Revista Internacional de Ciencia y Tecnología de la Información Geográfica
International Review of Geographical Information Science and Technology

Roca-Fernández, C., Ninyerola, M., Pons, X. (2024). Diagnosis for integrating the main Iberian networks of solar radiation meteorological stations. *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica* (Articles), 33, 43-75. http://dx.doi.org/10.21138/GF.850

DIAGNOSIS FOR INTEGRATING THE MAIN IBERIAN NETWORKS OF SOLAR RADIATION METEOROLOGICAL STATIONS

[1a]Catalina Roca-Fernández, [2]Miquel Ninyerola, [1b]Xavier Pons
[1,2]Grumets Research Group
[1]Departament de Geografia. Edifici B. Universitat Autònoma de Barcelona. 08193, Bellaterra, Barcelona, Catalonia, Spain
[2]Departament de Biologia Animal, Biologia Vegetal i Ecologia. Edifici C. Universitat Autònoma de Barcelona. 08193, Bellaterra, Barcelona, Catalonia, Spain
[1a]catalina.roca@uab.cat, [2]miquel.ninyerola@uab.cat, [1b]xavier.pons@uab.cat

ABSTRACT

This work carries out a diagnosis of the data from the two main networks of solar radiation stations in the Iberian Peninsula: the *Agencia Estatal de Meteorología* (AEMET, Spain, with daily data 1980–2020), and the *Sistema Nacional de Informação de Recursos Hídricos* (SNIRH, Portugal, with hourly data 2001–2020). One of the goals is to advance towards the synergistic use of the two networks and, hence, to improve the integrated knowledge of solar radiation in this interesting territory. The diagnosis was carried out considering measured magnitudes and instruments, error filtering and number of available stations, data completeness results for the networks and for each meteorological station, and territorial representativeness. An important finding is that it is necessary to guarantee greater temporal completeness for the two networks, which is sometimes very low (SNIRH, 35.1 %). Furthermore, it is necessary to increase spatial completeness in the AEMET network to avoid some large territorial gaps (with points located 140 km from the nearest station). Quality control (QC) should be reinforced in the SNIRH network and diffuse radiation data provided. Solar radiation is a key factor in atmospheric dynamics and climatic phenomena (including drought), and with this study we aim to contribute to the knowledge on solar radiation and provide elements to stimulate the consistency of data in a territory as sensitive and complex as the Iberian Peninsula.

Keywords: global solar radiation; direct solar radiation; diffuse solar radiation; data completeness; quality control; meteorological observations.

DIAGNOSIS PARA LA INTEGRACIÓN DE LAS PRINCIPALES REDES IBÉRICAS DE ESTACIONES METEOROLÓGICAS DE RADIACIÓN SOLAR

RESUMEN

Este trabajo efectúa una diagnosis de los datos de las dos redes principales de estaciones de radiación solar en la Península Ibérica: la Agencia Estatal de Meteorología (AEMET, España, con datos

Received: 12/03/2024 Accepted: 26/04/2024
Academic Editor: Dr. Joan Cristian Padró www.geofocus.org

diarios 1980–2020), y el *Sistema Nacional de Informação de Recursos Hídricos* (SNIRH, Portugal, con datos horarios 2001–2020). Uno de los objetivos es avanzar hacia el uso sinérgico de las dos redes y, así, mejorar el conocimiento integrado de la radiación solar en este interesante territorio. La diagnosis se lleva a cabo teniendo en cuenta, principalmente: magnitudes e instrumentos de medida, filtrado de errores y número de estaciones disponibles, completitud de los datos en las redes y para cada estación meteorológica, y representatividad territorial. Un hallazgo importante es que es necesario garantizar una mayor completitud temporal de las dos redes, que a veces es especialmente baja (SNIRH, 35.1 %). Además, es necesario aumentar la completitud espacial de la red de la AEMET para evitar grandes vacíos territoriales (con puntos situados a 140 km de la estación más cercana). Se debe reforzar el control de calidad (QC) de la red del SNIRH y proporcionar datos sobre radiación difusa. La radiación solar es un factor clave de la dinámica atmosférica y los fenómenos climáticos (incluida la sequía), y con este estudio se quiere contribuir al conocimiento sobre la radiación solar y aportar elementos que estimulen la consistencia de los datos en un territorio tan sensible y complejo como la Península Ibérica.

Palabras clave: radiación solar global; radiación solar directa; radiación solar difusa; completitud de datos; control de calidad; observaciones meteorológicas.

1. Introduction

It is necessary to measure environmental variables to gain the most possible knowledge about our planet. These variables can be used to analyse variability, distribution, impacts, and future trends, which are necessary aspects for a complete understanding of our environment (Bárcena *et al.* 2011, OECC 2022).

Terrestrial Mediterranean ecosystems, due to their topo-climatic heterogeneity, between arid and temperate regions, and their sensitivity to global atmospheric changes, are laboratories of reference for the researching of complex interactions in the environment (Carnicer *et al.* 2019). Indeed, their transitional climate and the high seasonal and interannual variability of their environmental conditions at different study scales on the Iberian Peninsula make them ideal for this purpose, besides to have to consider the Atlantic influence, with a higher cloudiness than in the Mediterranean climate, and therefore differentiated solar radiation patterns between both scenarios (Calbó *et al.* 2008). Strong differences in solar radiation and topographic gradients are paradigmatic examples that contribute to this variability (Aalto *et al.* 2017). At the same time, these ecosystems are very vulnerable, especially due to the current climate change scenario, which involves a trend towards an increased aridity as well as the frequency, intensity, and extent of drought episodes (Doblas-Miranda *et al.* 2017, Vicente-Serrano *et al.* 2019).

In this context, it is extremely important to have as much information as possible about the spatiotemporal variability of solar radiation. Although radiation was traditionally thought to be roughly constant over the years, several studies have shown not only its temporal variation, but also its impact (Sánchez-Lorenzo *et al.* 2013). For example, it has been found to influence temperatures, particularly in industrialized areas (Wild 2012), on changes of several components of the global climate system (Brunet *et al.* 2007, Wild *et al.* 2007), thawing processes (Ohmura *et al.* 2007), the terrestrial carbon cycle and vegetation growth through the regulation of photosynthesis (Gu *et al.* 2002, Wild *et al.* 2012), and agricultural production (Iglesias and Quiroga 2007). Understanding the temporal variation of solar radiation requires considering the dynamics of atmospheric transparency (Stanhill and Cohen 2001, Wild 2009). Among the advances made in this field, it is important to mention the trend towards an increase in diffuse solar radiation as a response to lower atmospheric transparency and an increase in the dispersion and absorption of solar energy (Europe, USA, China, South Africa, India) (Sánchez-Lorenzo *et al.* 2013). Significant impacts on the photosynthetic activity of vegetation have also been demonstrated (Gu *et al.* 2002, Mercado *et al.* 2009).

Some of these studies model solar radiation through astronomical calculations integrated into Geographic Information Systems (GIS), others approximate the data based on remote sensing information, and others combine these two approaches. The first approach, the solar radiation modelled through astronomical calculations integrated into GIS, combines the geometry of solar illumination with

the orography (elevations, slopes, aspects, projected shadows, etc), as for example in Pons (1996), Pons and Ninyerola (2008), Ruiz-Arias *et al.* (2011) and Zhang *et al.* (2015). The second approach, the solar radiation modelled on remote sensing information, uses information from satellites, for example in geostationary orbits, which provide almost continuous data as they can cover different spatiotemporal scales without gaps, such as the SARAH product (Surface Solar Radiation Data Set – Heliosat, Pfeifroth *et al.* 2019); they are used in applications such as calculating the fraction of photosynthetically active radiation absorbed by vegetation (fPAR) in combination with the remote sensing index NDVI (Normalized Difference Vegetation Index), such as in Myneni and Williams (1994), Li *et al.* (2015) and Rahman *et al.* (2015). Examples of the third approach, that combines the solar radiation modelled on astronomical calculations integrated into GIS and remote sensing information, are the Global Solar Atlas product by Solargis (Solargis 2016) and the work of Roca-Fernández *et al.* (2022). The latter proposes two novelties for calculating solar radiation with respect to previous works, based on improving the geometry of solar illumination and integrating satellite remote sensing data (atmospheric optical thickness) combined with the *in-situ* stations. This makes it possible to reflect the spatiotemporal variability of solar radiation more rigorously.

Regardless of the methodology used to calculate solar radiation surfaces at ground level, the data series obtained in meteorological stations are essential for both calibration tasks and for validating the models. These data are commonly obtained with pyranometers and pyrheliometers, which, when well calibrated and maintained, provide reliable long-term solar radiation data at specific locations. They are frequently used for studies of solar radiation and its interaction processes with the atmosphere (Dozier 1989, Trenberth *et al.* 2009). Recently, data from photovoltaic stations have also been used for estimating global solar radiation (Suri *et al.* 2014, Virtuani *et al.* 2017). In all cases, however, the spatial density of the *in-situ* data is very irregular, and its distribution does not, in general, fulfil any scientific criteria that would make it possible to sufficiently describe its spatialization, especially in flat areas (Ninyerola *et al.* 2000, Ruiz-Arias *et al.* 2011).

Although terrestrial measurements tend to be more accurate than modelled data (Salazar *et al.* 2020), the data contain measurement errors or drawbacks due to the calibration and maintenance of the instruments, the capture and storage of the data in the meteorological station data logger, the location and distribution of the stations over the territory, their temporal representativeness, the environmental conditions, etc. This is why it is necessary to perform harmonized quality controls (QC) to obtain the best results and to reduce the uncertainty of the measured dataset (Salazar *et al.* 2020, Forstinger *et al.* 2021). These analyses also make it possible to verify metadata (if available) and determine temporal accuracy, as well as to detect and/or correct problems such as inconsistencies, spatiotemporal gaps, etc (Sánchez-Lorenzo *et al.* 2013, Salazar *et al.* 2020, Forstinger *et al.* 2021, Padial-Iglesias *et al.* 2022).

The objective of this work is to carry out a diagnosis of the data quality of the two main networks of solar radiation stations of the Iberian Peninsula: the *Agencia Estatal de Meteorología* (AEMET, Spain) and the *Sistema Nacional de Informação de Recursos Hídricos* (SNIRH, Portugal) (Figure 1). We also determine the possibilities of using the two networks in combination to be able to better study the solar radiation in this complex territory in a more integrated way. The diagnosis was carried out considering the following aspects: measured magnitudes and instruments, error filtering and number of available stations, data completeness results for the networks and for each meteorological station, and territorial representativeness.

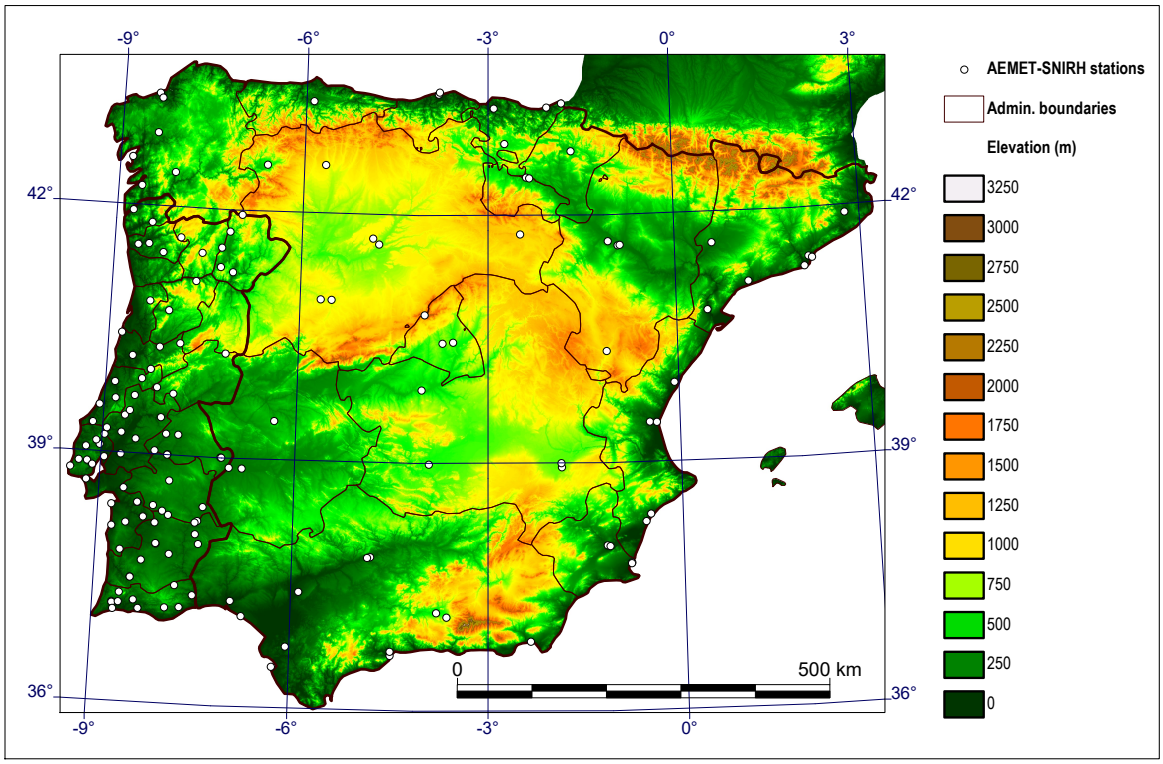

Figure 1. Location of AEMET-SNIRH meteorological stations network in the Iberian Peninsula.
Reference System: UTM-30N-ETRS89. Source: own elaboration based on the AEMET and SNIRH networks.

2. The ground-based solar radiation network AEMET-Spain, its treatment, and data completeness results

2.1. Measured magnitudes and instruments

AEMET provides solar radiation data in 10 kJ·m^{-2}·day^{-1} (read tens of kJ per m^2 and per day), *i.e.*, in units of energy per unit of surface and per unit of time. These units are the same as those used by the World Radiometric Reference (WRR) in accordance with the scale established in 1980 (Pons 1996). Radiation is measured in three different ways depending on the meteorological station: global (GHI), direct (DNI) and diffuse (DIF). In this network, solar radiation has been measured for the daily period 1980–2020. These data can be requested from AEMET through its digital headquarters < https://www.aemet.es/es/sede_electronica >. Data were provided in separate files for each of the major watersheds.

For comparisons with GHI and DIF data, DNI data can be converted to the horizontal plane by multiplying it by the sine of the solar elevation angle at the time of measurement (Equation 1).

$$\mathrm{DHI}=\sin(\theta_s)\cdot\mathrm{DNI} \quad (1)$$

Where:

DHI is the DNI calculated on the horizontal plane.

θ_s is the solar elevation angle.

DNI is direct solar radiation.

Since AEMET does not provide the values of θ_s for each measurement, these have been obtained with the Astres application of MiraMon (Pons 2024), based on the geographic coordinates, the date, and the time in UTC (as provided by AEMET).

GHI and DIF are obtained on the horizontal plane because they are usually obtained with a pyranometer that measures radiation on a flat surface. The GHI includes the radiation received directly

from the Sun, as well as the DIF scattered through the atmosphere. To measure the DIF, the take on DNI into the instrument is avoided through a screen system (AEMET 2024). On the other hand, the DNI is not obtained on the horizontal plane, since this is usually captured with a pyrheliometer that measures the radiation continuously following the Sun and receives the light perpendicular to the plane of the sensor. Pyrheliometers are mounted on a mechanism that allows them to follow the Sun in azimuth and elevation (AEMET 2024). Given that this reading is perfectly oriented towards the Sun, the DNI values are much higher, in most cases even higher than those of the GHI: ~+500 (median of the data of the AEMET 1980–2020 in 10 $kJ \cdot m^{-2} \cdot day^{-1}$). However, if the DNI is calculated on the horizontal plane (DHI) the results are lower than those of the GHI, which is logical because this does not include the DIF. The following conclusions have been drawn by considering the median of the GHI, DNI, DHI and DIF of the period 1980–2020 and of all the AEMET stations that provide these measurements, and comparing the results according to whether the DNI or DHI is used in the calculations:

- Taking DNI into account, the average of the differences between the observed GHI and that calculated from the sum of DNI and DIF is 703.3, the median is 780.0, the standard deviation is 232.4, and the mean absolute deviation around the median is 150.4 (results in 10 $kJ \cdot m^{-2} \cdot day^{-1}$).
- Taking the DHI into account, the average of the differences between the observed GHI and that calculated from the sum of the DHI and the DIF is 257.4, the median is 262.0, the standard deviation is 47.2, and the mean absolute deviation around the median is 36.2 (results in 10 $kJ \cdot m^{-2} \cdot day^{-1}$).

Specifically, Figure 2 shows the daily results of solar radiation obtained at meteorological station 9981A during the month of December 2017. It presents the observed GHI, DNI and DIF by AEMET, an estimated GHI obtained from the sum of the observed DNI and DIF (GHIe=DNI+DIF), the DNI calculated on the horizontal plane using Equation 1 (DHI), an estimated GHI obtained from the sum of the DHI and the observed DIF ($GHIe_h$=DHI+DIF), and an estimated DHI obtained from the difference between the observed GHI and DIF (DHIe=GHI-DIF).

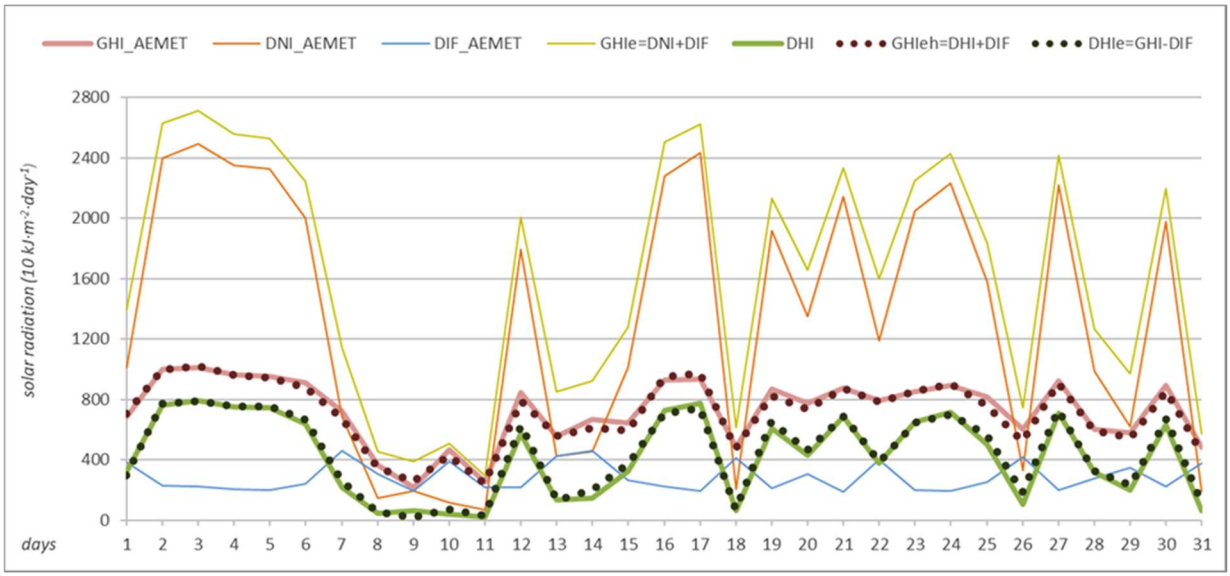

Figure 2. Daily solar radiation at the meteorological station with ID-AEMET network 9981A (December 2017).

Figure 2 compares the different types of solar radiation, showing for station 9981A that:

- On most days in December 2017, the observed direct radiation (DNI) is much higher than the observed global radiation (GHI) for the reason explained above.
- The sum of the observed direct and diffuse radiation (GHIe) is always higher than the observed direct radiation (DNI) and they have a similar temporal behaviour. It should be noted that, as is

logical, this estimated global solar radiation (GHIe) is very different from the observed global radiation (GHI).

- The direct radiation calculated on the horizontal plane (DHI) using Equation 1 is always lower than the observed global radiation (GHI) and has a very similar temporal pattern.
- Despite showing some small differences, the comparison of the observed global radiation (GHI) with the sum of the horizontal direct and observed diffuse radiation ($GHIe_h$) reveals that they have very similar values.
- Similarly, the values obtained from the difference between the observed global and diffuse radiation (DHIe) and the values of the direct radiation calculated on the horizontal plane (DHI) are also very similar.

The AEMET station network adheres to the specifications of the World Meteorological Organization (WMO), forming a National Radiometric Network (NRN) made up of the main stations of the National Radiometric Centres (NRC) and by radiation measuring instruments of other entities. The NRN must meet the following requirements (AEMET 2024):

- The NRCs must have and maintain at least one absolute radiometer for use as a standard reference for checking and calibrating instruments in the rest of the network. In the AEMET network, all the instruments of the stations that measure solar radiation are reviewed and calibrated with those of the NRC every two years.
- The absolute radiometers of the NRC must be calibrated at least every five years with a referenced radiometer from the World Standardization Group (WSG) of the 15 stations that make up the World Radiometric Reference (WRR), which are inter-calibrated at least once every year.
- The NRCs are responsible for maintaining all necessary technical documentation related to the operation and maintenance of the NRN and must ensure accuracy and reliability in the capture and storage of recorded data. They also process and filter the data. In the AEMET network, the NRCs store minute data.

The purchase, installation, calibration, and maintenance of all NRN equipment is carried out by the NRCs. In addition, all data are sent to the World Radiation Data Centre (WRDC) (AEMET 2024).

2.2. Error filtering and number of available stations

To obtain a network of solar radiation data stations for the territory of peninsular Spain, we excluded those stations outside this territory, such as the one located in Melilla and those located in the Balearic Islands and the Canary Islands. AEMET provides a data quality field (ID_FLAG), marking data with the values 0 or 1 as correct (manual or automatic validation, respectively), and marking data as unverified or questionable with the values 10 (unverified), 20 and 21 (questionable data by manual or automatic validation, respectively). Consequently, AEMET solar radiation data with any of these last three values (10, 20, 21) were removed. Therefore, the QC of the AEMET is integrated into this study. After applying this filtering, the GHI AEMET station network was composed of 65 stations (considering the period 1980–2020), of 41 stations (1980–2000 period), and of 54 stations (2001–2020 period, common with the stations of the SNIRH, section 3). The number of stations varies according to the temporal grouping analysed (monthly and/or yearly).

The authors carried out a complementary analysis of the data and did not detect situations that require additional filtering.

2.3. Data completeness results

Tables 1–5 show the temporal completeness results of the AEMET network data. These tables highlight the completeness in percentage based on the number of potential observations considering the number of stations that provide observations of solar radiation for each year and day, as well as considering the maximum number of stations as if for all years of the temporal series all stations had been active. This second calculation makes it possible to determine the degree of completeness

according to the maximum possible number of stations (65). Consequently, it provides information about the degree of density of the network. We studied the complete period 1980–2020 (Tables 1 and A1), and the periods 1980–2000 and 2001–2020 (Tables 2–3 and 4–5, respectively), to show the evolution of the AEMET network. This last period is used in combined studies with the Portugal SNIRH network.

Table 1 is shown in full in Appendix 1 of this work, Table A1, which presents the temporal data completeness for each year of the period 1980–2020. The "by year" version of this table has not been included in this section because we consider it is better to represent these results according to the periods 1980–2000 and 2001–2020 (Tables 2 and 4, respectively). Note that the partial results for each year are the same, regardless of the period considered, while the total results do depend on the period considered.

Studying the complete 1980–2020 series shows that, although at first glance it may seem that the completeness is high, in reality it is not. As shown in Table 1, the temporal completeness of the data considering the number of real annual stations that provide solar radiation observations throughout the period is 88.4 % (GHI), 86.3 % (DNI) and 85.3 % (DIF). However, the percentages of completeness considering the maximum possible number of stations are much lower: 40.4 % (GHI), 32.2 % (DNI) and 37.6 % (DIF). This shows that, considering the period 1980–2020, the number of actual annual stations is quite a bit lower in relation to the maximum possible number of stations (as if all of them had been active every year), which indicates their low spatial completeness.

To complete the diagnosis, we decided to separate the periods 1980–2000 and 2001–2020, firstly because the latter can be used in combined studies with the SNIRH network, and secondly because it begins roughly at the same time as the increase in the number of available stations. Indeed, in the first period the completeness is (Table 3) 33.0 % (GHI), 23.1 % (DNI) and 22.9 % (DIF) considering the maximum possible number of stations. However, in the second period it is (Table 5) 91.0 % (GHI), 87.1 % (DNI) and 88.0 % (DIF) considering the number of actual annual stations, and 73.5 % (GHI), 57.9 % (DNI) and 70.0 % (DIF) considering the number of maximum possible stations.

Comparing Tables 2 and 4 year by year, it can be observed how the number of AEMET stations (GHI, DNI and DIF) is gradually increasing over the years. At the same time a certain agreement is detected between this increase in the number of stations and the increase in the temporal completeness of the data. It is indeed important to emphasize the completeness calculation according to the maximum number of stations because, as can be seen in Table 2, between the years 1981 and 1985 AEMET provided 100 % of the DNI days; however, it only had one active meteorological station that offered this measure. Therefore, in these cases it can be stated that the temporal completeness of the data is total, but with a minimum availability of the stations over the territory. A similar case is that of the GHI of the year 1994 (Table 2), which had 95.5 % temporal completeness of the data based on 13 stations out of the 41 possible (28 stations of difference = spatial completeness of 31.7 %), compared to the GHI of 2016 (Table 4), with 99.3 % completeness based on 49 stations out of 54 possible (5 stations of difference = 90.7 % spatial completeness). Likewise, it is also worth highlighting the number of stations that take GHI, DNI and DIF measurements, with a distribution of the network representatively wider than those that provide the GHI (65[1], 41[2] and 54[3] stations: density of $1.32 \cdot 10^{-4}$, $0.83 \cdot 10^{-4}$ and $1.09 \cdot 10^{-4}$ stations/km^2, respectively), followed by those that measure the DIF (38[1], 20[2] and 35[3] stations: density of $0.77 \cdot 10^{-4}$, $0.41 \cdot 10^{-4}$ and $0.71 \cdot 10^{-4}$ stations/km^2, respectively), and lastly those that measure the DNI (27[1], 9[2] and 27[3] stations: density of $0.55 \cdot 10^{-4}$, $0.18 \cdot 10^{-4}$ and $0.55 \cdot 10^{-4}$ stations/km^2, respectively). The number of stations increased between the two different periods: 1980–2000 with 41 (GHI), 20 (DIF) and 9 (DNI) stations, and 2001–2020 with 54 (GHI), 35 (DIF) and 27 (DNI) stations.

Considering the minimum and maximum values of the temporal completeness of the data, the GHI remains between 68.1 % and 95.5 %, with an availability of 18 and 13 stations out of the 41 possible (spatial completeness of 43.9 % and 31.7 %, respectively; 1980–2000, Table 2), and between 72.0 % and 99.3 %, with an availability of 46 and 49 stations out of the 54 possible (spatial completeness of 85.2 % and 90.7 %, respectively; 2001–2020, Table 4). The temporal completeness of the DNI data remains between 53.3 % and 100.0 %, with an availability of 2 and 1 stations out of the 9 possible (spatial completeness of 22.2 % and 11.1 %, respectively; 1980–2000, Table 2), and between 58.4 %

[1]Period 1980–2020. [2] Period 1980–2000. [3] Period 2001–2020.

and 98.1 %, with an availability of 18 and 21 stations out of the 27 possible (spatial completeness of 66.7 % and 77.8 %, respectively; 2001–2020, Table 4). Finally, the temporal completeness of the DIF data remains between 61.0 % and 100.0 %, with an availability of 5 and 1 stations out of a possible number of 20 stations (spatial completeness of 25.0 % and 5.0 %, respectively; 1980–2000, Table 2), and between 57.8 % and 98.1 %, with an availability of 25 and 31 stations out of the 35 possible (spatial completeness of 71.4 % and 88.6 %, respectively; 2001–2020, Table 4).

Table 1. GHI, DNI, DIF temporal data completeness of AEMET network (1980–2020).

GHI.DC AEMET 1980–2020			
n.days		n.days (%)	How % is computed
393 701 [1]	445 268 [2]	85.8	AVG
		86.9	MED
	445 268 [2]	88.4	1980–2020
	973 440 [3]	40.4	

DNI.DC AEMET 1980–2020			
n.days		n.days (%)	How % is computed
130 169 [1]	150 855 [2]	87.0	AVG
		93.9	MED
	150 855 [2]	86.3	1980–2020
	404 352 [3]	32.2	

DIF.DC AEMET 1980–2020			
n.days		n.days (%)	How % is computed
214 112 [1]	250 941 [2]	81.2	AVG
		80.8	MED
	250 941 [2]	85.3	1980–2020
	569 088 [3]	37.6	

How n.days are computed: [1] number of true days having solar radiation measurements from AEMET stations for the whole period: each station only contributes with its days with data; [2] number of potential days in AEMET stations for the whole period: each station contributes with complete years (~365.25 days) for each year having some data; [3] maximum number of potential days in AEMET stations for the whole period: each station contributes with complete years (~365.25 days) regardless it has data or not. How the percentages of n.days are computed: average of the different years (AVG); median of the different years (MED); whole period (1980–2020).

Table 2. Annual GHI, DNI, DIF temporal data completeness of AEMET network (1980–2000).

GHI.DC AEMET 1980–2000				
Year	n.stations	n.days [4]	n.days [5]	n.days (%)
1980	8	2 559	2 928	87.4
1981	8	2 589	2 920	88.7
1982	10	2 771	3 650	75.9
1983	16	4 598	5 840	78.7
1984	20	5 732	7 320	78.3
1985	18	4 475	6 570	68.1
1986	17	5 150	6 205	83.0
1987	16	4 930	5 840	84.4
1988	14	3 743	5 124	73.0
1989	13	3 871	4 745	81.6
1990	14	3 881	5 110	75.9
1991	15	4 177	5 475	76.3
1992	15	4 242	5 490	77.3
1993	14	4 269	5 110	83.5
1994	13	4 531	4 745	95.5
1995	13	3 880	4 745	81.8
1996	22	6 098	8 052	75.7
1997	23	7 493	8 395	89.3
1998	24	7 773	8 760	88.7
1999	25	8 336	9 125	91.4
2000	28	8 554	10 248	83.5

DNI.DC AEMET 1980–2000				
Year	n.stations	n.days [4]	n.days [5]	n.days (%)
1980	1	356	366	97.3
1981	1	365	365	100.0
1982	1	365	365	100.0
1983	1	365	365	100.0
1984	1	366	366	100.0
1985	1	365	365	100.0
1986	1	362	365	99.2
1987	1	363	365	99.5
1988	2	390	732	53.3
1989	2	497	730	68.1
1990	2	603	730	82.6
1991	2	688	730	94.2
1992	2	720	732	98.4
1993	2	727	730	99.6
1994	2	682	730	93.4
1995	2	726	730	99.5
1996	4	1 073	1 464	73.3
1997	4	1 293	1 460	88.6
1998	4	1 231	1 460	84.3
1999	9	1 864	3 285	56.7
2000	9	2 540	3 294	77.1

DIF.DC AEMET 1980–2000				
Year	n.stations	n.days [4]	n.days [5]	n.days (%)
1980	1	362	366	98.9
1981	1	365	365	100.0
1982	2	458	730	62.7
1983	4	1 268	1 460	86.8
1984	7	1 620	2 562	63.2
1985	7	1 698	2 555	66.5
1986	6	1 599	2 190	73.0
1987	5	1 426	1 825	78.1
1988	5	1 116	1 830	61.0
1989	4	1 094	1 460	74.9
1990	4	1 239	1 460	84.9
1991	6	1 508	2 190	68.9
1992	6	1 561	2 196	71.1
1993	6	1 556	2 190	71.1
1994	6	1 819	2 190	83.1
1995	6	1 533	2 190	70.0
1996	7	1 775	2 562	69.3
1997	7	2 042	2 555	79.9
1998	8	2 430	2 920	83.2
1999	15	3 825	5 475	69.9
2000	17	4 775	6 222	76.7

How n.days are computed: [4] number of true days having solar radiation measurements from AEMET stations for each year: each station only contributes with its days with data; [5] number of potential days in AEMET stations for each year: each station contributes with complete years (~365.25 days) for each year having some data. The minimum value of the GHI, DNI and DIF is represented in pink colour, and the maximum value in green; the homologous cells regarding n.stations are also represented in the same colours for the sake of comparison.

Table 3. GHI, DNI, DIF temporal data completeness of AEMET network (1980–2000).

GHI.DC AEMET 1980–2000			
n.days		n.days (%)	How % is computed
103 652 [1]	126 397 [2]	81.8	AVG
		81.8	MED
	126 397 [2]	82.0	1980–2000
	314 511 [3]	33.0	

DNI.DC AEMET 1980–2000			
n.days		n.days (%)	How % is computed
15 941 [1]	19 729 [2]	88.8	AVG
		97.3	MED
	19 729 [2]	80.8	1980–2000
	69 039 [3]	23.1	

DIF.DC AEMET 1980–2000			
n.days		n.days (%)	How % is computed
35 069 [1]	47 493 [2]	75.9	AVG
		73.0	MED
	47 493 [2]	73.8	1980–2000
	153 420 [3]	22.9	

Legend explained in Table 1.

Table 4. Annual GHI, DNI, DIF temporal data completeness of AEMET network (2001–2020).

GHI.DC AEMET 2001–2020				
Year	n.stations	n.days [4]	n.days [5]	n.days (%)
2001	28	8 058	10 220	78.8
2002	27	7 432	9 855	75.4
2003	33	9 381	12 045	77.9
2004	38	12 470	13 908	89.7
2005	41	11 732	14 965	78.4
2006	43	13 752	15 695	87.6
2007	45	15 322	16 425	93.3
2008	45	15 811	16 470	96.0
2009	46	12 088	16 790	72.0
2010	46	15 070	16 790	89.8
2011	46	16 225	16 790	96.6
2012	48	16 448	17 568	93.6
2013	48	17 349	17 520	99.0
2014	48	17 248	17 520	98.4
2015	49	17 477	17 885	97.7
2016	49	17 800	17 934	99.3
2017	50	17 552	18 250	96.2
2018	49	17 280	17 885	96.6
2019	48	16 927	17 520	96.6
2020	46	14 627	16 836	86.9

DNI.DC AEMET 2001–2020				
Year	n.stations	n.days [4]	n.days [5]	n.days (%)
2001	10	2 488	3 650	68.2
2002	10	2 807	3 650	76.9
2003	10	2 694	3 650	73.8
2004	10	2 827	3 660	77.2
2005	16	3 530	5 840	60.4
2006	18	3 839	6 570	58.4
2007	19	5 713	6 935	82.4
2008	19	6 296	6 954	90.5
2009	19	4 851	6 935	69.9
2010	20	6 190	7 300	84.8
2011	19	6 692	6 935	96.5
2012	20	6 873	7 320	93.9
2013	21	7 293	7 665	95.1
2014	21	7 450	7 665	97.2
2015	21	7 502	7 665	97.9
2016	21	7 542	7 686	98.1
2017	22	7 513	8 030	93.6
2018	21	7 269	7 665	94.8
2019	21	7 443	7 665	97.1
2020	21	7 416	7 686	96.5

DIF.DC AEMET 2001–2020				
Year	n.stations	n.days [4]	n.days [5]	n.days (%)
2001	17	4 938	6 205	79.6
2002	20	5 241	7 300	71.8
2003	21	6 272	7 665	81.8
2004	21	6 209	7 686	80.8
2005	25	5 278	9 125	57.8
2006	27	7 204	9 855	73.1
2007	28	8 916	10 220	87.2
2008	29	9 703	10 614	91.4
2009	29	7 453	10 585	70.4
2010	29	9 143	10 585	86.4
2011	29	10 195	10 585	96.3
2012	31	10 531	11 346	92.8
2013	32	11 260	11 680	96.4
2014	32	11 255	11 680	96.4
2015	31	11 102	11 315	98.1
2016	31	11 134	11 346	98.1
2017	32	11 064	11 680	94.7
2018	31	10 757	11 315	95.1
2019	31	10 782	11 315	95.3
2020	31	10 606	11 346	93.5

Legend explained in Table 2.

Table 5. GHI, DNI, DIF temporal data completeness of AEMET network (2001–2020).

GHI.DC AEMET 2001–2020			
n.days		n.days (%)	How % is computed
290 049 [1]	318 871 [2]	90.0	AVG
		93.5	MED
	318 871 [2]	91.0	2001–2020
	394 470 [3]	73.5	

DNI.DC AEMET 2001–2020			
n.days		n.days (%)	How % is computed
114 228 [1]	131 126 [2]	85.2	AVG
		92.0	MED
	131 126 [2]	87.1	2001–2020
	197 235 [3]	57.9	

DIF.DC AEMET 2001–2020			
n.days		n.days (%)	How % is computed
179 043 [1]	203 448 [2]	86.9	AVG
		92.1	MED
	203 448 [2]	88.0	2001–2020
	255 675 [3]	70.0	

Legend explained in Table 1.

3. The ground-based solar radiation network SNIRH-Portugal, its treatment, and data completeness results

3.1. Measured magnitudes and instruments

SNIRH provides solar radiation data in $W \cdot m^{-2}$, *i.e.*, in units of power per unit of surface. The measured radiation is GHI, for the hourly period 2001–2020. The data were downloaded from < https://snirh.apambiente.pt/ >. The server does not have any stations in the Portuguese island territories and does not provide DNI or DIF data. The data were supplied in different files for each of the measuring stations.

To be able to validate and compare these data with those of AEMET, we transformed $W \cdot m^{-2}$ in one hour to $kWh \cdot m^{-2}$ during the same hour [value $W \cdot m^{-2} \cdot h^{-1} \cdot 0.001$ = value $kWh \cdot m^{-2} \cdot h^{-1}$], and from $kWh \cdot m^{-2} \cdot h^{-1}$ to 10 $kJ \cdot m^{-2} \cdot day^{-1}$, which is the energy measurement unit used by AEMET, and which is more convenient for solar radiation studies in space-time as the amount of energy recorded [value $kWh \cdot m^{-2} \cdot h^{-1} \cdot 3600/10$ = value 10 $kJ \cdot m^{-2} \cdot h^{-1}$]. In summary, to convert the data in $W \cdot m^{-2}$ from SNIRH stations to data in 10 $kJ \cdot m^{-2} \cdot h^{-1}$ it is necessary to multiply by 0.36. Finally, to obtain the data in the units of the AEMET stations (10 $kJ \cdot m^{-2} \cdot day^{-1}$) it is necessary to sum up the 24 values (the accumulated solar radiation of the measurements over the 24 h of the day).

The measuring instrument that SNIRH uses in its station network is a silicon photovoltaic cell (Si-01TCext) (e-mail communication, 12th October 2023), installed at 2 m on any surface (ground or roof). It measures the amount of solar energy in the form of light and heat. It records data between 0 and 2000 $W \cdot m^{-2}$ with an accuracy of ± 5 $W \cdot m^{-2}$. The operation temperature is between -20 °C and +70 °C. The primary capture is made in 15-minute intervals, and the data are sent in real time via GSM (Global System for Mobile Communications). Although solar panels are more efficient in capturing DNI because this is the main energy source for photovoltaic systems, they also capture DIF (Iturbe 2019,

Factorenergia 2023); however, the value of solar radiation recorded through a solar panel is total, and it is currently not possible to measure DNI and DIF separately with these instruments.

3.2. Error filtering and number of available stations

Although the SNIRH provides daily data, this work uses the hourly data because the daily data are provided at 09:00:00 h of the day, understood as the sum of the hours from 10:00:00 h of the previous day until 08:59:59 h of the day for which the data are given (24 values). However, hourly data can be accumulated for a calendar day, from 00:00:00 h to 23:59:59 h (24 values). This means we can create a harmonized daily dataset, to be used individually for analyses of Portugal and also in an integrated way with AEMET for analyses of the Iberian Peninsula as a whole.

To make the cumulative calculation of the 24 h it is necessary to consider the night hours (hours with negative solar elevation angles), for which we would expect a value of 0; however, we sometimes also found assigned a small value of power per unit of surface ($W \cdot m^{-2}$). We do not know the reasons for this, but it is possibly noise from the instrument or pollution from some nearby artificial lighting, etc. Therefore, we assigned a value of 0 to the night hours. An e-mail was sent to the SNIRH asking about the possible reasons why their sensors record solar radiation data during the night, and we received the logical answer: when there is no radiation the sensor should register 0 (e-mail communication, 12th October 2023). Nevertheless, and as said, the data show some readings at night. In relation to this, Salazar *et al.* (2020) assessed the quality of terrestrial data based on comparison with global estimates of solar radiation and preferred not to consider nighttime hours in the data filtering process, as we have done in the present article.

The hours in which the solar elevation angle is negative (night hours) can be eliminated from the temporal series of the data with the Astres application, and radiation can be integrated into the daily calculation of the period 2001–2020, which effectively goes from sunrise to sunset. This procedure was carried out for each meteorological station, hour, day, and month of the central year of the series, considered representative enough for all years to find no differences in the sign of the angle at the same hour, day, and month. After checking solar elevation angles with Astres and cancelling nighttime readings, we assessed the fact that 0 solar irradiance data were sometimes found during daytime hours. In this case, solar radiation was reclassified from 0 to NoData, assuming it to be an error. In relation to this point, the SNIRH reported that it could be that these data with a value of 0 in daytime hours are still in the validation process, while if the sensor fails no value should appear (e-mail communication, 12th October 2023). In any case, these hours are not considered in the calculation of the daily series.

From the same solar elevation angles obtained with Astres, which are necessary to be able to differentiate between daytime and nighttime hours, we calculated the total of actual daylight hours for each meteorological station, day and month of the year. It was determined that there is no day with more than 16 hours of sunlight at any of the Portugal stations. With this value in mind, a screening was carried out for daily data calculated from the daytime hours. We found 80 days with n_values ≥ 17 (hours) distributed between stations 03G_02C, 04N_01C, 22F_03C and 22M_05F, as a result of SNIRH recording the same hour with different random minutes and different measurements of solar radiation, for the years 2004, 2006, 2007, 2014 and 2016, during the months of February, March, June, November and December, and for the days between 1 and 30. In addition, a daily record with several hours of daylight more than what should be (14th December 2006, at station 04N_01C) was removed, using 16 hours in the calculation according to SNIRH; however, as indicated by the different solar positions determined with Astres, on this day and in this location only 9 hours of light are possible.

The daily data that did not reach a minimum of daytime hours in the solar radiation calculation were also removed. In addition, the daily data were discarded if the minimum of n_values was not 80–90 % or higher, according to the number of hours of daylight that a given day must have, also calculated with Astres. Considering each meteorological station for the 2001–2020 series day by day, 90 %, 80 % and 70 % of the total possible hours were calculated. Once the n_values (hours) recorded by the SNIRH had been compared with those acceptable in the calculation of the daily series according to these fractions calculated based on the actual daylight hours, differences of 1–2 hours were found using the 90 %, 2–3 hours using 80 % and 3–4 hours using 70 % data. The 70 % threshold (3–4 hours without solar radiation data in daylight hours) was considered to imply that there was too much missing data,

and that day was considered entirely NoData. We therefore accepted the daily data that reached 90 % of the daytime hours involved in the calculation. For those that only reached 80 %, the day was considered valid only when the count differed by 2 hours or less. In total, only 1.2 % of days with measurements were lost with this procedure.

Once the inconsistent records had been removed and the hours that should be involved in the calculation of the daily solar radiation had been checked, the last filtering of the SNIRH data was carried out according to the amount of energy that reaches the sensor. We applied a threshold of 4000 (10 $kJ \cdot m^{-2} \cdot day^{-1}$), considering that the daily GHI data of the AEMET never exceeds a value of 3700. We removed the 60 daily SNIRH records that exceed this threshold (0.02 % of the total), reaching radiation values up to 6500, and distributed among 9 stations (ID 02E_02GC, 15H_01C, 16C_01C, 18K_01C, 21J_03C, 23E_01C, 25E_03C, 26F_02C, 26M_01C). They were recorded in the years 2002, 2003, 2004, 2008, 2012, 2015 and 2016, during the months of May, June, July, August, September, and October, and for the days between 1 and 31.

As a general example, we detail below some overall results (average of the period 2001–2020) considering the radiation recorded by the SNIRH only at night hours (Table 6), showing the false increase in radiation that would be described if the filtering explained above was not carried out.

Table 6. Average of GHI during nighttime hours. Example in 12 meteorological stations of SNIRH network (2001–2020).

ID-SNIRH	GHI.NIGHT.AVG 2001–2020
21A_01C	277.2
30E_03C	178.1
25E_03C	131.7
16C_01C	128.3
14D_03C	77.2
19O_02C	55.6
17C_07C	45.8
19C_04C	35.7
19J_04C	28.5
06O_06C	28.5
03G_02C	24.0
02P_01C	10.9

Units: 10 $kJ \cdot m^{-2} \cdot day^{-1}$.

As can be seen in Table 6, some of the stations provide a high average GHI in nighttime hours, reaching 277.2 in station 21A_01C for the period 2001–2020. The distribution of these errors is strongly skewed, as shown by the fact that the median at this station is only 2.5, the standard deviation 560.9, the mean absolute deviation around the median 276.0, the 1st quartile 1.1 and the 3rd quartile 7.6 (results in 10 $kJ \cdot m^{-2} \cdot day^{-1}$).

Importantly, although Table 6 shows example data at 12 locations, all SNIRH stations that provide GHI data (88 stations: density of $9.90 \cdot 10^{-4}$ stations/km^2) provide some data during the night for some days, months, and years. The minimum value recorded in all 88 stations is 3.3 (average value at station 04G_06C). Considering the 88 stations, the average for the period 2001–2020 is 28.7, the median 14.4, the standard deviation 39.8, the mean absolute deviation around the median 19.9, the 1st quartile 9.2, and the 3rd quartile 29.3 (results in 10 $kJ \cdot m^{-2} \cdot day^{-1}$).

Going into more detail, we present below the statistics obtained during the years 2003, 2011 and 2019 (Table 7), as well as those obtained at station 21A_01C during the years 2017, 2019 and 2020 (all the years for which this station provides nighttime solar radiation data) (Table 8). No temporal pattern was detected in GHI measurements during night hours.

Table 7. Monthly statistics of GHI during nighttime hours at SNIRH network (examples for 2003, 2011, 2019).

	2003	2011	2019	2003									2011								2019								
Month	n.stations			AVG	MED	STDV	Dmed	MIN	1st.Q	3rd.Q	P.95	MAX	AVG	MED	STDV	Dmed	MIN	1st.Q	3rd.Q	P.95	MAX	AVG	MED	STDV	Dmed	MIN	1st.Q	3rd.Q	P.95	MAX
Jan	60	20	48	17.2	7.5	24.8	12.6	3.1	4.2	13.0	81.1	109.0	10.0	9.7	6.0	5.1	1.1	4.2	13.9	21.3	21.4	37.2	5.3	80.5	34.5	0.8	2.4	35.7	227.8	468.8
Feb	61	21	51	17.9	7.2	25.8	12.8	2.9	5.1	12.2	78.1	111.7	17.0	11.5	19.6	10.3	2.1	5.1	15.8	56.9	87.5	31.5	7.8	62.9	26.8	2.1	4.3	31.6	108.9	397.7
Mar	63	22	53	20.9	7.0	36.4	16.5	2.7	4.0	14.3	84.4	230.0	13.7	10.0	15.9	7.6	1.3	5.8	14.7	53.3	82.7	32.8	8.8	58.6	27.9	0.8	4.6	31.9	146.7	355.9
Apr	65	23	55	16.2	7.8	21.0	10.8	0.7	5.6	13.1	67.1	94.5	15.3	9.7	17.0	8.8	2.6	7.2	14.0	46.0	80.3	27.4	7.4	50.0	23.6	0.7	3.5	32.8	136.3	306.1
May	67	22	58	14.3	6.3	18.8	10.4	2.0	3.6	12.5	64.7	72.8	14.4	10.1	16.0	9.1	0.8	4.3	14.0	58.3	65.3	26.6	7.2	49.2	22.7	1.1	3.7	27.1	181.8	264.8
Jun	69	21	58	18.5	10.5	19.1	10.7	0.4	8.0	16.8	68.3	86.0	13.4	12.0	7.7	5.7	2.6	7.3	16.4	28.7	34.9	27.4	9.6	43.3	21.7	1.8	5.5	34.2	118.9	276.1
Jul	69	20	57	18.0	10.8	17.6	10.1	4.5	8.0	17.0	67.1	73.5	15.7	12.9	12.8	8.2	1.9	7.2	16.9	49.8	56.7	31.5	10.3	53.3	25.5	2.3	5.4	32.7	139.7	274.1
Aug	69	20	58	13.6	5.5	19.5	10.4	0.9	3.0	10.0	69.5	77.1	15.4	7.2	32.3	11.8	1.1	3.7	11.0	89.7	153.5	39.1	5.1	100.8	36.6	0.6	2.1	35.8	291.4	664.4
Sep	69	19	64	17.7	9.0	21.8	11.7	2.8	6.1	15.0	77.6	87.6	25.4	14.4	39.3	17.6	1.7	8.6	23.6	182.1	182.1	75.9	11.7	270.1	69.8	0.4	6.3	40.6	218.6	1842.0
Oct	69	19	65	20.7	8.7	27.2	14.7	3.4	5.7	16.7	92.7	107.7	30.8	14.2	49.3	21.8	3.8	7.7	24.2	194.9	194.9	55.9	7.4	171.1	51.1	0.9	4.6	38.5	174.7	1324.7
Nov	70	18	62	17.3	4.9	28.9	15.4	0.6	1.5	13.2	94.5	118.4	18.7	6.7	43.8	16.2	0.9	2.4	13.2	196.6	196.6	98.1	4.9	381.3	96.6	0.4	1.0	39.9	292.1	2653.8
Dec	69	17	63	16.7	5.5	27.2	15.1	0.4	0.8	12.9	85.3	106.1	7.4	6.4	7.8	5.3	0.9	1.4	10.9	33.7	33.7	100.0	8.6	374.0	97.5	0.4	0.7	59.5	234.1	2744.9
ANNUAL	71	24	67	16.9	7.9	21.8	11.9	0.4	5.0	14.6	76.0	92.7	15.1	10.9	16.7	9.2	1.1	5.6	15.9	57.6	84.8	51.4	10.8	114.1	46.0	1.9	5.4	53.1	180.0	654.2

Units: 10 kJ·m^{-2}·day^{-1}. Statistics computed for each month and year: average (AVG); median (MED); standard deviation (STDV); mean absolute deviation around the median (Dmed); minimum (MIN); first quartile (1st.Q); third quartile (3rd.Q); 95th percentile (P.95); maximum (MAX).

Table 8. Monthly statistics of GHI during nighttime hours at the meteorological station with ID-SNIRH network 21A_01C (2017, 2019, 2020).

	2017	2019	2020	2017									2019								2020								
Month	n.days			AVG	MED	STDV	Dmed	MIN	1st.Q	3rd.Q	P.95	MAX	AVG	MED	STDV	Dmed	MIN	1st.Q	3rd.Q	P.95	MAX	AVG	MED	STDV	Dmed	MIN	1st.Q	3rd.Q	P.95	MAX
Jan	9	13	20	1.3	0.7	1.2	0.9	0.4	0.4	2.2	4.0	4.0	1.9	1.4	1.2	0.9	0.4	0.7	2.2	4.3	4.3	1.2	0.7	1.6	0.7	0.4	0.4	1.1	5.4	7.6
Feb	4	11	11	5.2	4.1	3.1	2.2	2.2	3.1	7.4	10.4	10.4	3.4	1.4	3.8	2.9	0.4	0.4	7.9	11.5	11.5	0.9	0.7	0.5	0.3	0.4	0.4	1.1	2.2	2.2
Mar	18	30	29	1.9	1.1	2.3	1.3	0.4	0.5	2.7	10.1	10.1	2.5	2.0	2.3	1.5	0.4	0.7	3.1	9.2	11.9	2.2	1.8	1.2	1.0	0.4	1.1	2.9	4.0	5.8
Apr	4	7	12	3.7	2.2	3.1	1.9	1.4	1.8	5.6	9.0	9.0	3.5	2.9	2.9	2.6	0.4	0.4	6.1	7.6	7.6	2.2	0.9	2.5	1.6	0.4	0.5	3.1	7.9	7.9
May	18	25	28	2.1	1.4	1.9	1.2	0.4	0.7	3.1	7.9	7.9	2.2	1.4	1.8	1.4	0.4	0.7	4.3	5.0	5.0	2.2	2.0	1.3	1.0	0.4	1.4	3.1	4.5	4.7
Jun	15	16	16	2.9	1.8	2.5	1.8	0.4	1.1	3.6	8.6	8.6	4.0	2.9	3.0	2.0	0.4	2.0	5.0	12.6	12.6	4.5	4.1	1.7	1.5	2.2	3.1	6.1	7.6	7.6
Jul	17	20	23	3.4	2.2	3.1	2.1	0.4	1.1	3.6	11.5	11.5	3.9	3.1	2.6	2.0	0.4	1.8	5.9	9.0	9.7	4.0	3.6	2.7	1.8	1.1	1.8	5.4	7.6	13.3
Aug	ND	30	17	ND	ND	ND	ND	ND	ND	ND	ND	ND	664.4	674.8	596.7	529.3	0.4	2.0	1221.3	1734.8	1748.2	1.5	1.4	0.7	0.6	0.4	0.7	2.2	2.5	2.5
Sep	ND	30	27	ND	ND	ND	ND	ND	ND	ND	ND	ND	1210.8	1196.3	328.9	278.4	578.9	963.5	1478.0	1777.1	1840.0	2.6	2.2	1.7	1.3	0.7	1.1	3.6	5.4	7.6
Oct	ND	31	21	ND	ND	ND	ND	ND	ND	ND	ND	ND	1324.7	1337.4	410.5	347.4	575.6	990.7	1629.4	1906.9	2250.7	2.2	2.2	1.5	1.2	0.4	0.7	3.2	4.7	5.4
Nov	ND	30	20	ND	ND	ND	ND	ND	ND	ND	ND	ND	1503.5	1408.1	372.9	321.7	799.2	1211.2	1880.6	2194.6	2365.6	0.7	0.5	0.6	0.4	0.4	0.4	0.7	2.3	2.9
Dec	ND	11	4	ND	ND	ND	ND	ND	ND	ND	ND	ND	1265.1	1147.7	421.4	342.8	617.8	920.2	1574.6	2154.6	2154.6	0.4	0.4	0.0	0.0	0.4	0.4	0.4	0.4	0.4
ANNUAL	85	254	228	2.6	1.8	2.6	1.7	0.4	0.7	3.2	9.0	11.5	616.9	8.1	703.3	614.4	0.4	1.8	1233.5	1843.7	2365.6	2.2	1.8	1.9	1.4	0.4	0.7	3.2	6.1	13.3

Units: 10 kJ·m^{-2}·day^{-1}. Legend explained in Table 7. ND: NoData.

It is necessary to emphasize that the presence of data during the night is not anecdotic or limited to a few days during each month, as explained in detail in Appendix 2. This highlights the importance of filtering solar radiation measurements during night hours. Otherwise, a large amount of noise would be introduced into the dataset.

Several hypotheses can be formulated regarding why the sensor can capture electromagnetic energy during the night hours:

- *Artificial light arriving to the sensor.* Solar panels are designed to convert any type of light into electricity (EcoInventos 2022, Solarni Paneli 2023); in fact, there are solar panels that can be charged with LED lights, incandescent and fluorescent lights, and with public lighting, if the energy is sufficiently intense.
- *Electromagnetic radiation being captured in the thermal infrared (TIR).* Although work is being carried out on using this radiation in photovoltaic panels (Nielsen *et al.* 2022, I'MNOVATION 2023), it is currently not a technology available in installed solar panels, and therefore we cannot consider this hypothesis as valid.
- *The full Moon radiation arriving to the sensor.* The light of the full Moon is not enough for the sensor to register energy, since this is a million times less bright than the Sun (Carrasco and Carramiñana 1998, MANSUR 2023, Wikipedia 2023, Williams 2024).
- *Batteries being used to store the generated electricity.* The batteries do not affect the data recorded by the solar panel sensor during the night because they simply make use of the surplus energy, previously recorded and stored when the solar panel is not generating electricity, or it is very reduced (Iturbe 2019, SolarPlak 2023).
- *Errors in the capture system.* Since no periodic or systematic problems have been detected, this hypothesis can be discarded.

As a result of the above considerations, the first hypothesis is the most plausible to explain the recording of the GHI data during night hours.

3.3. Data completeness results

Tables 9 and 10 present the temporal completeness results of the SNIRH network data. Completeness is expressed as a percentage based on the number of potential observations through two approximations: with respect to the number of stations providing solar radiation observations for each year and day, and with respect to the maximum number of stations (as if during all the years of the temporal series all the stations had been active). This second calculation approximation makes it possible to determine the degree of completeness according to the maximum possible number of stations (88) to assess the degree of density of the network.

Examining the 2001–2020 series, it can be seen that, although it may appear that completeness is high, it is not. Indeed, and as shown in Table 10, the temporal completeness of the data considering the number of real annual stations that provide observations of solar radiation throughout the period is 78.0 % (GHI). However, the percentage of completeness considering the maximum possible number of stations is lower: 54.3 % (GHI). This highlights that, considering the period 2001–2020, the number of actual annual stations is lower than the maximum possible number of stations (as if all stations were active each year), indicating a moderate spatial completeness.

As can be seen in Table 9, the number of stations increases throughout the annual series of the GHI network of the SNIRH, but in an unstable way both in completeness and density. However, availability waxes and wanes over the period 2001–2020, as does the temporal completeness of the data. Note that a less dense network can give high data completeness (years 2002, 2011, and 2012), while a denser network can give rather low completeness (years 2010, 2014, 2017, and 2018). For example, in the first case, the year 2012 presents 77.6 % temporal completeness of the data based on 21 stations out of the 88 possible (67 stations of difference = spatial completeness of 23.9 %), and in the second case, the year 2010 presents 39.2 % temporal completeness of the data based on 67 stations (21 stations of difference = spatial completeness of 76.1 %). The year 2002 is also noteworthy, as it explains 93.8 % temporal completeness of the data based on 59 stations (29 stations of difference = spatial completeness of 67.0 %), presenting completeness values like that of the years with a denser network (2003–2007).

Table 9. Annual GHI temporal data completeness of SNIRH network (2001–2020).

GHI.DC SNIRH 2001–2020				
Year	n.stations	n.days [4]	n.days [5]	n.days (%)
2001	56	10 815	20 440	52.9
2002	59	20 208	21 535	93.8
2003	71	23 680	25 915	91.4
2004	72	25 320	26 352	96.1
2005	74	25 757	27 010	95.4
2006	75	26 210	27 375	95.7
2007	75	26 518	27 375	96.9
2008	75	23 279	27 450	84.8
2009	70	20 713	25 550	81.1
2010	67	9 593	24 455	39.2
2011	**24**	6 805	8 760	77.7
2012	**21**	5 968	7 686	77.6
2013	**17**	3 898	6 205	62.8
2014	44	5 645	16 060	35.1
2015	77	22 643	28 105	80.6
2016	79	23 954	28 914	82.8
2017	78	17 237	28 470	60.5
2018	58	9 456	21 170	44.7
2019	67	20 061	24 455	82.0
2020	66	21 321	24 156	88.3

Legend explained in Table 2. Bold figures in year rows express the minimum values.

Table 10. GHI temporal data completeness of SNIRH network (2001–2020).

GHI.DC SNIRH 2001–2020			
n.days		n.days (%)	How % is computed
349 081 [1]	447 438 [2]	76.0	AVG
		81.6	MED
	447 438 [2]	78.0	2001–2020
	642 840 [3]	54.3	

Legend explained in Table 1.

Considering the minimum and maximum values of the temporal completeness of the data, the GHI remains between 35.1 % and 96.9 %, with an availability of 44 and 75 stations out of the 88 possible (spatial completeness of 50.0 % and 85.2 %, respectively; Table 9).

4. The ground-based solar radiation network from the Iberian Peninsula (AEMET-SNIRH): spatiotemporal results from the data completeness integrated treatment

Figure 3 shows the distance from any point on the Iberian Peninsula to its nearest meteorological station of the AEMET-SNIRH network. It can be seen that the SNIRH-Portugal provides a very dense network, while the spatial distribution of the AEMET-Spain stations has certain geographical areas in which the distance to the nearest solar radiation station is between 100 and 140 km. These regions are in three main areas: 1) the central and western Pyrenean area; 2) the corridor through the SE area of the Cantabrian Mountains, passing through the Iberian System and part of the eastern strip of the Southern Sub-Plateau, and reaching the Penibaetic System; and 3) the western area of the Central Plateau and the Southern Sub-Plateau, from the Central System to Sierra Morena.

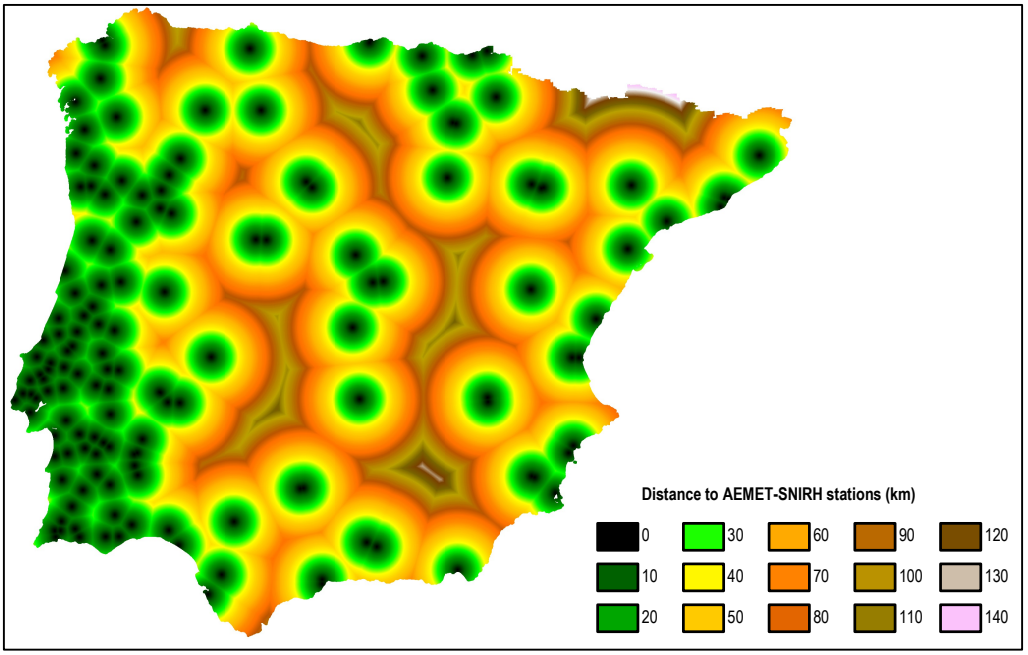

Figure 3. Distances to the nearest peninsular meteorological station of AEMET-SNIRH network if all of them would provide annual data.

Remembering that the comparative analysis can only be made based on GHI data for the common period 2001–2020, Table 11 presents the results of the solar radiation measured at the AEMET and SNIRH stations, as well as their comparison at peninsular level based on the calculation of correlations and the integration of data from the two networks. The annual results were calculated from the medians

of the observed measurements at each station for the three scenarios worked on (Spain, Portugal, and the Iberian Peninsula). Likewise, we provide the average, median, standard deviation, and mean absolute deviation around the median calculated from the annual values, as well as the results considering the entire period. Similarly, and to provide a view of the variation of the stations over time, the figure is shown for each case.

In studying the 2001–2020 series (Table 11), it is evident that the magnitudes of the Pearson correlation coefficient (r) between the solar radiation measurements of the AEMET and the SNIRH are strong (0.98–0.99) and significant (p-value < 0.05). At the same time, although the radiation values are higher considering the AEMET network (due to the fact that it contains more stations in Mediterranean areas with less cloudiness), it can be observed that the dispersion obtained from the annual medians is quite similar between the two networks: considering the standard deviation of 90.3 according to the AEMET stations and of 87.6 according to the SNIRH stations, and considering the mean absolute deviation around the median of 74.6 according to the AEMET stations and of 74.2 according to the SNIRH stations. Observing the same results based on integrating the two networks, the respective dispersion values decrease, obtaining 79.1 for the standard deviation and 64.9 for the mean absolute deviation around the median.

Table 11. Annual comparison of GHI between AEMET and SNIRH networks (2001–2020).

Year	GHI.MED AEMET+SNIRH 2001–2020						
	n.stations			GHI.MED (10 kJ·m^{-2}·dia^{-1})			
	AEMET	SNIRH	AEMET+SNIRH	AEMET	SNIRH	r	AEMET+SNIRH
2001	28	56	84	1589.3	1568.4	0.993	1578.2
2002	27	59	86	1589.3	1428.5	0.985	1468.2
2003	33	71	104	1736.5	1509.0	0.990	1586.3
2004	38	72	110	1622.0	1592.8	0.991	1599.8
2005	41	74	115	1770.0	1565.6	0.995	1618.1
2006	43	75	118	1683.3	1443.2	0.992	1512.8
2007	45	75	120	1719.3	1623.2	0.994	1672.0
2008	45	75	120	1756.3	1522.4	0.994	1596.6
2009	46	70	116	1850.3	1673.6	0.987	1729.0
2010	46	67	113	1674.8	1417.0	0.997	1566.0
2011	46	**24**	70	1723.8	1532.3	0.999	1656.9
2012	48	**21**	69	1803.5	1412.8	0.989	1691.1
2013	48	**17**	65	1628.8	1491.4	0.993	1600.0
2014	48	44	92	1767.0	1437.8	0.994	1678.5
2015	49	77	126	1766.3	1607.3	0.994	1688.6
2016	49	79	128	1779.3	1589.1	0.995	1661.7
2017	50	78	128	1829.5	1626.5	0.992	1718.3
2018	49	58	107	1613.5	1472.2	0.985	1564.9
2019	48	67	115	1890.5	1697.1	0.992	1770.2
2020	46	66	112	1614.0	1457.5	0.996	1514.0
AVG	**43.7**	**61.3**	**104.9**	**1720.3**	**1533.4**	**0.992**	**1623.5**
MED	**46.0**	**68.5**	**112.5**	**1730.1**	**1527.4**	**0.993**	**1609.1**
STDV	**6.9**	**19.5**	**20.1**	**90.3**	**87.6**	**0.004**	**79.1**
Dmed	**4.5**	**13.4**	**15.0**	**74.6**	**74.2**	**0.004**	**64.9**
2001–2020	**54**	**88**	**142**	**1756.0**	**1527.7**	**0.987**	**1678.5**

The median is used to compute the annual GHI from de daily data (GHI.MED), and the Pearson correlation coefficient (r) is computed using AEMET *versus* SNIRH. Statistics computed for each year: average (AVG); median (MED); standard deviation (STDV); mean absolute deviation around the median (Dmed). The minimum value of the GHI is represented in pink colour, and the maximum value in green; the homologous cells regarding n.stations are also represented in the same colours for the sake of comparison. Bold figures in year rows express the minimum values.

The dispersion of the number of stations is higher considering the SNIRH network, which demonstrates the annual irregularity of the density of this network along the temporal series (Table 11): 19.5 standard deviation in the SNIRH network and 6.9 in the AEMET network, and 13.4 mean absolute deviation around the median in the SNIRH network and 4.5 in the AEMET network. Integrating the two networks, the dispersion values of the number of stations increase (20.1 standard deviation and 15.0 mean absolute deviation around the median). Although for most years they exceed 100 locations, for 2 years, less than 90 stations were active (2001 and 2002), and for 3 years, 70 stations or less were active (2011, 2012 and 2013). This is also demonstrated by observing the variability between the total number of active stations for the period 2001–2020 and the results obtained on average or median calculated from the annual values (Table 11), in which the differences in the number of stations considering the integrated network are: 142 active stations throughout the period (density of $2.44 \cdot 10^{-4}$ stations/km^2), and 113 stations (median) providing data over the years (density of $1.94 \cdot 10^{-4}$ stations/km^2). In this context, it is important to emphasize that none of the three scenarios reaches the totality of active stations possible in any year. However, the increased availability of ground stations from integrating the two networks provides more spatiotemporal detail of solar radiation data.

Tables 12 and 13 present the results of the temporal completeness of the GHI data of the integrated AEMET-SNIRH network. The results are also shown separately for each of the networks, already discussed previously in Tables 4, 5, 9 and 10. The number of stations increases throughout the annual series of the integrated AEMET-SNIRH network of GHI, following approximately the same temporal pattern as that presented by the SNIRH network (Table 12).

Examining the 2001–2020 series shows that, although it may appear that completeness is high, it is not. Indeed, and as shown in Table 13, the temporal completeness of the data considering the number of real annual stations that provide observations of solar radiation throughout the period is 83.4 % (GHI). However, the percentage of completeness considering the maximum possible number of stations is lower: 61.6 % (GHI). This highlights that, considering the period 2001–2020, the number of actual annual stations is lower than the maximum possible number of stations (as if all stations were active each year), indicating a moderate spatial completeness. As expected, given that the temporal completeness of the AEMET data is higher than the one of the SNIRH data, the completeness of the integrated network is higher than the one of SNIRH, and this pattern is repeated yearly (Table 13).

Indeed, the annual completeness of the data of the integrated network is lower according to the individual network that is the lower of the two. However, it should be considered that, in any case, the availability of stations increases using the integrated AEMET-SNIRH network, although its spatial completeness may become comparatively lower, as for example in the year 2012 (Table 12). In this case, the spatial completeness of AEMET is 88.9 % (48 stations in operation out of 54 possible), that of SNIRH is 23.9 % (21 stations in operation out of 88 possible), and spatial completeness of the integrated network is 48.6 % (69 stations in operation out of 142 possible). A contrary case is shown in 2005, with a spatial completeness of 75.9 % for AEMET (41 stations in operation out of the 54 possible), 84.1 % for SNIRH (74 stations in operation out of the 88 possible) and 81.0 % for the integrated network (115 stations in operation out of the possible 142). This shows that the increase in the number of stations and the temporal completeness of the data due to integrating the two networks always result in lower values than those of the network with better representativeness.

Considering the minimum and maximum values of the temporal completeness of the data, the GHI remains between 59.8 % and 95.5 %, with an availability of 113 and 120 stations out of the 142 possible (spatial completeness of 79.6 % and 84.5 %, respectively; Table 12). The temporal completeness of the AEMET data is maintained between 72.0 % and 99.3 %, with an availability of 46 and 49 stations out of the 54 possible (spatial completeness of 85.2 % and 90.7 %, respectively; Table 12), and the temporal completeness of the SNIRH data between 35.1 % and 96.9 %, with an availability of 44 and 75 stations out of the 88 possible (spatial completeness of 50.0 % and 85.2 %, respectively; Table 12).

Table 12. Annual comparison of GHI temporal data completeness between AEMET and SNIRH networks (2001–2020).

Year	GHI.DC AEMET+SNIRH 2001–2020									
	n.stations			AEMET			SNIRH			AEMET+SNIRH
	AEMET	SNIRH	AEMET+SNIRH	n.days [4]	n.days [5]	n.days (%)	n.days [4]	n.days [5]	n.days (%)	n.days (%)
2001	28	56	84	8 058	10 220	78.8	10 815	20 440	52.9	61.6
2002	27	59	86	7 432	9 855	75.4	20 208	21 535	93.8	88.1
2003	33	71	104	9 381	12 045	77.9	23 680	25 915	91.4	87.1
2004	38	72	110	12 470	13 908	89.7	25 320	26 352	96.1	93.9
2005	41	74	115	11 732	14 965	78.4	25 757	27 010	95.4	89.3
2006	43	75	118	13 752	15 695	87.6	26 210	27 375	95.7	92.8
2007	45	75	120	15 322	16 425	93.3	26 518	27 375	96.9	95.5
2008	45	75	120	15 811	16 470	96.0	23 279	27 450	84.8	89.0
2009	46	70	116	12 088	16 790	72.0	20 713	25 550	81.1	77.5
2010	46	67	113	15 070	16 790	89.8	9 593	24 455	39.2	59.8
2011	46	**24**	70	16 225	16 790	96.6	6 805	8 760	77.7	90.1
2012	48	**21**	69	16 448	17 568	93.6	5 968	7 686	77.6	88.8
2013	48	**17**	65	17 349	17 520	99.0	3 898	6 205	62.8	89.6
2014	48	44	92	17 248	17 520	98.4	5 645	16 060	35.1	68.2
2015	49	77	126	17 477	17 885	97.7	22 643	28 105	80.6	87.2
2016	49	79	128	17 800	17 934	99.3	23 954	28 914	82.8	89.1
2017	50	78	128	17 552	18 250	96.2	17 237	28 470	60.5	74.5
2018	49	58	107	17 280	17 885	96.6	9 456	21 170	44.7	68.5
2019	48	67	115	16 927	17 520	96.6	20 061	24 455	82.0	88.1
2020	46	66	112	14 627	16 836	86.9	21 321	24 156	88.3	87.7

Legend explained in Table 2. Bold figures in year rows express the minimum values.

Table 13. Comparison of GHI temporal data completeness between AEMET and SNIRH networks (2001–2020).

GHI.DC AEMET+SNIRH 2001–2020							
AEMET			SNIRH			AEMET+SNIRH	How % is
n.days		n.days (%)	n.days		n.days (%)	n.days (%)	computed
290 049 [1]	318 871 [2]	90.0	349 081 [1]	447 438 [2]	76.0	83.3	AVG
		93.5			81.6	88.1	MED
	318 871 [2]	91.0		447 438 [2]	78.0	83.4	2001–2020
	394 470 [3]	73.5		642 840 [3]	54.3	61.6	

Legend explained in Table 1.

By studying the spatial distribution of the temporal completeness of the data of the integrated network for the period 2001–2020, it is evident that the area comprised by a small part of the western area of peninsular Spain and a large part of Portugal is the one that has the lowest completeness percentages (Figures 4 and 5). It is observed that most regions provide a completeness between 60 % and 80 %, two between 50 % and 60 %, two more between 40 % and 50 %, and one between 20 % and 30 %. The regions providing a completeness greater than or equal to 80 % from the area of influence that covers the station are found mostly in Spanish territory (Figure 5). The density of the stations is clearly higher in Portugal, although some less dense areas can be distinguished, such as the easternmost border with Spain. The density of the Spanish stations has a much lower coverage. Therefore, large areas of territory do not have solar radiation data provided for them, as was also observed in Figure 3.

To complete the diagnosis, we decided to represent the observed GHI measurements in each location based on calculating the median of the complete period 2001–2020. This makes it possible to observe that most of the values provided by each station are quite similar (Figure 4), which reinforces what has been commented on for Table 11.

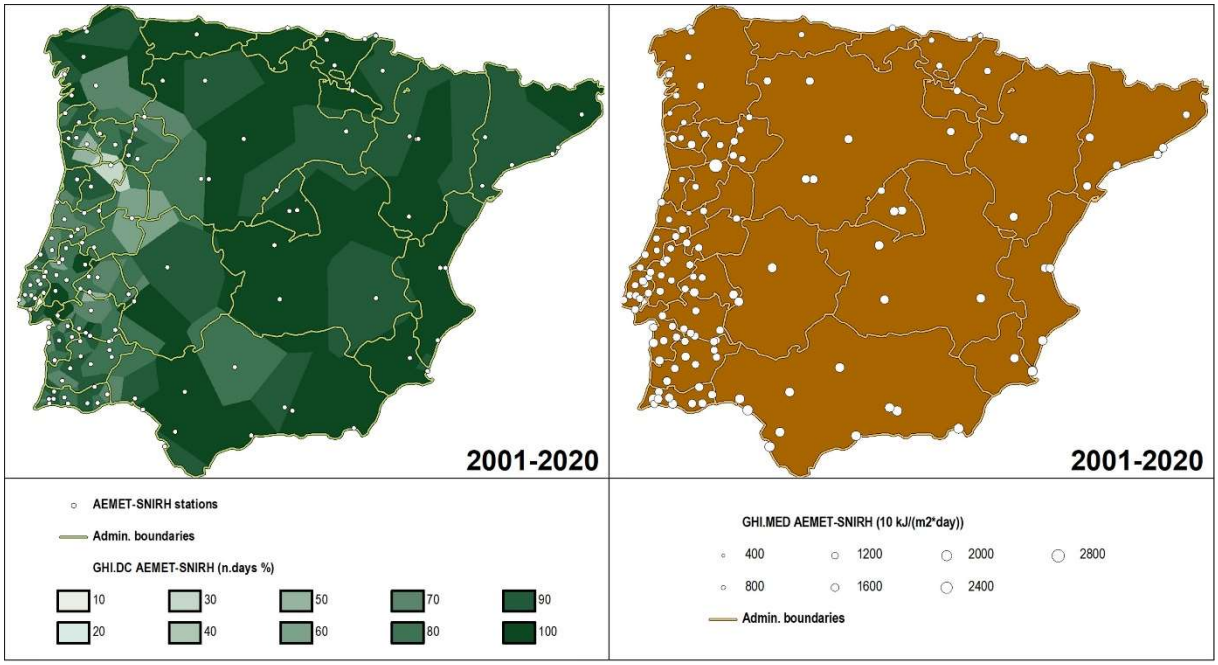

Figure 4. Spatiotemporal data completeness and GHI of AEMET-SNIRH network (2001–2020).
GHI.DC: GHI data completeness. GHI.MED: median GHI.

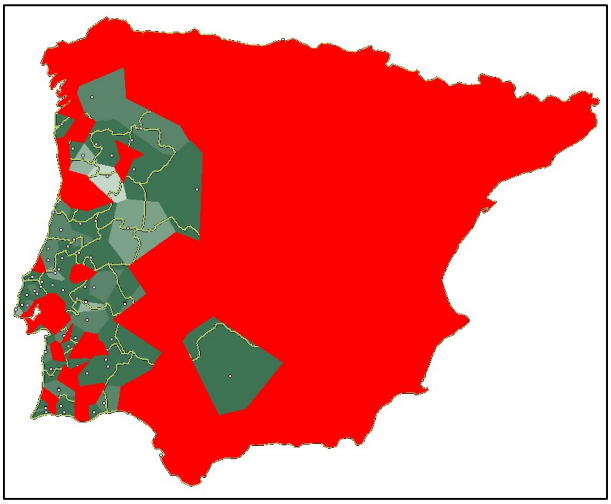

Figure 5. GHI spatiotemporal data completeness ≥ 80 % of AEMET-SNIRH network (2001–2020).
GHI data completeness ≥ 80 % in red.

5. Complementary ground-based solar radiation networks

Without any intention to make exhaustive diagnoses of other station networks that provide solar radiation data, it is valuable to provide a brief overview of them. With this aim, the most important networks that cover the three areas studied in this article (Spain, Portugal, and the Iberian Peninsula) are presented and briefly discussed below.

The Baseline Surface Radiation Network (BSRN) is a network of stations worldwide; however, for peninsular Spain it only offers one station in Sarrigurren (Navarra, CENER: *Centro Nacional de Energías Renovables*) that is not available in the AEMET network. There are no stations for Portugal.

The World Radiation Data Centre (WRDC) provides GHI and DIF data, but not DNI data, from the stations of the state meteorological agencies: AEMET for Spain, and IPMA (*Instituto Português do Mar e da Atmosfera*) for Portugal, available at < http://wrdc.mgo.rssi.ru/ >. First, we checked whether

the WRDC stations for Spain include all the AEMET stations. Of the 65 AEMET stations, the WRDC only contains 51. It should be noted that the station in Navacerrada (WMO Index 8215) contains erroneous coordinates, identical to those of the León station (WMO Index 8055); the station in Reus (WMO Index 8175) contains a wrong longitude sign; the station in Salamanca (WMO Index 8202) contains an offset of 21.1 km to the W; the station in Soria (WMO Index 8148) has an offset of 19.5 km to the S; and the station in Vitoria (WMO Index 8080) has a displacement of 10.4 km towards the E. The rest of the stations have a displacement of less than 5.1 km. Considering the thematic quality of the data and given that it is not possible to make a mass download, the comparison was made with five AEMET stations (367, 1111, 1495, 4121 and 5783). We verified that the data provided by the WRDC are identical to those of the AEMET. The solar radiation data of the stations in Portugal from the IPMA (which are not published on their web portal < https://www.ipma.pt/pt/index.html >) constitute a network of 19 GHI and 8 DIF stations, and some of them started taking measurements before 2001. We did not consider it appropriate to incorporate these stations in the present study because the SNIRH is a much larger network that covers the country more densely, so that the percentage of increase would be small considering the close proximity between the stations of the two networks: minimum 8.0 km, maximum 39.6 km, and median 21.0 km. However, in very detailed studies it could be interesting to incorporate IPMA stations in years in which the SNIRH network is not especially dense (*e.g.*, years 2011 [24 stations], 2012 [21] and 2013 [17]).

Another GHI terrestrial measurement dataset, in this case exclusively for Spain, is provided by the *Sistema de Información Agroclimática para el Regadío* (SiAR), available at < https://servicio.mapa.gob.es/websiar/SeleccionParametrosMap.aspx?dst=1 >. It is a very dense network of stations in some regions (*e.g.*, the Comunitat Valenciana, Murcia and Andalucía) and is more dispersed in others (*e.g.*, Galicia, Madrid, and the southern area of Aragón). Moreover, it does not provide any data source in Asturias, Cantabria, Catalonia, La Rioja, or the Basque Country. This network does not offer direct bulk download of all peninsular data, which makes a more detailed comparative analysis difficult in the context of this article.

In Andorra, the *Servei Meteorològic Nacional d'Andorra* has five meteorological stations (Envalira, Bony de les Neres, Aixàs, Sorteny and La Margineda), providing GHI data since the year 2000 for the Envalira station, since 2015 for the Bony de les Neres station, since 2016 for the Aixàs and Sorteny stations, and since 2018 for the La Margineda station. They are all located between 950 m and 2500 m. Data are available at < https://www.meteo.ad/estacions >; however, there is a low temporal availability compared to the data presented in the previous sections. Nevertheless, it could be useful for regional studies combined with other station networks, such as those provided by the *Xarxa d'Estacions Meteorològiques Automàtiques* (XEMA) of the *Servei Meteorològic de Catalunya* (SMC) for the countries of Pallars Sobirà, Alt Urgell and Cerdanya, and thus increase the locations available in high mountain areas. The data can be downloaded individually for each station.

Gibraltar has five meteorological stations (Bleak House, Bruce's Farm, Caleta Palace Hotel, Cemetery Office, and Rosia Road), which have provided GHI data since the end of 2012 for the Bleak House and Rosia Road stations, since the end of 2013 for the Bruce's Farm and Cemetery Office stations, and since the beginning of 2014 for the Caleta Palace Hotel station. The data are available at < http://www.gibmetportal.gi/data/ >. They also have a low temporal availability compared to the data presented in the previous sections. They are not downloadable, but rather they must be requested; however, the waiting time is short. These stations can be used for regional studies combined with other nearby stations from AEMET and SiAR, and/or with those from the *Red de Información Agroclimática de Andalucía* (RIA), available at < https://www.juntadeandalucia.es/agriculturaypesca/ifapa/riaweb/web/datosabiertos > (Estévez and Gavilán 2008).

Finally, it is also worth mentioning the solar radiation data from meteorological stations of the different regional governments and/or councils, and other administrative entities within the Iberian Peninsula, which are particularly interesting for regional studies, but outside the scope of this article. The main reasons for not including them in this first approach are:

1) *Effort to download and/or validate the data.* It is common to have to download the data individually or even to have to request them through forms. It has already been seen in the previously mentioned networks that it is often necessary to perform an additional validation to that already carried

out by the distribution agencies, since errors are not uncommon. As indicated by Salazar *et al.* (2020) and Forstinger *et al.* (2021), a thorough and expert evaluation of the data provided by the stations requires a lot of checking and validation work, as the necessary QC must be applied before they can be used in further analyses.

2) *Territorial homogeneity.* A first overview of these networks reveals very different densities and series lengths depending on the area, which could end up decompensating certain integrated approaches. However, it would be interesting to apply the procedures developed in this work in analyses that use other station networks that provide observed data solar radiation, such as, for example, the GHI network of the SMC (XEMA), available at < https://www.meteo.cat/wpweb/serveis/formularis/petitio-dinformes-i-dades-meteorologiques/ >. Moreover, once the necessary filters and homogenizations have been applied, the integrated application of different networks would make it possible to have a higher density of stations in regions with complex orography and/or with singular atmospheric and climatic conditions.

6. Discussion

The study of data from stations providing solar radiation data in the Iberian Peninsula indicates that more efforts are needed to provide adequate radiation measurements, and thus reduce errors and ensure the spatial and temporal completeness of the datasets. Likewise, the homogeneity of the type of data collected is affected by the fact that the DNI and DIF measurements are more difficult to obtain than GHI measurements because the necessary instruments and/or capture mechanisms are different (pyranometers and pyrheliometers). In addition, the homogeneity of the type of data collected is also affected because the solar panels are not yet designed to differentiate between DNI and DIF. As shown in Tables 2 and 4, limitations are observed in the temporal distribution of stations that provide DNI and DIF data compared to those providing GHI.

The temporal completeness of the data was calculated by considering the number of potential observations based on the number of actual annual stations (those that provide observed data solar radiation), and the maximum possible number of stations (as if during all the years in the series all stations had been active). This showed that, although the temporal completeness of the data is total (100 %) in some years, according to the number of actual annual stations (DNI years 1981–1985, Table 2), their distribution over the territory is minimal (1 station). The number of stations in the networks has gradually increased over the years; however, the number of real annual stations is always lower in relation to the maximum possible number of stations[1], indicating a low spatial completeness during the periods 1980–2020 and 1980–2000 (Tables 1 and 3) and moderate spatial completeness during the period 2001–2020 (Table 5) in the case of AEMET, SNIRH (Table 10) and the integrated network (Table 13), although the temporal completeness of the data is high when we consider the number of real annual stations.

In the comparative analysis based on integrating the GHI observations from the temporal series 2001–2020 common to both networks, none of the three scenarios analysed (Spain, Portugal, and the Iberian Peninsula) reaches the totality of active stations possible in any year. However, the increased availability of ground stations due to combining the two networks provides more spatiotemporal detail of solar radiation data, as demonstrated by the strong and significant correlations for all observed years in the series. Nevertheless, this increase in the number of stations and the temporal completeness of the data from the AEMET-SNIRH integrated network always implies lower values than those of the network with better representation. In this context, it is important to highlight that the result of the comparative analysis between the two networks has been positive, providing a network of stations with a more densified spatial distribution, covering the entire Iberian Peninsula.

For purposes of comparison between the GHI, DNI and DIF, it is necessary to remember the importance of calculating the DNI on the horizontal plane using Equation 1. It has been possible to

[1]Maximum possible number of GHI stations: AEMET 1980–2020 = 65; AEMET 1980–2000 = 41; AEMET 2001–2020 = 54; SNIRH 2001–2020 = 88; AEMET-SNIRH 2001–2020 = 142.

verify, in Figure 2, that the values obtained from the difference between the GHI and the DIF measured by the AEMET follow the same pattern as those of the DNI calculated on the horizontal plane (DHI). This implies that, using the network provided by the World Radiation Data Centre (WRDC) with data from GHI and DIF in Portugal, it would be possible to obtain a DHI that is quite well adjusted to the observed data, which can be useful in future studies that use the DHI and DIF over the entire peninsular territory.

The quantities provided are also different depending on the network: tens of $kJ \cdot m^{-2} \cdot day^{-1}$ (energy per unit area and per day) or $W \cdot m^{-2}$ (power per unit area). This implies an added difficulty in treating radiation data, since, as we have seen, depending on the instrument, it can record data even during the night and, depending on the conversion of the required unit of measurement, the time interval for which the data is provided must be taken into account.

Although it is practically impossible to achieve 100 % reliability (Salazar *et al.* 2020), it is important to analyse the data with the aim of achieving a balance between the size of the sample and its quality (Forstinger *et al.* 2021). In this sense, it should be emphasized that the presence of radiation data during night hours is not anecdotic or limited to a few days of each month (Tables 7 and 8, and Appendix 2), which highlights the importance of filtering these measurements. Otherwise, climate analyses with solar radiation data would be contaminated with a large amount of noise. The network of GHI stations of the SNIRH required a lot of effort to filter the data. The main problems detected were the radiation data at night hours and, in less cases, the measurements recorded in random different minutes from the same hour. The radiation measurements during night hours could be of interest in modelling studies that consider infrared thermal radiation. Likewise, they could also be useful in studies on the generation of photovoltaic energy (Nielsen *et al.* 2022).

The indicators collected in the present study often have a very non-normal distribution, which has been evidenced in the different average and median values shown in the results tables (*e.g.*, Tables 7 and 8). We would like to draw attention to this to urge researchers in this field to obtain more realistic indicators of centrality and dispersion through the median and the mean absolute deviation around the median, since these indicators are not as affected as the average and standard deviation by the extreme values of the sample.

For comparative purposes, the differences observed in the two datasets evaluated (AEMET-SNIRH) may have occurred for several reasons, other than due to the different cloudiness conditions of the two countries:

- The solar radiation measurements provided by AEMET are daily, while those of SNIRH are hourly, which may lead to small calculation differences between the various time scales in the sample.
- The measuring instruments are completely different, as well as the mechanisms for operation, capture, and storage of the data.
- For the research carried out, the observed solar radiation of AEMET is understood to be properly adjusted and all the corresponding QC processes properly applied, as the station network must comply with the World Standardization Group (WSG). In this context, and as has been demonstrated in the SNIRH network, it is preferable to use hourly data to have greater control over applying QC mechanisms.
- AEMET has a long track record in the capture, storage, and management of solar radiation data, and is a consolidated network at national level that includes the designation of National Radiometric Centres (NRC) validated according to the standards of the WSG. Indeed, during the period 1980–2020 there was an increase and more stability in the number of stations and the temporal completeness of the observed data. The SNIRH has functioned for a shorter time span, 2001–2020, but the results show that the network is of great interest due to its density and wide spatial distribution, and we encourage the continuation of the current data collection policy.

7. Conclusions

Regarding the AEMET data. It is a complete network considering the extent of the territory and has significant instrumental maturity (adhering to the specifications of the World Meteorological Organization, WMO). Radiation is measured in three different ways depending on the meteorological station: global (GHI), direct (DNI) and diffuse (DIF). This provides rich but inconsistent information, as there is a noticeable asymmetry between the observations carried out at each station (DNI data monitoring is the least frequent of the three). Despite this, conversions between global, direct, and diffuse radiation data have shown that the values are consistent. There are a total of 65 stations considering the period 1980–2020, and 54 stations considering the period 2001–2020 (shared with the Portuguese SNIRH stations). In the first period, the completeness taking into account the maximum possible number of stations is particularly low: the percentage of stations that provide data with respect to the potential is 40.4 % for GHI, 32.2 % for DNI and 37.6 % for DIF. In the second period, this completeness is higher: 73.5 % for GHI, 57.9 % for DNI and 70.0 % for DIF.

Regarding the SNIRH data. This network only provides GHI data for the period 2001–2020, although the availability of hourly data makes additional QC possible, which has been shown to be both useful and necessary to deal with nightly readings, meaningless daytime readings (too many hours of light, or absurdly high values) or NoData records, which appear without any apparent pattern and too frequently to be ignored. The completeness considering the maximum possible number of stations is 54.3 %. Although the density of the network is remarkable, there is no relationship between the number of stations available each year and the temporal completeness of observations. This temporal completeness ranges between 35.1 % and 96.9 %, while the spatial completeness ranges between 50.0 % and 85.2 %, depending on the year. When comparisons are made with AEMET data, daily data should be avoided because SNIRH data are not provided by calendar days. The daily data must be obtained by integrating the hourly data of the same calendar day.

If we consider the two networks combined (AEMET-SNIRH) in the common period (2001–2020) to achieve an integrated network for the Iberian Peninsula, the following considerations must be added. The two networks provide different quantities: energy and power, respectively. After the appropriate conversions have been made, the two networks exhibit very high and significant correlations, suggesting robustness in the combined network. The AEMET observations are higher, which can be attributed to the fact that the Portuguese network has an essentially Atlantic influence, with more cloudiness, while many of the AEMET stations are in a Mediterranean climate, with less cloudiness. The Portuguese observations are denser, while the Spanish observations have areas where the distance to the nearest solar radiation station reaches between 100 and 140 km.

Our main recommendation is, in the case of the two networks, to ensure greater temporal completeness, as there is a high amount of missing data. For the AEMET network, the second recommendation is to increase spatial completeness by adding some stations to the network (or transferring them from networks of other entities such as those reported in the article) to avoid some large territorial gaps that have been noticed. For the SNIRH network it is worth making three additional recommendations: to increase the temporal completeness even more, to strengthen the QC, and to provide DIF data.

Solar radiation is a key factor in atmospheric dynamics and climate phenomena (Sánchez-Lorenzo *et al.* 2013, AEMET 2024, IDEAM 2024). This study contributes and allows us to continue advancing in the knowledge of solar radiation, providing elements to stimulate the growth of an integrated and consistent network for collecting and processing data in a territory as sensitive as the Iberian Peninsula. It is important to consider the stations from other sources in these regions where the State networks are insufficient, for example in areas with a complex orography and/or with singular atmospheric and climatic conditions (*e.g.*, high mountain stations). However, these complementary networks often do not provide data easily, an issue that needs to be improved in order to use them more effectively considering the effort involved in running them.

The results presented here are important given the current climate change situation at the global level, and particularly due to the significant impacts on the entire Iberian Peninsula with an increase in extreme climate events as a result of its complexity and geographical and climatic heterogeneity. Indeed,

more observed data that are more accurately processed, including harmonization, are needed. Moreover, this must be achieved while guaranteeing the maximum possible number of stations with an appropriate spatiotemporal distribution. This will make it possible to study the new climate scenarios, and current adaptation needs (Pratt 2022, Olcina 2023, Rodríguez 2023).

8. Future research

In future work, the data from these networks, already processed in the indicated manner, will be used to calibrate, and validate new improved solar radiation models, and carry out new works of a biological or geographical nature, energy studies, etc. For example, and following the same line of the work by Roca-Fernández *et al.* (2022), we would like to apply this knowledge to vegetation distribution studies and to the modelling of general climate phenomena, particularly to study of the droughts. Indeed, better knowledge about solar radiation will improve the calculation of the PET (Potential Evapotranspiration, Penman-Monteith 1965, Hargreaves and Samani 1985, Samani 2000), variable related to temperatures and included in the SPEI (Standardised Precipitation-Evapotranspiration Index, Vicente-Serrano *et al.* 2010) climate drought index.

9. Acknowledgments

This study is part of the NEWFORLAND research project (RTI2018-099397-B-C21), developed by the GRUMETS research group. C.R-F. is the recipient of a FI-DGR training grant (2020FI-B00669), co-financed by the *Agència de Gestió d'Ajuts Universitaris i d'Investigació* (AGAUR) of the *Generalitat de Catalunya* and by the European Union (EU), through the European Social Fund (ESF), and by the *Universitat Autònoma de Barcelona* (UAB). X.P. is the recipient of an ICREA Academia Excellence in Research Grant (2023–2027).

We thank the state meteorological organization AEMET-Spain for the information provided on solar radiation (data, metadata, and technical reports) < https://www.aemet.es/es/sede_electronica >, as well as SNIRH-Portugal for its data distribution policy < https://snirh.apambiente.pt/ > and for the help provided in solving specific doubts.

Bibliographic references

Aalto, J.; Riihimäki, H.; Meineri, E.; Hylander, K.; Luoto, M. (2017). "Revealing topoclimatic heterogeneity using meteorological station data". *Int. J. Climato.,* 37(S1), pp. 544-556. Doi: 10.1002/joc.5020

AEMET. (2024). *La radiación solar.* Ministerio de Medio Ambiente y Medio Rural y Marino (Gobierno de España). https://www.aemet.es/documentos/es/eltiempo/observacion/radiacion/Radiacion_Solar.pdf

Bárcena, A.; Prado, A.; Samaniego, J.; Pérez, R. (2011). *Efectos del cambio climático en la costa de América Latina y el Caribe. Dinámicas, tendencias y variabilidad climática.* IHCantabria. https://www.cepal.org/sites/default/files/events/files/dinamicas_tendencias_y_variabilidad_climatica_y_el_caribe.pdf

Brunet, M.; Jones, P.D.; Sigró, J.; Saladié, O.; Aguilar, E.; Moberg, A.; Della-Marta, P.M.; Lister, D.; Walther, A.; Lopez, D. (2007). "Temporal and spatial temperature variability and change over Spain during 1850–2005". *Journ. Geophys. Res.,* 112(D12), D12117. Doi: 10.1029/2006JD008249

Calbó, J.; Sánchez-Lorenzo, A.; Martín-Vide, J.; Brunetti, M. (2008). "Aspectos climatológicos y evolución temporal de la nubosidad en la Península Ibérica (1961–2004)". *VI Congreso Internacional de la Asociación Española de Climatología*, 08-11 October 2008, Tarragona (Spain). https://www.divulgameteo.es/Aspectos-climatologicos-y-evolucion-temporal-de-la-nubosidad-en-la-Peninsula-Iberica-1961-2004/

Carnicer, J.; Domingo-Marimon, C.; Ninyerola, M.; Camarero, J.J.; Bastos, A.; López-Parages. J.; Blanquer, L.; Rodríguez-Fonseca, B.; Lenton, T.M.; Dakos, V.; Ribas, M.; Gutiérrez, E.; Peñuelas, J.; Pons, X. (2019). "Regime shifts of Mediterranean forest carbon uptake and reduced resilience driven by multidecadal ocean surface temperatures". *Global Change Biology,* 25(8), pp. 2825-2840. Doi: 10.1111/gcb.14664

Carrasco, E.; Carramiñana, A. (1998). *Cuando la Luna brilla más que el Sol.* Diario Síntesis. https://www.inaoep.mx/~rincon/gluna.html

Doblas-Miranda, E.; Alonso, R.; Arnan, X.; Bermejo, V.; Brotons, L.; De las Heras, J.; Estiarte, M.; Hódar, J.A.; Llorens, P.; Lloret, F.; López-Serrano, F.R.; Martínez-Vilalta, J.; Moya, D.; Peñuelas, J.; Pino, J.; Rodrigo, A.; Roura-Pascual, N.; Valladares, F.; Vilà, M.; Zamora, R.; Retana, J. (2017). "A review of the combination among global change factors in forests, shrublands and pastures of the Mediterranean Region: Beyond drought effects". *Global Planet. Change,* 148, pp. 42-54. Doi: 10.1016/j.gloplacha.2016.11.012

Dozier, J. (1989). "Spectral Signature of Alpine Snow Cover from the Landsat Thematic Mapper". *Remote Sens. Environ.,* 28, pp. 9-22. Doi: 10.1016/0034-4257(89)90101-6

EcoInventos. (2022). *¿Es posible cargar los paneles solares con luz artificial?* https://ecoinventos.com/cargar-los-paneles-solares-con-luz-artificial/#google_vignette

Estévez, J.; Gavilán, P. (2008). "Procedimientos de validación de datos de estaciones meteorológicas automáticas. Aplicación a la Red de Información Agroclimática de Andalucía". *II Jornadas de Gestores y Usuarios de Redes Agrometeorológicas,* December 2008, Murcia (Spain). https://www.researchgate.net/publication/280665071_Procedimientos_de_validacion_de_datos_de_estaciones_meteorologicas_automaticas_Aplicacion_a_la_Red_de_Informacion_Agroclimatica_de_Andalucia

Factorenergia. (2023). *¿Qué es la radiación solar directa e indirecta? Y su importancia en la fotovoltaica.* https://www.factorenergia.com/es/blog/autoconsumo-electrico/que-es-la-radiacion-solar-directa-e-indirecta-y-su-importancia-en-la-fotovoltaica/

Forstinger, A.; Wilbert, S.; Jensen, A.; Kraas, B.; Fernández-Peruchena, C.; Gueymard, C.A.; Ronzio, D.; Yang, D.; Collino, E.; Martinez, J.P. (2021). "Expert quality control of solar radiation ground data sets". *Proceedings of SWC 2021: ISES Solar World Congress* (International Solar Energy Society), pp. 1037-1048, October 2021, France. Doi: 10.18086/swc.2021.38.02

Gu, L.; Baldocchi, D.; Verma, S.B.; Black, T.A.; Vesala, T.; Falge, E.M.; Dowty, P.R. (2002). "Advantages of diffuse radiation for terrestrial ecosystem productivity". *Journ. Geophys. Res.,* 107(D6), 4050. Doi: 10.1029/2001JD001242

Hargreaves, G.H.; Samani, Z. (1985). "Reference crop evapotranspiration from temperature". *Applied Engin. Agric.,* 1(2). Doi: 10.13031/2013.26773

I'MNOVATION. (2023). *¿Energía solar de noche? Un sensor pionero podría ser la clave.* https://www.imnovation-hub.com/es/energia/energia-solar-noche-sensores/

IDEAM. (2024). *Características de la radiación solar.* Ministerio de Ambiente y Desarrollo Sostenible (Gobierno de Colombia). http://www.ideam.gov.co/web/tiempo-y-clima/caracteristicas-de-la-radiacion-solar#

Iglesias, A.; Quiroga, S. (2007). "Measuring the risk of climate variability to cereal production at five sites in Spain". *Climato. Res.,* 34(1), pp. 47-57. Doi: 10.3354/cr034047

Iturbe, M. (2019). *Radiación solar en paneles solares: orientación, inclinación y sombras.* https://www.caloryfrio.com/energias-renovables/energia-solar/instalacion-captadores-solares.html

Li, L.; Xiaozhou, X.; Hailong, Z.; Jiangfeng Y.; Qinhuo L.; Shanshan Y.; Jianguang W. (2015). "A method for estimating hourly photosynthetically active radiation (PAR) in China by combining geostationary and polar-orbiting satellite data". *Remote Sens. Environ.,* 165(2015), pp. 14-26. Doi: 10.1016/j.rse.2015.03.034

MANSUR. (2023). *¿Los paneles solares funcionan de noche con la luz de la luna?* https://www.mansur-solar.com/2021/10/los-paneles-solares-funcionan-de-noche/

Mercado, L.M.; Bellouin, N.; Sitch, S.; Boucher, O.; Huntingford, C.; Wild, M.; Cox, P.M. (2009). "Impact of changes in diffuse radiation on the global land carbon sink". *Nature,* 458(7241), pp. 1014-1017. Doi: 10.1038/nature07949

Myneni, R.B.; Williams, D.L. (1994). "On the relationship between FAPAR and NDVI". *Remote Sens. Environ.,* 49(1994), pp. 200-211. Doi: 10.1016/0034-4257(94)90016-7

Nielsen, M.P.; Pusch, A.; Sazzad, M.H.; Pearce, P.M.; Reece, P.J., Ekins-Daukes, N.J. (2022). "Thermoradiative power conversion from HgCdTe photodiodes and their current-voltage characteristics". *ACS Photonics,* 9, pp. 1535-1540. Doi: 10.1021/acsphotonics.2c00223

Ninyerola, M.; Pons, X.; Roure, J.M. (2000). "A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques". *Int. J. Climato.,* 20(14), pp. 1823-1841. Doi: 10.1002/1097-0088(20001130)20:14<1823::AID-JOC566>3.0.CO;2-B

OECC. (2022). *Cambio Climático: Impactos, Adaptación y Vulnerabilidad. Guía resumida del sexto informe de evaluación del IPCC. Grupo de trabajo II.* Ministerio para la Transición Ecológica y el Reto Demográfico (Gobierno de España). https://www.miteco.gob.es/content/dam/miteco/es/cambio-climatico/temas/impactos-vulnerabilidad-y-adaptacion/ipcc-guia-resumida-gt2-imp-adap-vuln-ar6_tcm30-548667.pdf

Ohmura, A.; Bauder, A.; Mueller, H.; Kappenberger, G. (2007). "Long-term change of mass balance and the role of radiation". *Ann. Glaciology,* 46, pp. 367-374. Doi: 10.3189/172756407782871297

Olcina, J. (2023). *El clima de la Península Ibérica está cambiando.* RTVE audio 24 horas (RNE). https://www.rtve.es/play/audios/24-horas/dana-jorge-olcina-clima-peninsula-iberica-cambiando/6961646/

Padial-Iglesias, M.; Pons, X.; Serra, P.; Ninyerola, M. (2022). "Does the gap-filling method influence long-term (1950-2019) temperature and precipitation trend analyses?". *GeoFocus,* 29, pp. 5-33. Doi: 10.21138/GF.773

Penman-Monteith, J.L. (1965). "Evaporation and environment". *Symposia of the Society for Experimental Biology,* 19, pp. 205-234. https://repository.rothamsted.ac.uk/download/8ae229c1c0ea4f617750d8e98d2ee6c356c306fc01a39bb584a18eb112f443e1/3879831/Monteith65.pdf

Pfeifroth, U.; Trentmann, J.; Kothe, S. (2019). *Product User Manual. Meteosat Solar Surface Radiation and Effective Cloud Albedo Climate Data Record. SARAH-2.1 climate data records. EUMETSAT CM SAF Climate Monitoring.* https://www.cmsaf.eu/SharedDocs/Literatur/document/2019/saf_cm_dwd_pum_meteosat_hel_sarah_2_4_pdf.pdf?__blob=publicationFile

Pons, X. (1996). "Estimación de la Radiación Solar a partir de modelos digitales de elevaciones. Propuesta metodológica". *VII Coloquio de Geografía Cuantitativa, Sistemas de Información Geográfica y Teledetección,* Juaristi, J. & Moro, I. (eds.) Vitoria-Gasteiz (Spain). https://ddd.uab.cat/pub/poncom/1996/200247/Pons_1996_Estimacion_de_la_radiacion_solar_a_partir_de_modelos_digitales_de_elevaciones_Propuesta_metodologica_OCR.pdf

Pons, X. (2024). *MiraMon. Geographic Information System and Remote Sensing software.* Centre de Recerca Ecològica i Aplicacions Forestals (CREAF). https://www.miramon.cat/Index_usa.htm. Help for used applications: Astres https://www.miramon.cat/help/eng/msa/astres.htm; Thiessen https://www.miramon.cat/help/eng/msa/thiessen.htm

Pons, X.; Ninyerola, M. (2008). "Mapping a topographic global solar radiation model implemented in a GIS and refined with ground data". *Int. J. Climatol.,* 28, pp. 1821-1834. Doi: 10.1002/joc.1676

Pratt, S.E. (2022). *Iberian Peninsula Drought.* Earth Observatory (NASA). https://earthobservatory.nasa.gov/images/149469/iberian-peninsula-drought

Rahman, M.M.; Lamb, D.W.; Stanley, J.N. (2015). "The impact of solar illumination angle when using active optical sensing of NDVI to infer fAPAR in a pasture canopy". *Agric. For. Met.,* 202(2015), pp. 39-43. Doi: 10.1016/j.agrformet.2014.12.001

Roca-Fernández, C.; Ninyerola, M.; Pons, X. (2022). "Mejoras en el cálculo de los Modelos Digitales de Radiación Solar a partir del tratamiento del MDE y de la incorporación del patrón espaciotemporal de la profundidad óptica atmosférica: Resultados preliminares para la Península Ibérica". *XIX Congreso de Tecnologías de la Información Geográfica – Las TIG al Servicio de los ODS* (Departamento de Geografía y Ordenación del Territorio de la Universidad de Zaragoza, y la Asociación de Geógrafos Españoles), pp. 391-400, 12-14 September 2022, Zaragoza (Spain). https://zaguan.unizar.es/record/119771

Rodríguez, H. (2023). *La Península Ibérica, la región más seca del continente.* National Geographic España. https://www.nationalgeographic.com.es/ciencia/por-que-cada-vez-llueve-menos-peninsula-iberica_18450

Ruiz-Arias, J.A.; Pozo-Vázquez, D.; Santos-Alamillos, F.J.; Lara-Fanego, V.; Tovar-Pescador, J. (2011). "A topographic geostatistical approach for mapping monthly mean values of daily global solar radiation: A case study in southern Spain". *Agr. For. Met.,* 151(2011), pp. 1812-1822. Doi: 10.1016/j.agrformet.2011.07.021

Salazar, G.; Gueymard, C.; Bezerra-Galdino, J.; de Castro-Vilela, O.; Fraidenraich, N. (2020). "Solar irradiance time series derived from high-quality measurements, satellite-based models, and reanalyses at a near-equatorial site in Brazil". *Renew. Sust. Energ. Rev.,* 117(2020), 109478. Doi: 10.1016/j.rser.2019.109478

Samani, Z. (2000). "Estimating solar radiation and evapotranspiration using minimum climatological data". *Journ. Irrigation and Drainage Engin.,* 126(4). Doi: 10.1061/(ASCE)0733-9437(2000)126:4(265)

Sánchez-Lorenzo, A.; Calbó, J.; Wild, M. (2013). "Global and diffuse solar radiation in Spain: Building a homogeneous dataset and assessing their trends". *Global Planet. Change,* 100(2013), pp. 343-352. Doi: 10.1016/j.gloplacha.2012.11.010

Solargis. (2016). *Solargis Solar Resource Database. Description and Accuracy.* https://solargis2-web-assets.s3.eu-west-1.amazonaws.com/public/Uploads/279e8bb216/Solargis-database-description-and-accuracy.pdf

Solarni Paneli. (2023). *Paneles solares. ¿Los paneles solares necesitan Sol o solo luz?* https://solarnipaneli.energy/es/los-paneles-solares-necesitan-sol-o-solo-luz/

SolarPlak. (2023). *Cómo funcionan las baterías solares.* https://solarplak.es/energia/como-funcionan-las-baterias-solares/

Stanhill, G.; Cohen, S. (2001). "Global dimming: a review of the evidence for a widespread and significant reduction in global radiation". *Agric. For. Met.,* 107(4), pp. 255-278. Doi: 10.1016/S0168-1923(00)00241-0

Suri, M.; Cebecauer, T.; Skoczek, A.; Marais, R.; Mushwana, C.; Reinecke, J.; Meyer, R. (2014). "Cloud cover impact on photovoltaic power production in South Africa". *Proceedings of the SASEC 2014 Conference,* 28 January 2014, Port Elizabeth, South Africa*,* A20, pp. 309-317. https://solargis.com/es/publications/2014/all/intermitency-analysis

Trenberth, K.E.; Fasullo, J.T.; Kiehl, J. (2009). "Earth's Global Energy Budget". *Bull. Amer. Met. Soc.,* 90(3), pp. 311-324. Doi: 10.1175/2008BAMS2634.1

Vicente-Serrano, S.M.; Azorín-Molina, C.; Peña-Gallardo, M.; Tomas-Burguera, M.; Domínguez-Castro, F.; Martín-Hernández, N.; Beguería, S.; El Kenawy, A.; Noguera, I.; García, M. (2019). "A high-resolution spatial assessment of the impacts of drought variability on vegetation activity in Spain from 1981 to 2015". *Digital.CSIC. Natural Earth System Sciences,* 19, pp. 1189-1213. Doi: 10.5194/nhess-19-1189-2019

Vicente-Serrano, S.M.; Beguería, S.; López-Moreno, J.I. (2010). "A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index". *Journ. Climato.,* 23, pp. 1696-1718. Doi: 10.1175/2009JCLI2909.1

Virtuani, A.; Pravettoni, M.; Parini, L.; Morganti, L.; Skoczek, A.; Betak, J.; Mussetta, M.; Marchionna, S. (2017). "Where has all the power gone? A health check of italian solar electricity in 2016". *33rd EU PVSEC,* 25-29 September 2017, Amsterdam, 6Bv.2.37, pp. 2401-2405. https://solargis.com/publications/2017/all/regional-studies

Wikipedia. (2023). *Moonlight* [last update: 05.09.2023]. https://en.wikipedia.org/wiki/Moonlight#:~:text=with%20its%20brightness%22.-,Illumination,reach%20up%20to%200.32%20lux

Wild, M. (2009). "Global dimming and brightening: a review". *Journ. Geophys. Res.,* 114(D10), D00D16. Doi: 10.1029/2008JD011470

Wild, M. (2012). "Enlightening global dimming and brightening". *Bull. Amer. Met. Soc.,* 93(1), pp. 27-37. Doi: 10.1175/BAMS-D-11-00074.1

Wild, M.; Ohmura, A.; Makowski, K. (2007). "Impact of global dimming and brightening on global warming". *Geophys. Res. Lett.,* 34(4), L04702. Doi:10.1029/2006GL028031

Wild, M.; Roesch, A.; Ammann, C. (2012). "Global dimming and brightening – Evidence and agricultural implications". *CAB Reviews,* 7(003), pp. 1-7. Doi: 10.1079/PAVSNNR20127003

Williams, D.R. (2024*). Moon Fact Sheet* [last update: 11.01.2024]. NASA. https://nssdc.gsfc.nasa.gov/planetary/factsheet/moonfact.html

Zhang, Y.; Li, X.; Bai, Y. (2015). "An integrated approach to estimate shortwave solar radiation on clear-sky days in rugged terrain using MODIS atmospheric products". *Sol. Energ.,* 113, pp. 347-357. Doi: 10.1016/j-solener.2014.12.028

APPENDIX 1. Additional annual tables and figures

Table A1. Annual GHI, DNI, DIF temporal data completeness of AEMET network (1980–2020).

GHI.DC AEMET 1980–2020				
Year	n.stations	n.days [4]	n.days [5]	n.days (%)
1980	8	2 559	2 928	87.4
1981	8	2 589	2 920	88.7
1982	10	2 771	3 650	75.9
1983	16	4 598	5 840	78.7
1984	20	5 732	7 320	78.3
1985	18	4 475	6 570	68.1
1986	17	5 150	6 205	83.0
1987	16	4 930	5 840	84.4
1988	14	3 743	5 124	73.0
1989	13	3 871	4 745	81.6
1990	14	3 881	5 110	75.9
1991	15	4 177	5 475	76.3
1992	15	4 242	5 490	77.3
1993	14	4 269	5 110	83.5
1994	13	4 531	4 745	95.5
1995	13	3 880	4 745	81.8
1996	22	6 098	8 052	75.7
1997	23	7 493	8 395	89.3
1998	24	7 773	8 760	88.7
1999	25	8 336	9 125	91.4
2000	28	8 554	10 248	83.5
2001	28	8 058	10 220	78.8
2002	27	7 432	9 855	75.4
2003	33	9 381	12 045	77.9
2004	38	12 470	13 908	89.7
2005	41	11 732	14 965	78.4
2006	43	13 752	15 695	87.6
2007	45	15 322	16 425	93.3
2008	45	15 811	16 470	96.0
2009	46	12 088	16 790	72.0
2010	46	15 070	16 790	89.8
2011	46	16 225	16 790	96.6
2012	48	16 448	17 568	93.6
2013	48	17 349	17 520	99.0
2014	48	17 248	17 520	98.4
2015	49	17 477	17 885	97.7
2016	49	17 800	17 934	99.3
2017	50	17 552	18 250	96.2
2018	49	17 280	17 885	96.6
2019	48	16 927	17 520	96.6
2020	46	14 627	16 836	86.9

DNI.DC AEMET 1980–2020				
Year	n.stations	n.days [4]	n.days [5]	n.days (%)
1980	1	356	366	97.3
1981	1	365	365	100.0
1982	1	365	365	100.0
1983	1	365	365	100.0
1984	1	366	366	100.0
1985	1	365	365	100.0
1986	1	362	365	99.2
1987	1	363	365	99.5
1988	2	390	732	53.3
1989	2	497	730	68.1
1990	2	603	730	82.6
1991	2	688	730	94.2
1992	2	720	732	98.4
1993	2	727	730	99.6
1994	2	682	730	93.4
1995	2	726	730	99.5
1996	4	1 073	1 464	73.3
1997	4	1 293	1 460	88.6
1998	4	1 231	1 460	84.3
1999	9	1 864	3 285	56.7
2000	9	2 540	3 294	77.1
2001	10	2 488	3 650	68.2
2002	10	2 807	3 650	76.9
2003	10	2 694	3 650	73.8
2004	10	2 827	3 660	77.2
2005	16	3 530	5 840	60.4
2006	18	3 839	6 570	58.4
2007	19	5 713	6 935	82.4
2008	19	6 296	6 954	90.5
2009	19	4 851	6 935	69.9
2010	20	6 190	7 300	84.8
2011	19	6 692	6 935	96.5
2012	20	6 873	7 320	93.9
2013	21	7 293	7 665	95.1
2014	21	7 450	7 665	97.2
2015	21	7 502	7 665	97.9
2016	21	7 542	7 686	98.1
2017	22	7 513	8 030	93.6
2018	21	7 269	7 665	94.8
2019	21	7 443	7 665	97.1
2020	21	7 416	7 686	96.5

DIF.DC AEMET 1980–2020				
Year	n.stations	n.days [4]	n.days [5]	n.days (%)
1980	1	362	366	98.9
1981	1	365	365	100.0
1982	2	458	730	62.7
1983	4	1 268	1 460	86.8
1984	7	1 620	2 562	63.2
1985	7	1 698	2 555	66.5
1986	6	1 599	2 190	73.0
1987	5	1 426	1 825	78.1
1988	5	1 116	1 830	61.0
1989	4	1 094	1 460	74.9
1990	4	1 239	1 460	84.9
1991	6	1 508	2 190	68.9
1992	6	1 561	2 196	71.1
1993	6	1 556	2 190	71.1
1994	6	1 819	2 190	83.1
1995	6	1 533	2 190	70.0
1996	7	1 775	2 562	69.3
1997	7	2 042	2 555	79.9
1998	8	2 430	2 920	83.2
1999	15	3 825	5 475	69.9
2000	17	4 775	6 222	76.7
2001	17	4 938	6 205	79.6
2002	20	5 241	7 300	71.8
2003	21	6 272	7 665	81.8
2004	21	6 209	7 686	80.8
2005	25	5 278	9 125	57.8
2006	27	7 204	9 855	73.1
2007	28	8 916	10 220	87.2
2008	29	9 703	10 614	91.4
2009	29	7 453	10 585	70.4
2010	29	9 143	10 585	86.4
2011	29	10 195	10 585	96.3
2012	31	10 531	11 346	92.8
2013	32	11 260	11 680	96.4
2014	32	11 255	11 680	96.4
2015	31	11 102	11 315	98.1
2016	31	11 134	11 346	98.1
2017	32	11 064	11 680	94.7
2018	31	10 757	11 315	95.1
2019	31	10 782	11 315	95.3
2020	31	10 606	11 346	93.5

Legend explained in Table 2.

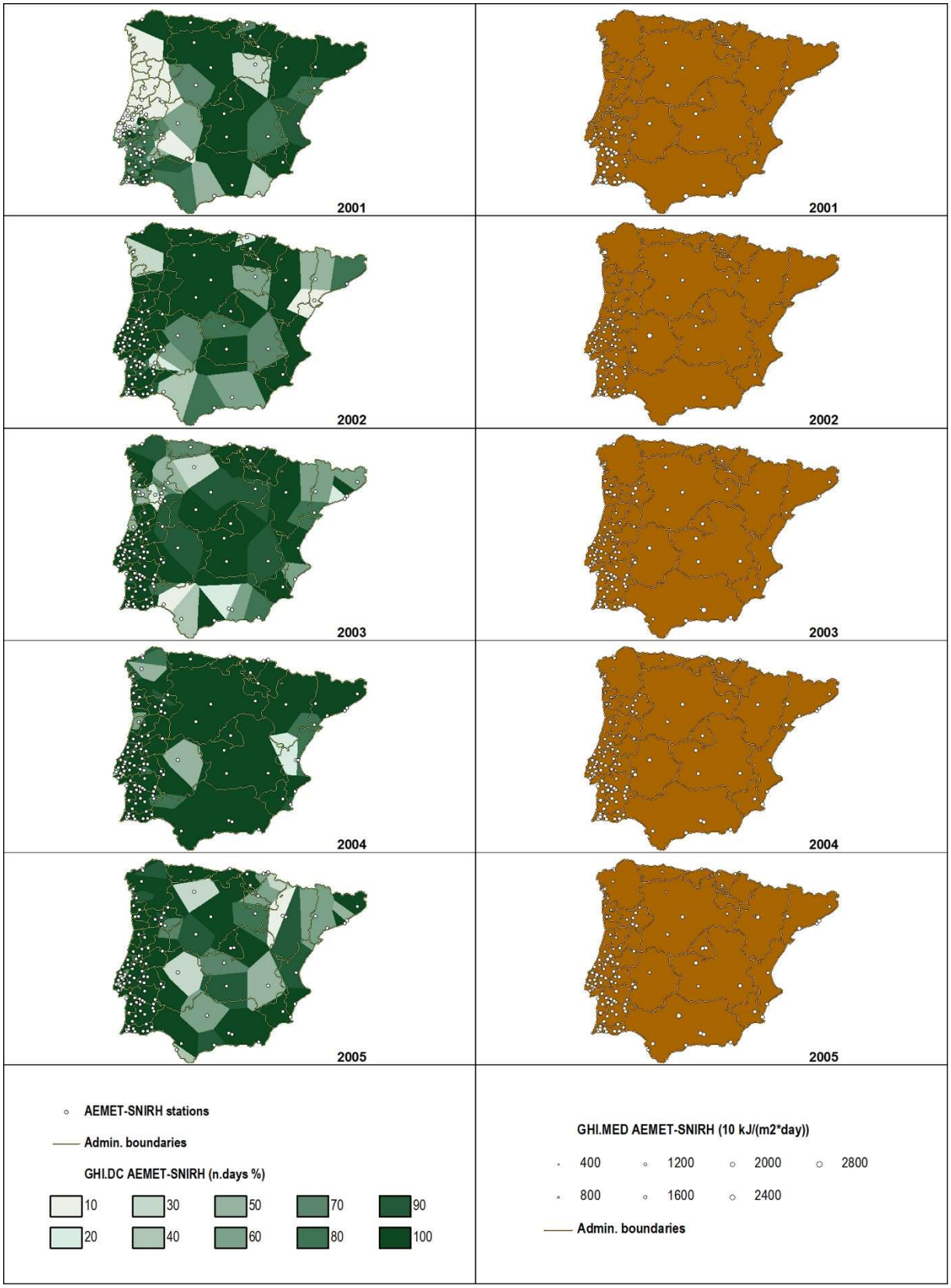

Figure A1 (1/4). Annual spatiotemporal data completeness and GHI of AEMET-SNIRH network (2001–2020).

Legend explained in Figure 4.

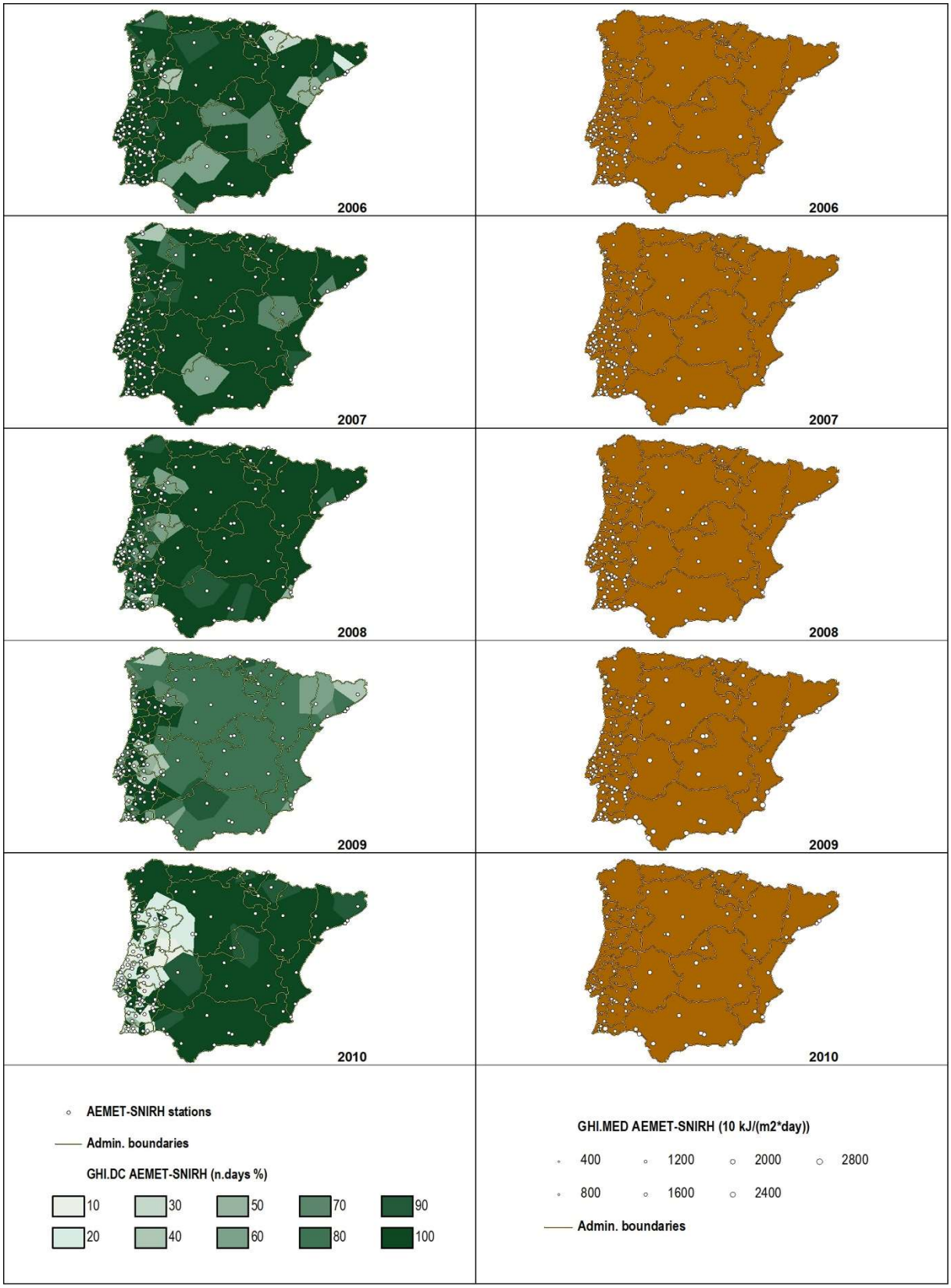

Figure A1 (2/4). Annual spatiotemporal data completeness and GHI of AEMET-SNIRH network (2001–2020).
Legend explained in Figure 4.

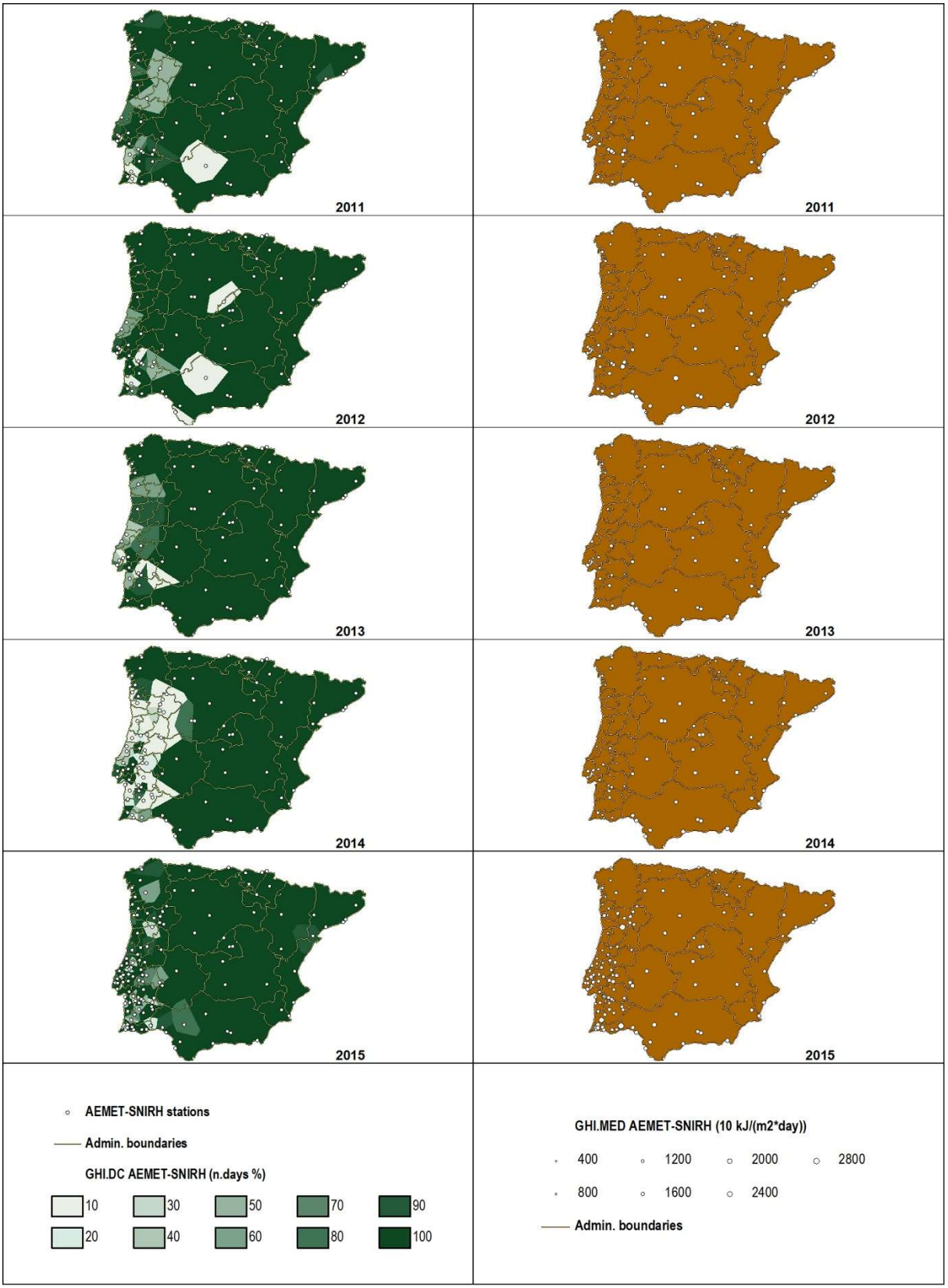

Figure A1 (3/4). Annual spatiotemporal data completeness and GHI of AEMET-SNIRH network (2001–2020).

Legend explained in Figure 4.

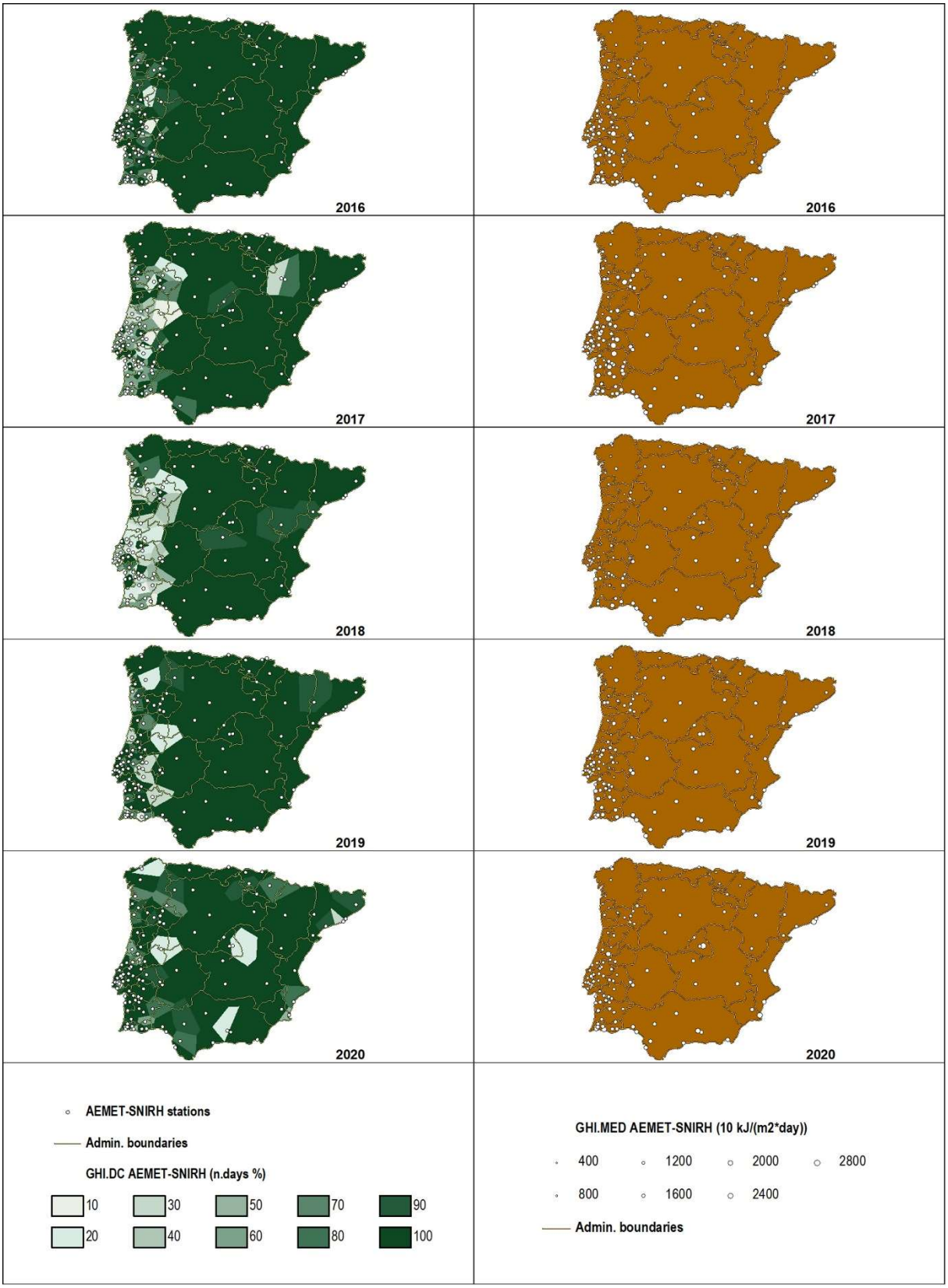

Figure A1 (4/4). Annual spatiotemporal data completeness and GHI of AEMET-SNIRH network (2001–2020).
Legend explained in Figure 4.

APPENDIX 2. Detailed analysis of the presence of GHI during nighttime hours in the SNIRH network

To illustrate that the presence of data during the night is not anecdotic or limited to a few days during each month in the network of SNIRH stations (results in 10 $kJ \cdot m^{-2} \cdot day^{-1}$), we can start by looking at the year 2003 (Table 7) and the GHI measured during night hours: 95 % of the data are below 76.0 and the minimum values are not zero (values are between 0.4 and 92.7). The 3rd quartile of the data (75 %) is below 14.6, with a median of 7.9 and the mean absolute deviation is 11.9, while the average is 16.9 and the standard deviation is 21.8. On a monthly basis, the maximum values for the months of January (109.0), February (111.7), March (230.0), October (107.7), November (118.4) and December (106.1) should be highlighted, as 95 % of the data are between 78.1 and 94.5, and 75 % below 16.7. The maximum dispersion of the temporal representativeness of each month is 4.3 days of mean absolute deviation around the median, and 6.4 days of standard deviation, both observed measures in April. The average of the GHI data is between 25.7 and 30.4 days, and the median between 28 and 31 days, coinciding with the 3rd quartile and the maximum value.

During the year 2011 (Table 7) the results are close to those obtained during the year 2003, despite the difference in stations distributed over the territory: 71 stations in 2003 and 24 in 2011. 95 % of the GHI measured at night hours is below 57.6 and the minimum values are not zero (values are between 1.1 and 84.8). A total 75 % of the data are below 15.9, with a median of 10.9 and mean absolute deviation of 9.2, while the average is 15.1 and the standard deviation is 16.7. On a monthly basis, the maximum values for the months of August (153.5), September (182.1), October (194.9) and November (196.6) should be highlighted, coinciding with 95 % of the data during the months of September, October, and November, and 75 % of the sample data are between 11.0 and 24.2. The maximum dispersion of the temporal representativeness of each month is 3.6 days of mean absolute deviation around the median in April, and 7.6 days of standard deviation in August. The average of the GHI data is between 24.4 and 30.1 days, and the median between 28 and 31 days, coinciding with the 3rd quartile and the maximum value.

The year 2019 shows results with a much higher dispersion than the two previous years (2003 and 2011) (Table 7). 95 % of the GHI measured at night is below 180.0 and can take values between 1.9 and 654.2. A total 75 % of the data is below 53.1, with a median of 10.8 and its mean absolute deviation is 46.0, while the average is 51.4 and the standard deviation is 114.1. Monthly, although the minimum, the 1st quartile (25 %) and the median have similar values in the three years (2003, 2011 and 2019) (Table 7). It is noteworthy that the maximum values exceed 264.8 (May) and reach up to 2744.9 (December), and that 95 % of the data are between 108.9 (February) and 292.1 (November). 75 % of the data for the year 2019, despite being higher than during the years 2003 and 2011, are between 27.1 (May) and 59.5 (December). The maximum dispersion of the temporal representativeness of each month is 7.8 days of mean absolute deviation around the median, and 9.5 days of standard deviation, both observed measures in December. The average of the GHI data is between 22.9 and 29.9 days, and the median between 28 and 31 days, coinciding with the 3rd quartile and the maximum value.

The results obtained from GHI in the night hours of station 21A_01C are shown in Table 8. It can be seen that the year 2019 would introduce a very large amount of noise in the radiation values. This would give a high dispersion to the data if measurements during the night were taken into account in the calculation of the temporal series (95 % of the data are below 1843.7, the minimum is 0.4, the maximum is 2365.6, the average is 616.9, the median is 8.1, the standard deviation is 703.3 and the mean absolute deviation around the median is 614.4). However, the temporal representativeness of the station is low in 2017 (85 days with data), although it increases in 2019 and 2020 (254 and 228 days with data).