle, R.J., Malhado, A., Justino, F.B., Aguiar, R.G. y da Costa, J.M.N., 2012:** Modeling the photosynthetically active radiation in South West Amazonia under all sky conditions. *Theoretical and Applied Climatology*, 108(3), 631-640.
- Denegri, M.J., 2014:** Medición de la Radiación Fotosintéticamente Activa (PAR) en planos con diferentes orientaciones, mediante el desarrollo y ensayo de un dispositivo ad-hoc, y verificación de modelos. Tesis para optar al grado de doctor. Universidad Nacional de Luján.
- Laguarda, A., Iturbide, P., Orsi, X., Denegri, M.J., Luza, S., Burgos, L., Stern, V., Alonso-Suárez, R., 2021:** Validación de modelos satelitales Heliosat-4 y CIM-ESRA para la estimación de irradiancia solar en la Pampa Húmeda. *Energías Renovables y Medio Ambiente (ERMA)*, Vol. 48, pp. 1 – 9, 2021.
- Long, C. y Shi, Y., 2008:** An automated quality assessment and control algorithm for surface radiation measurements. *The Open Atmospheric Science Journal*, 2008, 2, 23–37.
- Nwokolo, S.C. y Amadi, S.O., 2018:** A Global Review of Empirical Models for Estimating Photosynthetically Active Radiation. *Tr. Ren. Energy*, Vol.4, No.2, 236-327.
- Perez, R., Ineichen, P., Seals, R., Zelenka, A., 1990:** Making full use of the clearness index for parameterizing hourly insolation conditions. *Solar Energy* Vol. 45, No. 2 pp. 111-114, 1990.
- Rigollier C., Bauer O., Wald L., 2000:** On the clear sky model of the 4th European Solar Radiation Atlas with respect to the Heliosat method. *Solar Energy*, 68(1), 33-48.
- Zhang, X., Zhang Y. y Zhou Y., 2000:** Measuring and modelling photosynthetically active radiation in Tibet Plateau during April–October. *Agr. For. Met.* 102, 207–212.