Renewable Energy 189 (2022) 1176–1193

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

Spatio-temporal characterization of long-term solar resource using spatial functional data analysis: Understanding the variability and complementarity of global horizontal irradiance in Ecuador

Mariela Tapia [a, *], Detlev Heinemann [b], Daniela Ballari [c], Edwin Zondervan [d]

[a] Resilient Energy Systems Research Group, University of Bremen, Enrique-Schmidt-Str. 7, 28359, Bremen, Germany
[b] Energy Meteorology Group, Institute of Physics, University of Oldenburg, D-26111, Oldenburg, Germany
[c] IERSE - Science and Technology Faculty, Universidad del Azuay, Avenida 24 de mayo 7-77, 010151, Cuenca, Ecuador
[d] Laboratory of Process Systems Engineering, University of Twente, Horst - Meander, De Horst 2, 7522 LW, Enschede, the Netherlands

ARTICLE INFO

Article history:
Received 15 November 2021
Received in revised form
28 January 2022
Accepted 7 March 2022
Available online 9 March 2022

Keywords:
Solar irradiance
Spatio-temporal variability
Energy meteorology
Functional data analysis
Clustering
Ecuador

ABSTRACT

Understanding the spatio-temporal variability of the solar resource is crucial to effectively support solar power utilization. Unfortunately, long-term and high-resolved measurements of solar irradiance are generally scarce, challenging the characterization for larger areas. In this paper, we propose a methodology to characterize the spatio-temporal variability of global horizontal irradiance (GHI) at a regional scale using long-term satellite-derived data. Spatial functional data analysis (sFDA) is used to identify areas with similar intra-annual variability patterns. The methodology is applied to a 21-year period data on Ecuador retrieved from the National Solar Radiation Database. Being the first time that sFDA is used for this purpose, the results indicate that it provides an appropriate basis for the interannual variability and complementarity analyses. In Ecuador's mainland, twenty-two subregions with four seasonal patterns are identified. The highest GHI potential (5.4 $kWhm^{-2}d^{-1}$) with the lowest variability (3.4%) is found in the Inter-Andean valleys. Further, seasonal complementarities between the coast and western Andes are identified. In Galapagos, high values are found over all islands ($\geq$4.8 $kWhm^{-2}d^{-1}$), characterized by three subregions with one seasonal pattern. Our findings provide the first comprehensive spatio-temporal characterization of GHI in Ecuador, which aims at supporting a sustainable energy transition in the country.

1. Introduction

Solar energy plays a leading role in the global renewable capacity expansion. In 2020, there was an annual increase of solar capacity of 127 GW (+22%) [1]. This growing expansion is a key component that supports the energy transition towards decarbonization in many countries worldwide. However, due to the varying nature of the solar resource, the optimal planning and deployment of solar power applications requires detailed knowledge of the spatial and temporal variability of solar irradiance.

Long-term variability analyses of solar irradiance provide essential information to decision-makers for a valid selection of suitable locations and optimal system design, as well as for the assessment of risk and the financial viability of solar energy projects [2–7]. Historical datasets covering 30 years or longer are needed to statistically characterize the solar irradiance at any location [7]. Unfortunately, such long-term, high-quality solar irradiance measurements are generally scarce and sparsely distributed, which challenges the spatial characterization at a country or regional level, especially over complex terrain. To overcome this limitation, satellite-derived datasets are considered a reliable and practical option that cover a wide spectrum of spatial scales and are available for long time periods [6].

Previous studies have explored the spatio-temporal variability at a regional scale using gridded satellite datasets [8–14].[1] Some authors used classical climatological classification to condense the

* Corresponding author.
E-mail address: mariela.tapia@uni-bremen.de (M. Tapia).

[1] An overview of these studies can be found in the online Supplementary Material.

https://doi.org/10.1016/j.renene.2022.03.049
0960-1481/

large number of grid points into smaller groups [8], while other studies applied data-driven methods based on regionalization techniques to identify areas with similar solar radiation variability [9–14]. A common approach used in the latter is the k-means algorithm applied to a dimension-reduced dataset by Principal Component Analysis (PCA). Reducing the data dimensionality lowers the computational complexity; however, it might not consider the inherent spatio-temporal dependencies of the data [15,16].

Spatial functional data analysis (sFDA) represents an alternative for analyzing high-dimensional gridded satellite data, which considers the spatio-temporal dependencies of each grid point by means of spatial autocorrelation and complete time functions [16,17]. Although a spatial functional regionalization method was previously applied to gridded satellite precipitation data in Ref. [16], no studies have been found that applied spatial functional regionalization methods to gridded satellite GHI data in order to analyze the spatio-temporal variability of solar radiation.

In this paper, we propose a methodology to characterize the spatio-temporal variability and complementarity of long-term gridded satellite GHI data using sFDA. The methodology comprises multiple steps. First, a hierarchical regionalization method for spatial functional data is used to identify homogeneous areas with similar intra-annual variability patterns. Second, the characterization of the interannual GHI variability is performed using the coefficient of variation. Finally, the spatio-temporal complementarity between the resulting areas from the regionalization is analyzed through correlation coefficients.

As a case study, this methodology is applied to a 21-year period of gridded satellite GHI data on Ecuador's mainland and the Galapagos Islands, which is retrieved from the National Solar Radiation Database (NSRDB). This region features complex climatic characteristics and topography [18,19], which may cause significant spatio-temporal GHI variability [3,7,20], thus representing a relevant case for the application of the proposed methodology. Furthermore, Ecuador has an important solar energy potential, since almost 55% of its territory shows solar radiation levels above 4.1 $kWhm^{-2}d^{-1}$ [21]. However, the share of solar technology in the power mix is still minimal. In 2019, photovoltaic generation only accounted for 0.1% of the total electric power production, while the country continued relying on hydropower and fossil fuel thermal power for electricity generation (76.3% and 21.9%, respectively) [22].

Therefore, the relevance of this work is twofold. First, its novelty lies in the use of the sFDA method for the regionalization of long-term gridded satellite GHI data applied to a region characterized by complex climate and terrain. Second, the case study provides the first comprehensive spatio-temporal characterization of GHI in Ecuador that aims at supporting the Ecuadorian energy sector for the optimal planning and deployment of solar power systems in the country.

The paper is structured as follows. Section 2 presents the study area and data used. Section 3 describes the proposed methodology. Section 4 presents the results of the case study in different subsections: (4.1) GHI regionalization, (4.2) interannual spatio-temporal variability, and (4.3) seasonal complementarity. Section 5 discusses the methodological approach and the practical contribution of the findings. Finally, Section 6 presents the conclusions and gives an outlook to future work.

2. Study area and data

Ecuador is located in the northwest of South America between Colombia to the north, Peru to the east and south, and the Pacific Ocean to the west. The Galapagos Islands also belong to the country and are located about 1000 km west of Ecuador's mainland (Fig. 1). From a climatic point of view, the study area is influenced by the displacement of the Intertropical Convergence Zone (ITCZ), the Pacific Ocean sea surface temperature fluctuations, the trade winds, the influence of Hadley and Walker circulation cells, and moisture advection from the Amazon [16]. From a topographic point of view, the Andes cross over the continental area from north to south dividing it into three regions: the Pacific coastal plains, the Andean highlands, and the Amazon lowlands.

For this study, the satellite-derived GHI data of the Physical Solar Model (PSM) version 3 from the NSRDB [6] was used. The NSRDB is produced and disseminated by the National Renewable Energy Laboratory (NREL) and provides solar irradiance and other meteorological data at a high spatio-temporal resolution (4×4 km grid at hourly and half-hourly intervals) covering a large part of the American continent [6]. The comparison between GHI satellite estimations from the NSRDB and ground measurements from 53 stations across Ecuador's mainland was performed in Ref. [21]. According to their results, the hourly mean bias error (MBE) for clear sky conditions was less than 5%, while for cloudy sky conditions the MBE was found to be less than 10%, suggesting a good agreement between measured and satellite-estimated data. Those results were aligned with the findings from other comparisons at locations throughout the continental United States [6,25].

A 21-year period of hourly GHI dataset from 1998 to 2018 corresponding to the study area was downloaded through the Application Programming Interface (API)[2] provided by NREL. Two preprocessing steps were performed before applying the methodology described in section 3. First, the dataset was statistically checked in order to search for possible data anomalies that could be related to satellite artifacts. This analysis revealed that pixels at longitudes 78.02°W and 77.98°W between latitudes 0.13°N and 1.71°S showed anomalies during the period 2008/11/24 to 2017/12/31, which might be attributed to satellite artifacts (See Appendix A). The nearest-neighbor interpolation method was applied to the identified pixels to reduce the anomalies. Second, the spatial resolution of the dataset was increased from the native NSRDB resolution to 3×3 km, applying the first order conservative remapping method. Fig. 2 shows the long-term monthly mean GHI dataset used in this study. The change in resolution was required for the assessment of mutual complementarities between solar and wind resources for the study area, which uses wind resource data at a 3×3 km resolution simulated by the WRF-ARW[3] meso-scale model. The assessment will be reported in a forthcoming publication.

3. Methodology

Fig. 3 provides an overview of the proposed methodology. Each block represents one of the different steps: regionalization, interannual variability and complementarity analyses, which are explained in detail in the following subsections.

3.1. Regionalization

The hierarchical regionalization method for spatial functional data described in Ref. [26] is applied to the GHI dataset in order to identify homogeneous spatial areas with similar intra-annual variability patterns. This method combines hierarchical clustering algorithms for both geographically referenced and functional data in order to classify spatially correlated curves [26]. It requires a

[2] https://developer.nrel.gov/docs/solar/nsrdb/psm3-download/.
[3] https://www.mmm.ucar.edu/weather-research-and-forecasting-model.

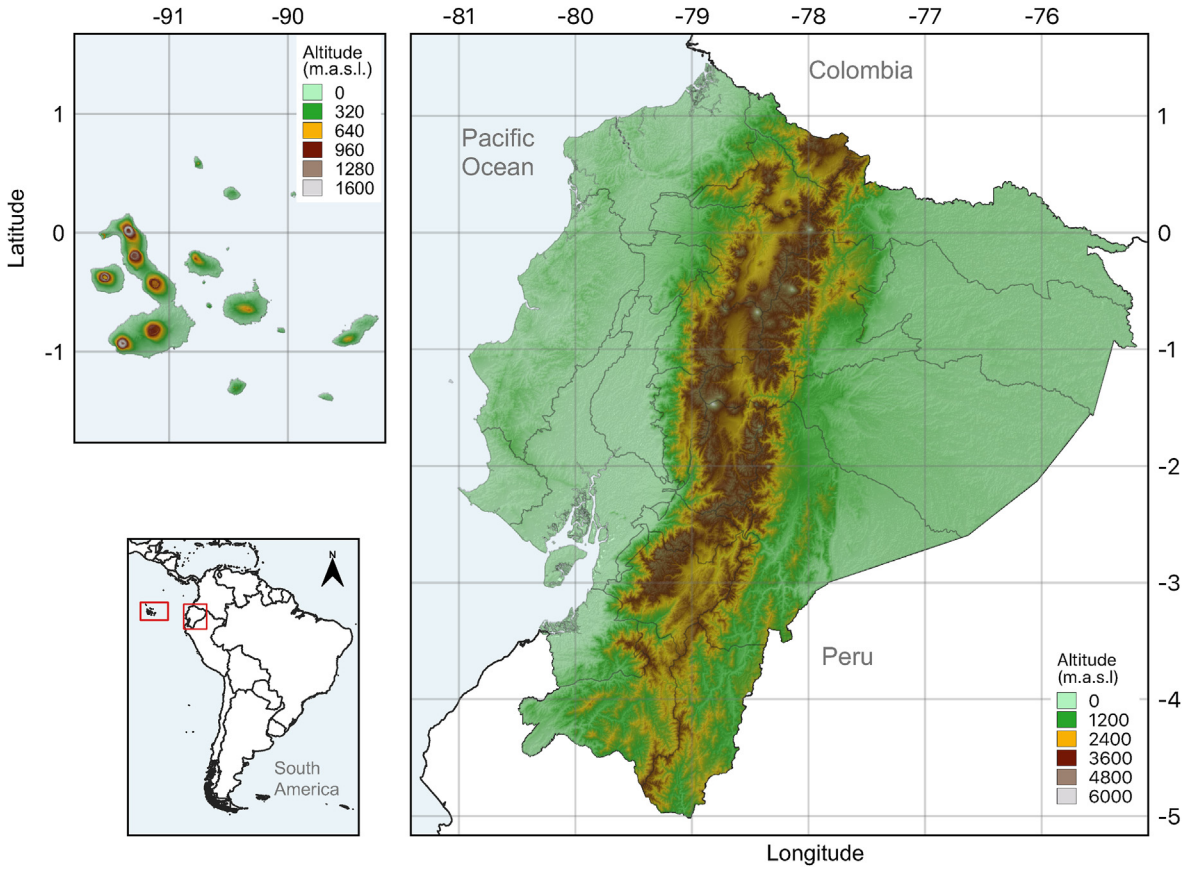

Fig. 1. Digital elevation map of Ecuador's mainland and the Galapagos Islands. The map of the continental area also shows the provincial boundaries. Source: Own representation, data from Refs. [23,24].

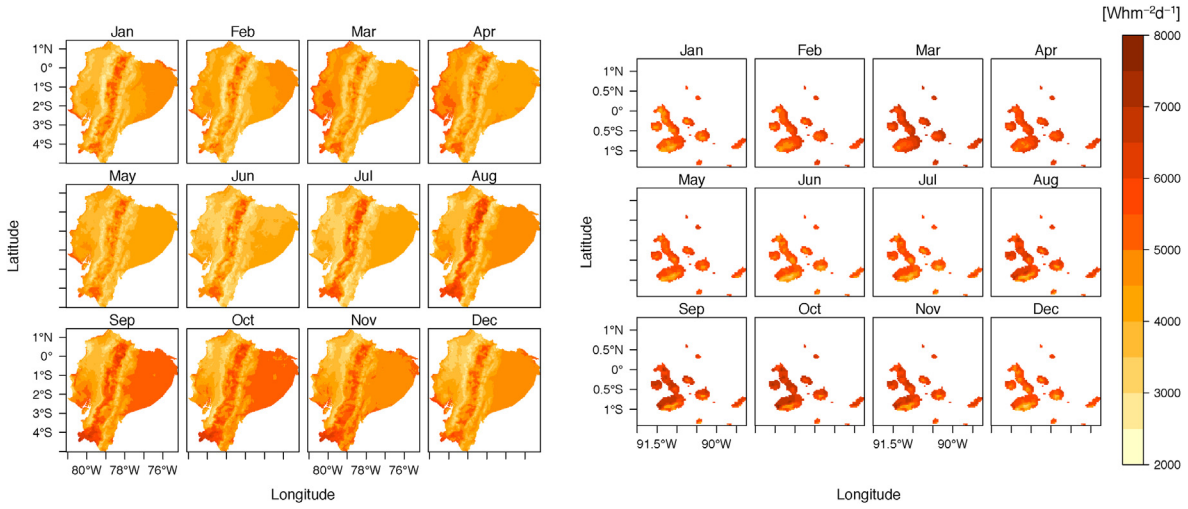

Fig. 2. Long-term monthly mean daily total GHI in [$Whm^{-2}d^{-1}$] from 1998 to 2018 at a 3 × 3 km resolution on Ecuador's mainland (left) and the Galapagos Islands (right). Source: Own representation, data retrieved from the NSRDB.

weighted dissimilarity matrix based on the L_2 norm that accounts for the dissimilarities among curves and the so-called trace-variogram function [27] that accounts for their spatial correlation. A detailed description of the method can be found in Refs. [16,26]. Here we summarize the required steps that are displayed in the regionalization block of Fig. 3.

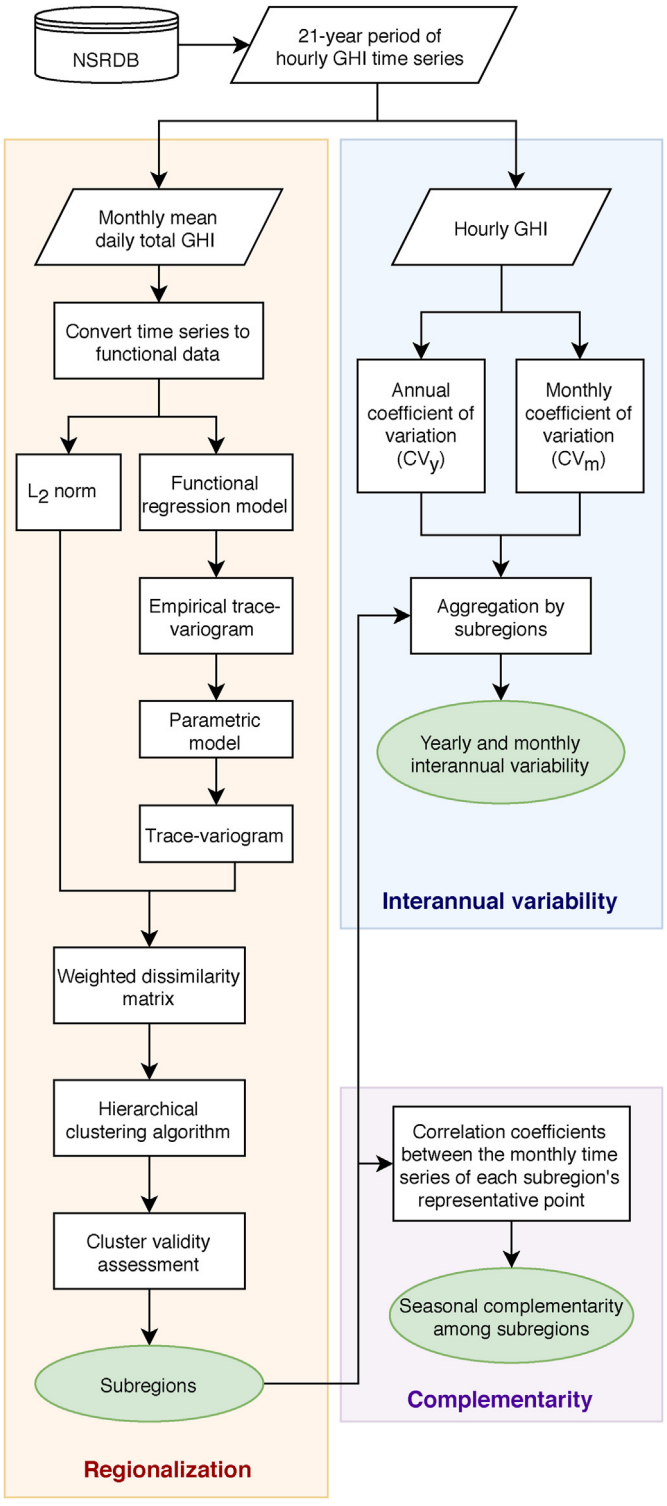

Fig. 3. Flowchart of the proposed methodology to characterize the spatio-temporal variability and complementarity of long-term gridded satellite GHI data.

- ***Convert time series to functional data:*** The discrete time series of monthly mean daily total GHI (12-dimension vectors per grid point) are converted into curves using a Fourier basis with 11 functions. The Fourier basis system is chosen as the smoothing method assuming the periodicity of the GHI time series [28]. The resulting dataset comprises functional curves for all grid points (27,175 in Ecuador's mainland and 874 in the Galapagos Islands).
- ***Weighted dissimilarity matrix:*** The weighted dissimilarity measure is expressed as:

$$d\omega\left(\chi_{s_i}(t),\chi_{s_j}(t)\right)=d\left(\chi_{s_i}(t),\chi_{s_j}(t)\right)\gamma_{s_is_j}(h) \tag{1}$$

where $d(\chi_{s_i}(t),\chi_{s_j}(t))=\sqrt{\int_T(\chi_{s_i}(t)-\chi_{s_j}(t))^2dt}$ is the L_2 norm of the distance between two curves, and $\gamma_{s_is_j}(h)$ is the trace-variogram function calculated for the distance between locations s_i and s_j [16]. The method assumes that the spatial functional process is second-order stationary [26]; however, this assumption no longer holds true when there is a pronounced geographic trend in the dataset [29], which is the case with solar radiation. For this reason, it is necessary to remove the spatial trend before calculating the trace-variogram function [26]. This is done by using a functional regression model [28] with functional response (smoothed GHI curves) and three scalar covariates (longitude, latitude and altitude). Once the regression model is estimated, the functional residuals are calculated to compute the empirical trace-variogram function. Then, a parametric model is fitted to the empirical trace-variogram following classical geostatistical estimation procedures [26]. Finally, the trace-variogram $\gamma_{s_is_j}(h)$ is calculated by subtracting the covariance function $C(h)$ for the distance between each pair of locations from the variance $C(0)$ obtained from the fitted parametric model [16].

- ***Hierarchical clustering algorithm:*** Once the weighted dissimilarity matrix is calculated, the Ward hierarchical agglomerative clustering method is applied. This method forms hierarchical groups of mutually exclusive subsets on the basis of their similarity with respect to specified characteristics [30].
- ***Cluster validity assessment:*** The optimal number of clusters is selected based on both the analysis of the average silhouette width (ASW) [31] and the analysis of the functional boxplots of the resulting clusters. The ASW provides a score based on how similar a curve is to all the curves in its respective cluster compared to curves in other clusters [16], whereas the functional boxplots provide a visual comparison of the similarities/dissimilarities of the curves grouped in each cluster.

3.2. Interannual spatio-temporal variability

The coefficient of variation (CV) is used to analyze the interannual GHI variability during the 21-year period at yearly and monthly timescales. The CV is defined as the ratio of the standard deviation to the mean value [9]. The results are expressed as percentages and represent a measure of the GHI variability over time at the cell's geographical location [3]. For the analysis of the yearly interannual variability, the annual coefficient of variation (CV_y) is calculated as:

$$CV_y(\%)=\frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(GHI_{y_i}-\overline{GHI}_y\right)^2}}{\overline{GHI}_y}\times 100 \tag{2}$$

where N is the number of years, GHI_{y_i} is the annual mean daily total

Table 1
Interpretation of negative correlation coefficient values for the complementarity analysis (Based on [32]).

Correlation coefficient	Complementarity level
$-0.3 < r < 0.0$	Weak
$-0.6 < r \leq -0.3$	Moderate
$-0.9 < r \leq -0.6$	Strong
$-1.0 \leq r \leq -0.9$	Very strong

GHI of each individual *i* year, and $\overline{GHI}_y$ is the mean of the mean daily total GHI during the 21-year period [8].

Similarly, for the analysis of the monthly interannual variability, the monthly coefficient of variation (CV_m) is calculated based on monthly bins of data (e.g., Januaries, Februaries, etc.) as:

$$CV_m(\%) = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(GHI_{m_i} - \overline{GHI}_m)^2}}{\overline{GHI}_m} \times 100 \quad (3)$$

where N is the number of years, GHI_{m_i} is the monthly mean daily total GHI of each individual *i* year, and $\overline{GHI}_m$ is the 21-year monthly mean daily total GHI for the corresponding month [3,8].

Both CV_y and CV_m are aggregated by the resulting clusters from the regionalization to facilitate the intercomparison among the spatial areas.

3.3. Complementarity

The Pearson product-moment correlation coefficient (r) is used to assess the spatio-temporal complementarity among the different clusters and is calculated as:

$$r_{ij} = \frac{Cov(i,j)}{\sigma_i\,\sigma_j} \quad (4)$$

where $Cov(i,j)$ is the covariance between the monthly time series corresponding to the representative points of the clusters i and j, and σ_i is the standard deviation of the monthly time series of the representative point of cluster i. The correlation coefficient can range from −1 to 1. Complementarity is associated to the negative values, whereas similarity is related to the positive values. The level of complementarity is evaluated according to the interpretation given in Table 1 [32].

3.4. Implementation

The methodology for this study was implemented in R [33], using the packages *fda* [34] and *geofd* [35], among others.

4. Results

4.1. Regionalization results

This section presents the regionalization of GHI in Ecuador's mainland and the Galapagos Islands, described and analyzed in relation to the climate characteristics of the resulting spatial areas.

4.1.1. Regionalization of GHI in Ecuador's mainland

Twenty-two clusters are selected as the optimal partition for Ecuador's mainland according to the cluster validity assessment explained in Appendix B. These clusters (hereafter called subregions) represent spatially homogeneous areas featured with similar intra-annual GHI variability. From the functional boxplots of the subregions (Appendix C), four main seasonal patterns are identified; consequently, the subregions are grouped into four spatial areas (hereafter called regions) according to the corresponding seasonal pattern. The spatial distribution of the 4 regions and their respective subregions is shown in Fig. 4, where it can be seen that the regions are spatially contiguous and compact areas, distributed longitudinally through Ecuador's mainland. From east to west, region A is located in the Ecuadorian Amazon, regions EH and WH are located in the eastern and the western side of the Andean highlands, respectively, and region C is located in the coastal area. The Ecuadorian Andes consist of two parallel mountain ranges (the Eastern and Western Cordilleras) separated by the tectonic Inter-Andean Depression [36]. They are the major climate divide that separates the humid lowland forests of the Amazon basin from the coastal areas [37]. These topographical and climate characteristics are reflected in the spatial distribution of the 4 regions, where the borders of regions C and A closely match the contour of the highlands. Furthermore, the spatial location of regions EH and WH resembles the distribution of the Andean highlands.

Fig. 5 shows the intra-annual GHI variability of each region. Fig. 5a shows the dispersion of the functional curves per region, whereas Fig. 5b compares their monthly and annual means. High GHI values are found in region WH, which shows a unimodal pattern with a peak in September. Region A displays a bimodal pattern with the annual maxima occurring in September–October and a peak of smaller magnitude in April. Region EH shows a similar seasonal pattern to region A but of smaller GHI magnitude, with the annual maxima occurring in October. In contrast to the previous regions, region C shows a bimodal pattern with a high peak in March and another of smaller magnitude in September. The annual minima of GHI in all regions (except for WH) occur in June–July, following the mid-year low sun elevation that corresponds to the austral winter. From Fig. 5b (left) it is evident that seasonal complementarities exists between region C and the other regions.

Figs. 6 and 7 show the spatial distribution and intra-annual GHI variability of the 22 subregions grouped by their corresponding region. Here it is noticeable that the subregions in the corresponding region share similar seasonal patterns, but they differ in the magnitude of GHI. This can also be seen in Table 2, which summarizes the annual mean daily total GHI per subregion.

Similar to the spatial variability, the intra-annual GHI variability may be explained by the climate characteristics – mainly cloud and rainfall dynamics – occurring in the different subregions that can affect GHI in complex ways. In Ecuador's mainland, the climatic regimes closely depend on the characteristics of the air masses, which in turn are influenced or produced by three main factors: (i) the seasonal displacement of the ITCZ towards the north or south that determines the input of air masses with different temperature and humidity conditions; (ii) the mountain ranges that play a fundamental role in the formation, displacement and isolation of local or regional air masses; and (iii) the Pacific Ocean currents and

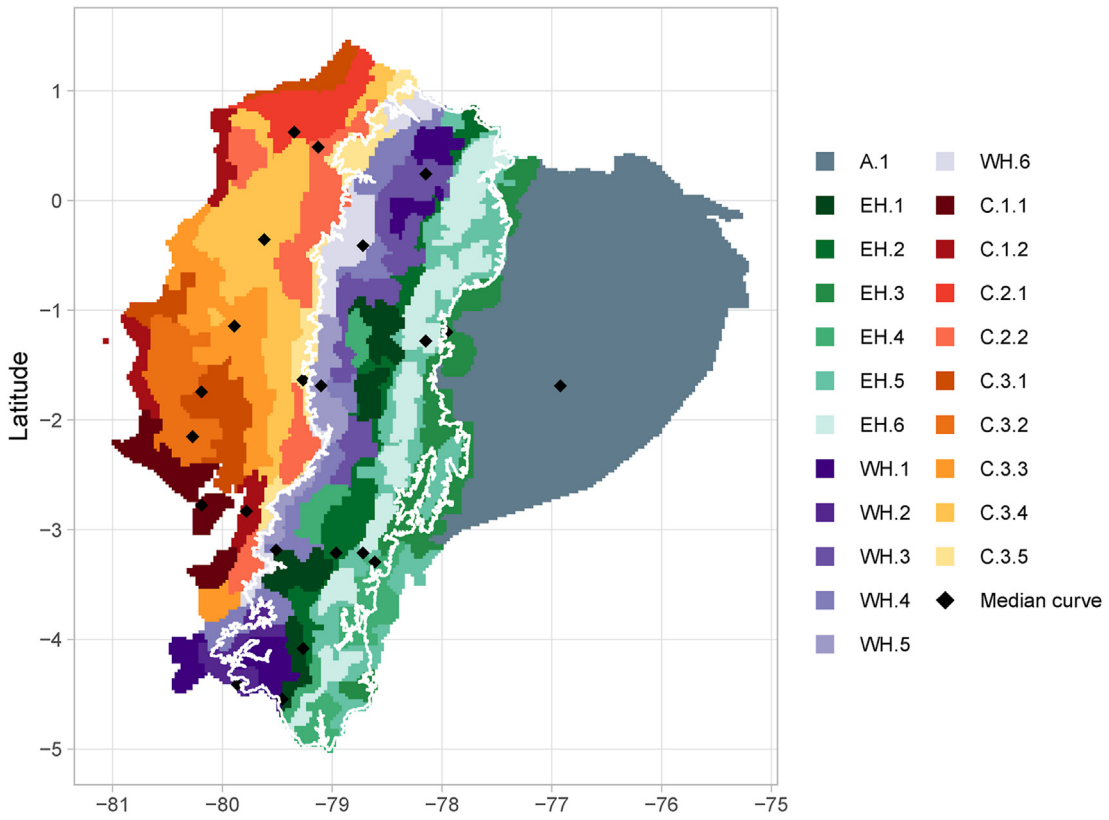

Fig. 4. Spatial distribution of the 22 subregions in Ecuador's mainland after applying the sFDA regionalization method. The subregions are grouped and named according to their corresponding region: Amazon (A), Eastern highlands (EH), Western highlands (WH), and Coast (C). The white line depicts the altitudinal contour at 1000 meters above sea level (m.a.s.l.). The black points represent the location of the median curve of each subregion.

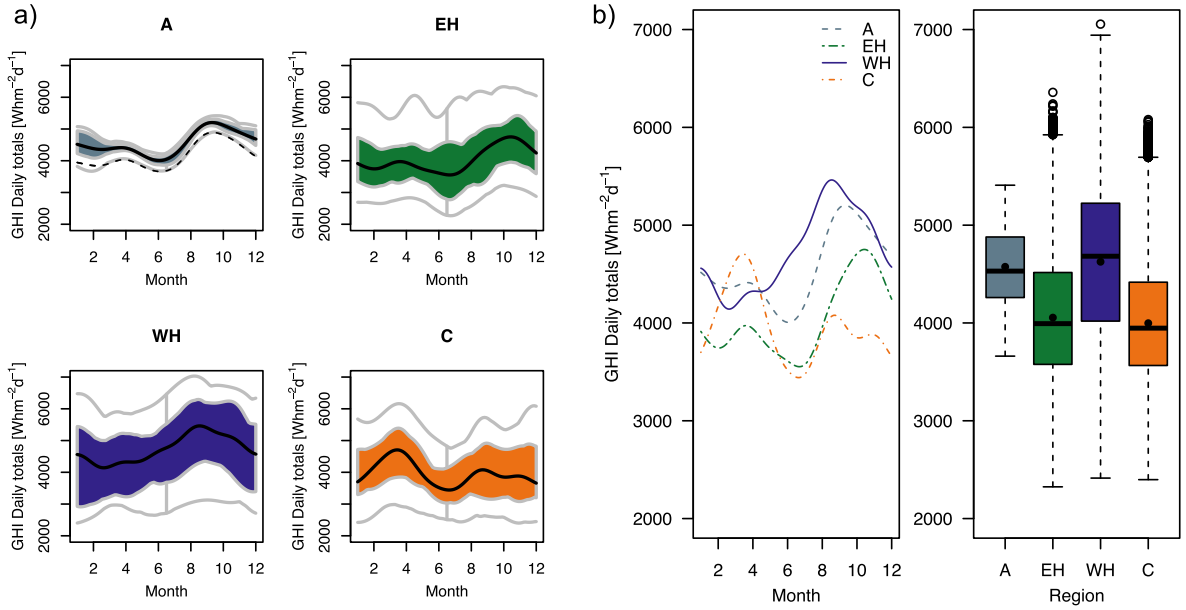

Fig. 5. Intra-annual variability of the four regions: Amazon (A), Eastern highlands (EH), Western highlands (WH), and Coast (C). The functional boxplots on a) show the dispersion of the functional curves per region, where the colored shapes represent the interquartile range, the gray external lines depict the minimum and maximum curves, and the black lines are the median curves interpreted as the main seasonal patterns in each region. The black dotted line in region A represents the outlier curve in the region. The left side of b) shows the comparison of the four main seasonal patterns. The right side of b) shows the dispersion of the annual mean daily total GHI per region, where the black points represent the values of the median curve for each region, and the black circles are the outliers.

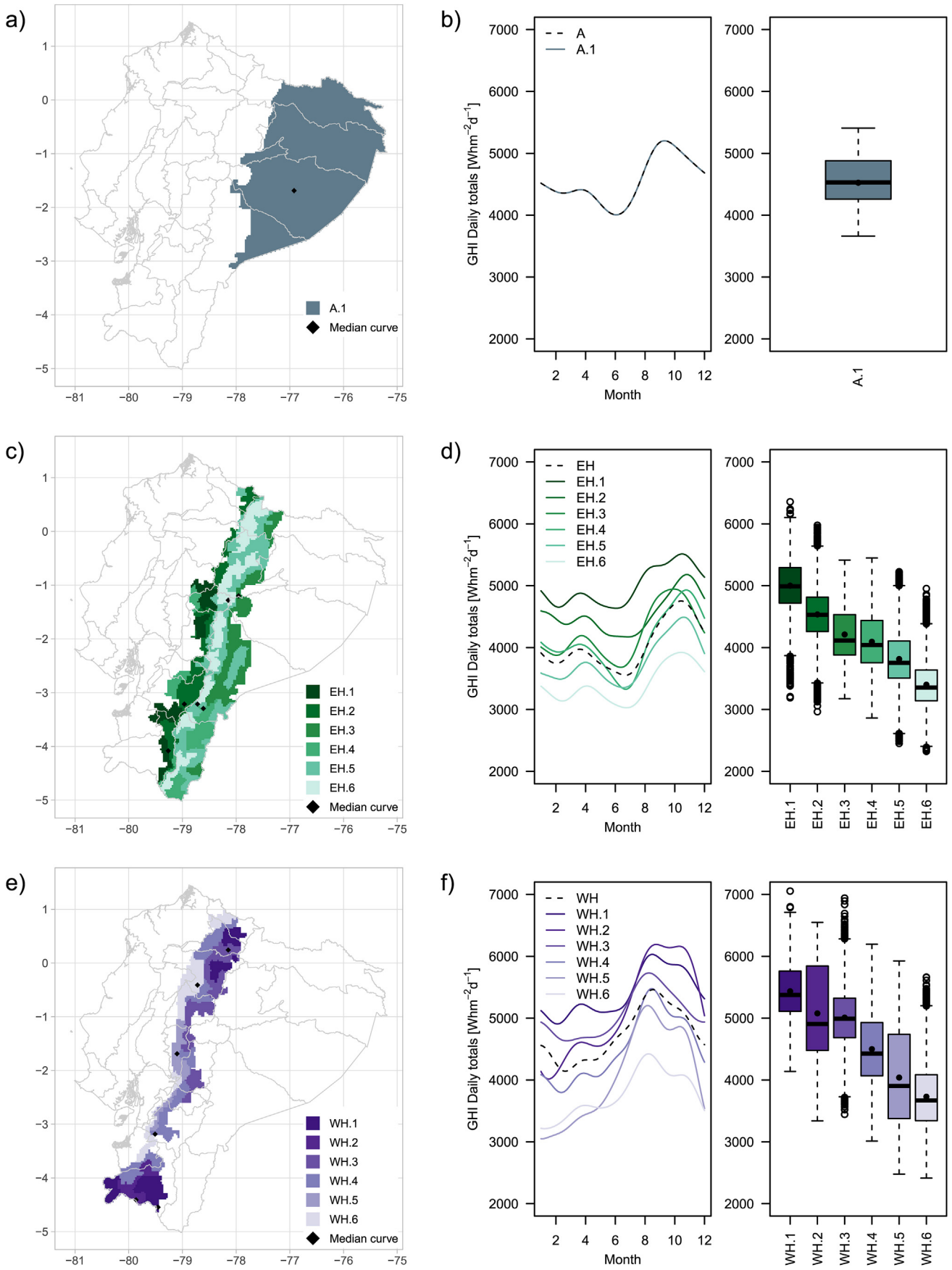

Fig. 6. Spatial and temporal variability of the subregions in regions: Amazon (A), Eastern highlands (EH), and Western highlands (WH). The maps on a), c) and e) show the spatial distribution of the subregions within each region (gray polygons depict the province boundaries). The left sides of b), d) and f) show the comparison of the main GHI seasonal patterns for the corresponding region (dotted black lines) and the median curves obtained from the functional boxplot of each subregion (Fig. C.1). The right sides of b), d) and f) show the dispersion of the annual mean daily total GHI for each subregion, where the black points represent the values of the median curve of each subregion.

the maritime air masses that influence the rainfall patterns in the Coast and the Andean region [18,38]. The influence of these factors on the spatio-temporal GHI variability will become more evident when describing the subregions of each region in the following subsections.

4.1.1.1. Amazon. One subregion (A1) is identified over the Amazon lowlands (Fig. 6a). Interestingly, this homogeneous subregion covers an extended geographical area and it is characterized by high annual mean values (4.6 $kWhm^{-2}d^{-1}$) and low intra-annual variability (Fig. 6b). This may be attributed to the high specific humidity and convective activity throughout the year in the area [39], which results in low total annual sunshine hours [18].

4.1.1.2. Eastern Andean Highlands. Six subregions (EH.1 – EH.6) are identified along the eastern Andean highlands and partly over the Inter-Andean valleys (Fig. 6c). The highest annual mean values are found in EH.1, followed by EH.2 and EH.4 (5.0, 4.5, 4.1 $kWhm^{-2}d^{-1}$, respectively) (Fig. 6d), which are located along the Inter-Andean valleys. In contrast, the lowest annual values are found in EH.6, followed by EH.5 (3.4, 3.8 $kWhm^{-2}d^{-1}$, respectively), located along the eastern Andean flanks.

The low GHI values and low intra-annual variability found in EH.5 and EH.6 may be explained by the influence of a high cloudiness and precipitation band called Andes-Occurring System (AOS) [40]. This area is dominated by moisture-bearing easterlies that originate over the tropical Atlantic and Amazon basin [41,42]. Most of the moisture is precipitated in the form of long duration drizzle due to orographic lifting [37].

The higher values of GHI seen in EH.1 and EH.2 may be attributed to lower rainfall amounts occurring in the Inter-Andean valleys because air masses transported from the Amazon and the Pacific lose much of their humidity on the eastern and western flanks of the Andes [41]. EH.3 is partly located along the outer eastern Andean flanks, and therefore shows a similar seasonal pattern to EH.5 and EH.6, but of higher GHI magnitude, since EH.3 covers part of the Amazon lowlands.

All the EH subregions depict annual maxima in October–November. This is aligned to ground measurements from stations located in southern Ecuador that found November as the month with the greatest clear sky probability [43]. The annual minima occur in June–August, which coincides with the lowest sun elevation of the year and more persistent overcast skies brought in by the strong easterlies [43].

4.1.1.3. Western Andean Highlands. Six subregions (WH.1 – WH.6) are identified along the western Andean highlands and the Inter-Andean valleys (Fig. 6e). WH.1 shows the highest annual mean values (5.4 $kWhm^{-2}d^{-1}$) and low intra-annual variability. WH.2 shows high annual mean values (5.1 $kWhm^{-2}d^{-1}$) as well; however, the intra-annual variability is higher compared to WH.1 (Fig. 6f). WH.3 is located at high elevations of the Western and Eastern Cordillera and shows a similar seasonal pattern to WH.1, but of lower GHI magnitude (Fig. 6f).

WH.4, WH.5 and WH.6 are spatially distributed along the western Andean slopes towards the coastal area (Fig. 6e) and show lower annual mean values (4.5, 4.0 and 3.7 $kWhm^{-2}d^{-1}$, respectively) (Fig. 6f). The reduced GHI magnitude in this area may be attributed to the influence of the ITCZ over the eastern Pacific, which is responsible for continuous moisture in the form of rainfall, clouds and fog due to orographic uplift [42]. These aspects are especially reflected in the lower intra-annual GHI variability of WH.6 (Fig. 6f).

As mentioned above, the annual minima in the WH region do not match the lowest sun elevation of the year and vary from December to March among the subregions. This singularity was analyzed in Ref. [43] using measured data from stations located in southern Ecuador. The author concluded that the barrier effect produced by the eastern mountains benefits the leeward flanks, because in June–August the share of the water vapor and the clouds retained by the mountain range increases. Consequently, extended times of direct insolation compensate for the lowest annual solar radiation [43].

4.1.1.4. Coast. Nine subregions with three slightly different seasonal patterns are identified in the coastal area and grouped accordingly as C.1.1, C.1.2, C.2.1, C.2.2, C.3.1 – C.3.5 (Fig. 7). In general, all the subregions show annual maxima in March–April and another peak of reduced intensity in September–October. Fig. 7b shows the first type of seasonal pattern found in C.1.1 and C.1.2. The highest annual mean value is found in C.1.1 (4.8 $kWhm^{-2}d^{-1}$) located in the lowlands near the coastline. This area is particularly dry with low rainfall rates (in the form of drizzle) due to the influence of the cold Humboldt current [18]. C.1.2 has a similar seasonal pattern to C.1.1, but of lower GHI magnitude due to the spatial dispersion and the humidity level of the three partitions that comprised this subregion (ranging from arid and dry to super-humid and sub-humid [44]).

Fig. 7c shows the second type of seasonal pattern that is found in C.2.1 and C.2.2, characterized by low intra-annual variability and low GHI values. This may be attributed to the influence of the ITCZ [45] and the location of the subregions along the lower slopes of the western Andes, where orographic rainfall occurs.

Fig. 7d shows the third seasonal pattern that is found in C.3.1 – C.3.5, where a gradual decrease of GHI magnitude can be seen. From Fig. 7a, it is noticeable that this gradient occurred from west to east and is related to the geographical location and altitude of the subregions, since rainfall increases between the low-altitude coastal Cordillera and the Andean foothills [46].

4.1.2. Regionalization of GHI in the Galapagos Islands

Three subregions are selected as the optimal partition for the Galapagos Islands according to the cluster validity assessment explained in Appendix B. Fig. 8a shows the spatial distribution of the subregions (G.1.1, G.1.2 and G.1.3), which depict a bimodal pattern, but they differ in GHI magnitude (Fig. 8b). G.1.1 shows the highest annual mean value (6.0 $kWhm^{-2}d^{-1}$) with the annual maxima in October. G.1.2 and G.1.3 show lower values (5.5 and 4.8 $kWhm^{-2}d^{-1}$, respectively) with the annual maxima in March. The minima of all subregions occur in June–July.

Similar to Ecuador's mainland, the spatio-temporal variability of GHI in Galapagos may be explained by the climate characteristics in this area, which result from a complex interaction of oceanic currents that surround the islands and the predominant trade winds from the southeast [47]. During January to May (hot season), G.1.1 and G.1.2 show similar patterns and high GHI values. In contrast, from June to December (cool season) a difference in GHI magnitude can be seen (Fig. 8b). On the one hand, G.1.2 is located along the

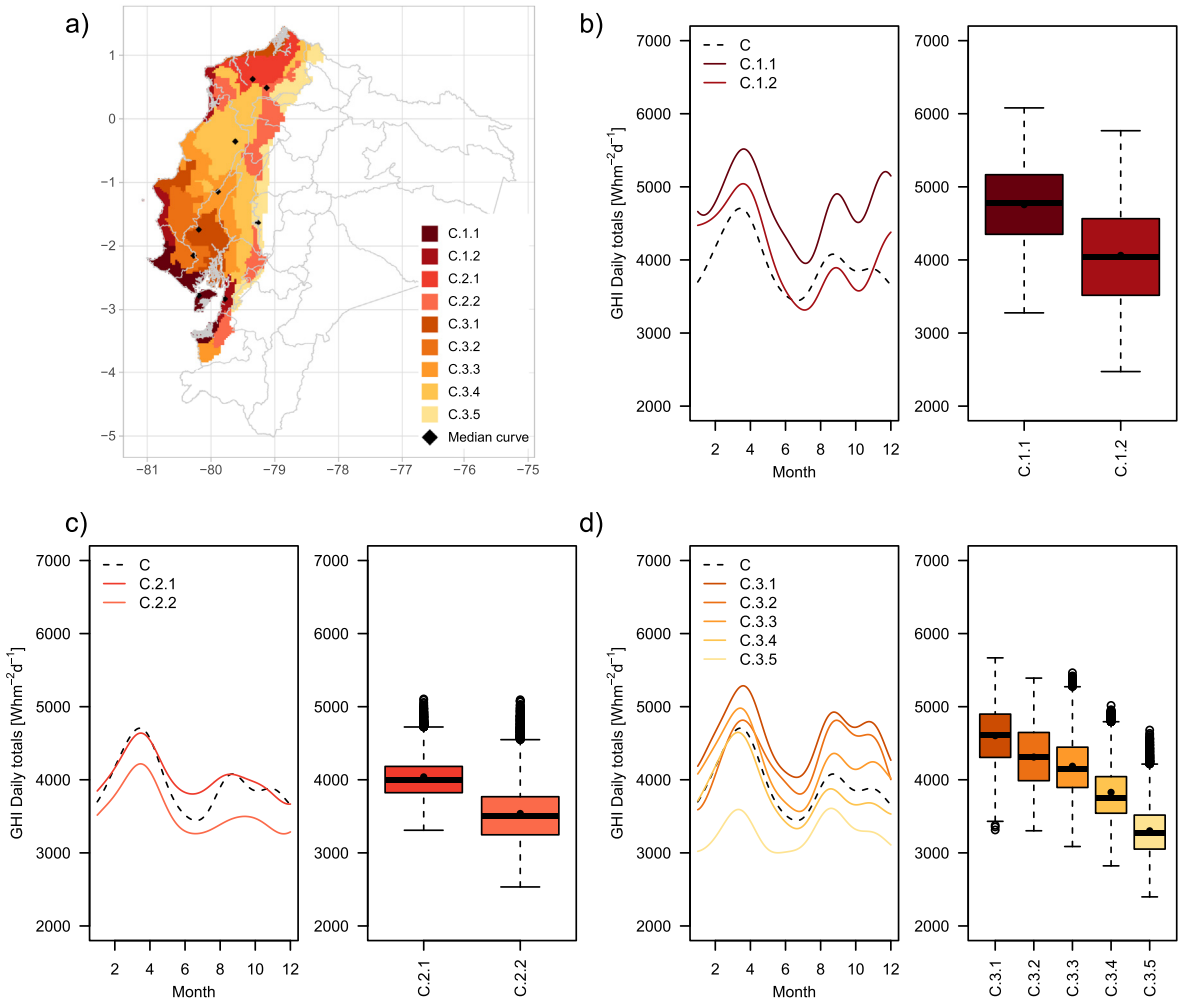

Fig. 7. Spatial and temporal variability of the subregions in the Coast region (C). The map on a) shows the spatial distribution of the subregions (gray polygons depict the province boundaries). The left sides of b) – d) show the comparison of the main GHI seasonal pattern of region C (dotted black lines) and the median curves obtained from the functional boxplot of each subregion (Fig. C.1). The right sides of b) – d) show the dispersion of the annual mean daily total GHI for each subregion, where the black points represent the values of the median curve of each subregion.

windward side of both the islands and the volcanoes, which are more humid during the cool season [47], thus reducing the incoming radiation in these areas. On the other hand, G.1.1 comprises the lowlands and the tops of higher volcanoes, areas that remain dry during the cool season [47], thus showing higher GHI values. G.1.3 shows lower GHI values throughout the year compared to the other subregions. This may be attributed to its location over the highlands and south-facing slopes, since humidity and rainfall increase considerably with altitude due to the influence of the Humboldt Current [18].

4.2. Interannual spatio-temporal variability

Fig. 9 shows the annual coefficient of variation CV_y, where it can be seen that the highest yearly interannual variabilities are along the eastern and western flanks of the Andes, the northwest coastline, and over the south-facing slopes of the Galapagos Islands. In contrast, the Amazon lowlands and the Inter-Andean valleys show the lowest variabilities. Table 2 and Fig. 10 show the variability per subregion in detail. The highest mean CV_y values (>6.0%) are found in EH.6, G.1.3, C.3.5, EH.5, and C.2.2, while the lowest mean values (<4.0%) are found in WH.1 and A.1. Fig. 10 also reveals a clear trend of increasing variability among the subregions within their corresponding regions, meaning that the subregions that show higher CV_y values are those with lower GHI magnitude in their respective region.

Figs. 11 and 12 show the monthly coefficient of variation CV_m, where it is notable that the highest monthly interannual variabilities are along the eastern flanks of the Andes, the western coastline, and over the south-facing slopes of the Galapagos Islands. Fig. 13 takes a more detailed look at these findings, aggregated per subregions. The highest mean CV_m values are in the EH region (especially in EH.6) during two intervals: December to February and July to September. On the coast, C.1.2 shows high variabilities from May to December. Further, in Galapagos, G.1.3 shows the highest variabilities from December to January. In contrast, the

Table 2
Annual mean daily total GHI [kWhm^{-2}d^{-1}], yearly coefficient of variation (CV_y) [%] and summary statistics of elevation [m.a.s.l.] in each subregion. Note that the subregions are named and sorted in decreasing order according to the GHI value in their respective region.

Subregion	GHI	CV_y	Elevation min	Elevation mean	Elevation max
A.1	4.6	3.8	184.2	312.3	1109.2
EH.1	5.0	4.0	1015.1	2645.5	4190.7
EH.2	4.5	4.9	1636.1	3186.4	4820.1
EH.3	4.2	4.8	346.1	951.7	1884.8
EH.4	4.1	5.4	787.3	2120.0	5253.3
EH.5	3.8	6.2	533.0	2016.7	4799.2
EH.6	3.4	8.0	1118.1	2685.9	4192.9
WH.1	5.4	3.4	146.8	1763.1	4577.3
WH.2	5.1	4.5	330.3	1148.2	2418.2
WH.3	5.1	4.7	1877.8	3414.7	5302.4
WH.4	4.5	5.3	403.6	2546.0	4870.9
WH.5	4.0	6.1	923.2	2243.8	3984.0
WH.6	3.7	6.0	214.2	1790.8	3993.0
C.1.1	4.8	4.4	0.3	31.2	176.5
C.1.2	4.1	6.0	0.0	101.5	523.9
C.2.1	4.0	4.9	17.1	140.0	488.3
C.2.2	3.6	6.2	4.9	319.2	1313.5
C.3.1	4.6	4.0	0.0	52.0	400.4
C.3.2	4.3	4.2	24.5	246.3	706.0
C.3.3	4.2	4.5	1.9	98.7	833.8
C.3.4	3.8	5.5	6.0	176.9	837.4
C.3.5	3.3	7.0	8.5	956.4	2989.1
G.1.1	6.0	3.9	2.8	259.1	1373.0
G.1.2	5.5	5.7	5.6	212.8	1321.2
G.1.3	4.8	7.3	15.9	292.2	900.9

lowest variabilities with minimum changes during the twelve months are in WH.1, located in the Inter-Andean valleys.

The high variability in EH.6 is associated with the AOS band of clouds and precipitation, as mentioned in section 4.1.1.2. However, from Fig. 9 it can be inferred that this variability may also be attributed to the satellite image artifact (Appendix A), which particularly affects this subregion. Further, the high variability in subregions C.2.2, C.3.5, WH.4, WH.5, WH.6 and G.1.3 may be attributed to the effect of climatic phenomena occurring in the area, such as El Niño–Southern Oscillation (ENSO). This phenomenon, recognized in its cold phase as La Niña and its warm phase as El Niño, influences the rainfall interannual variability on the Pacific coast [48]. El Niño triggers convection that leads to large rainfall events along the coast of Ecuador and Peru [19,49] and the western Andean slopes between 1° and 3°S [41]. Similarly, the surface ocean around the Galapagos Islands warms substantially during El Niño events, producing significantly more rainfall compared to normal years [50]. Studies regarding the effects of ENSO on solar radiation in Ecuador were not found in the literature. Nevertheless, Henao et al. [51] found that solar radiation anomalies in Colombia increase during El Niño and decrease during La Niña, which are related to the absence and excess of convective clouds. Similar dynamics might be expected in Ecuador; however, further investigation is needed to determine the underlying factors of the interannual GHI variability in Ecuador.

4.3. Seasonal complementarity

The correlation matrix for all the pairs of subregions in Ecuador's mainland and the Galapagos Islands is shown in Fig. 14. The negative coefficients highlight that the region C and the other regions are complementary, since their seasonal patterns are opposite (Fig. 5). Different levels of complementarity are identified: (i) strong complementarity between C.1.2 and all WH subregions, as well as between C.2.2 and WH.3, WH.4 and WH.5; (ii) moderate complementarity between subregions in C (except C.3.1 and C.3.2) and E.1 and WH subregions; and (iii) weak complementarity between C and A, and the other subregions in EH and WH. From the positive correlation coefficients, it is also evident that all subregions within each respective region show high similarities, thus highlighting the ability of the sFDA regionalization method to find groups with similar temporal patterns.

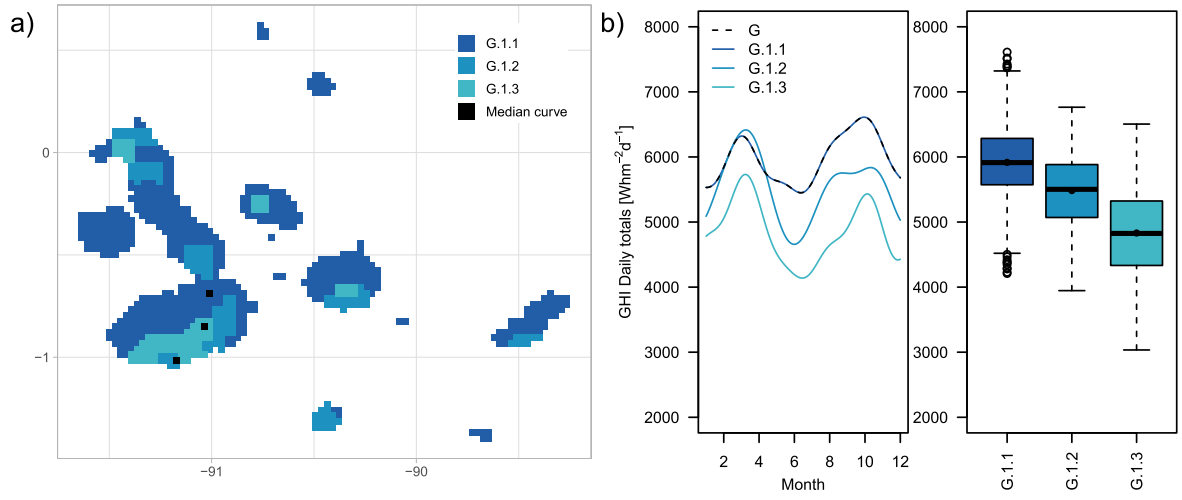

Fig. 8. Spatial and temporal variability of the subregions in the Galapagos Islands (region G) after applying the sFDA regionalization method. The map on a) shows the spatial distribution of the subregions. The left side of b) shows the comparison of the main GHI seasonal pattern of region G (dotted black lines) and the median curves obtained from the functional boxplot of each subregion (Fig. C.1). The right side of b) shows the dispersion of the annual mean daily total GHI for each subregion, where the black points represent the values of the median curve of each subregion.

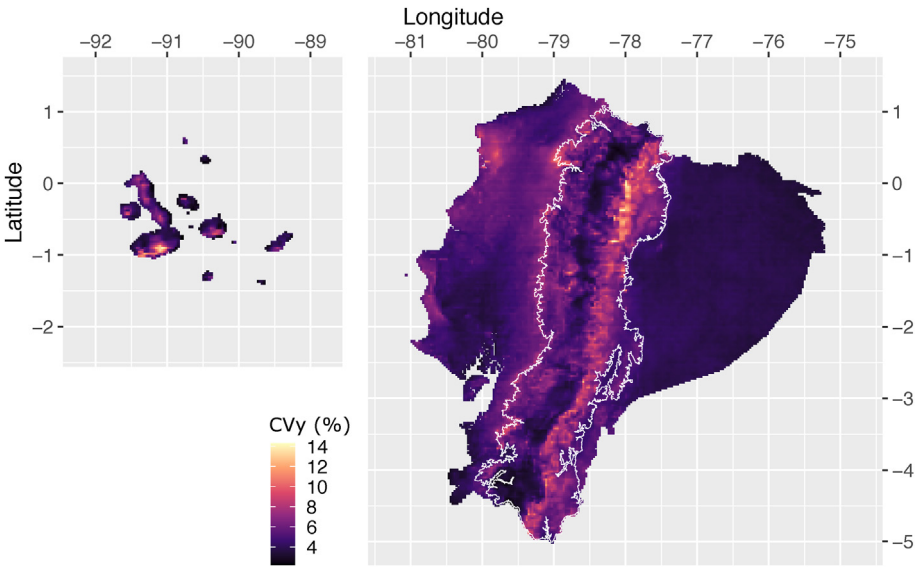

Fig. 9. Annual coefficient of variation CV_y (%) of mean daily total GHI during 1998–2018 over Ecuador's mainland and the Galapagos Islands. The white line over the continental part depicts the altitudinal contour at 1000 m.a.s.l.

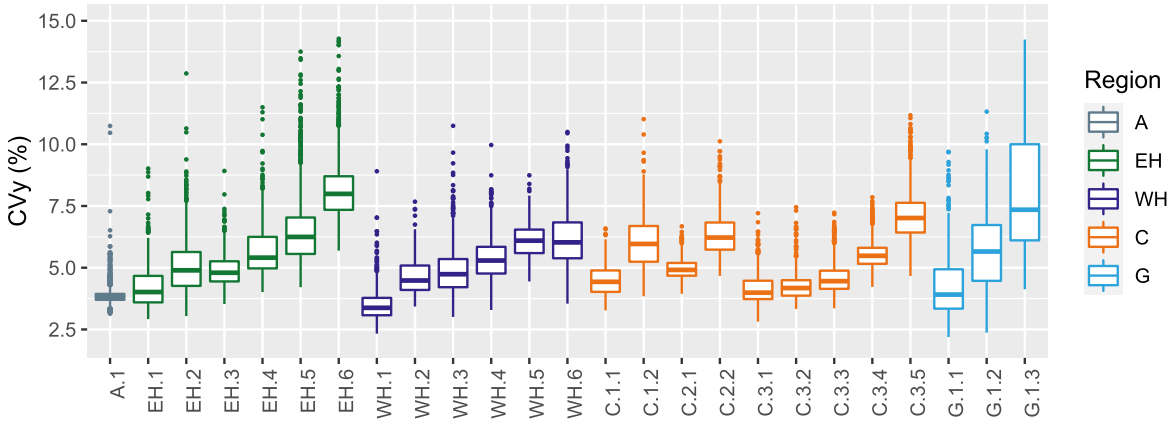

Fig. 10. Annual coefficient of variation CV_y (%) of mean daily total GHI during 1998–2018 per subregion and grouped by region: Amazon (A) Eastern highlands (EH), Western highlands (WH), Coast (C), and Galapagos (G).

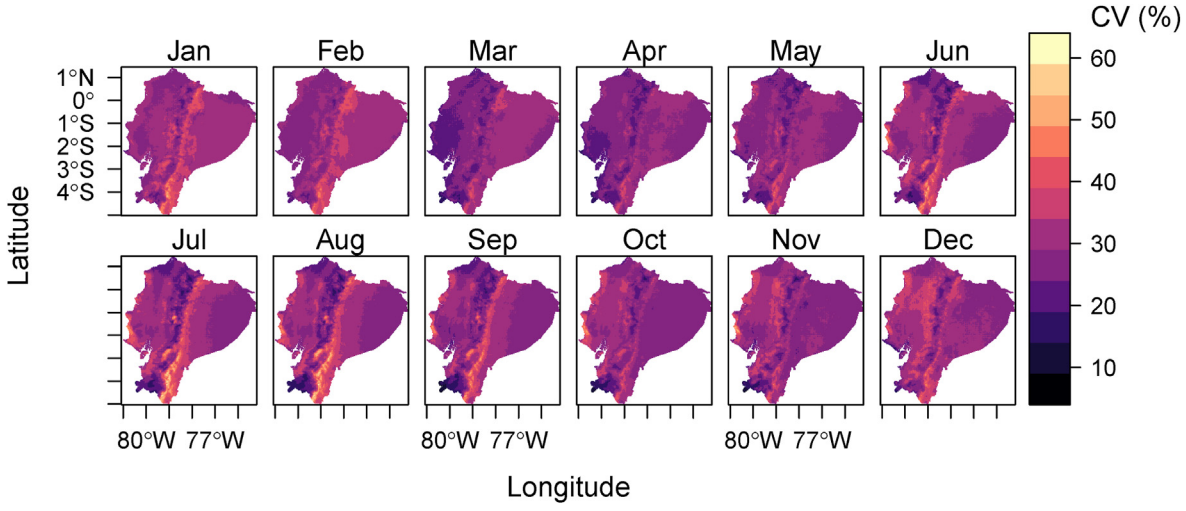

Fig. 11. Monthly coefficient of variation CV_m (%) of mean daily total GHI during 1998–2018 over Ecuador's mainland.

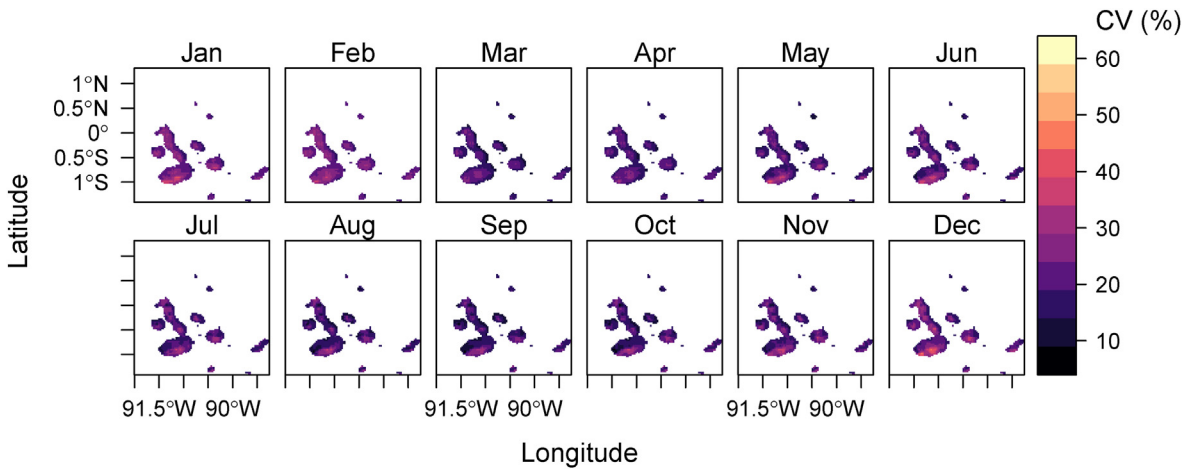

Fig. 12. Monthly coefficient of variation CV_m (%) of mean daily total GHI during 1998–2018 over the Galapagos Islands.

5. Discussion

The results from our case study indicate that the sFDA regionalization method applied to gridded satellite GHI data effectively identifies spatially homogeneous areas featured by similar GHI intra-annual variability patterns. The variation of the magnitude of GHI among areas that share similar patterns can also be identified. This approach differs from previous studies dealing with the analysis of the spatio-temporal variability of solar resource in a number of respects. Firstly, the sFDA regionalization is a data-driven method that allows for the classification of the dataset considering its inherent spatio-temporal characteristics. This objective classification differs from classical climatological classifications based on temperature and precipitation. Habte et al. [8] used the Köppen-Geiger (KG) climate classification [52] to analyze the variability of the solar resource over America. Other studies used also the KG classification in various solar applications at a global scale [53,54]. Such an approach may be appropriate for the analysis over large geographical areas; however, misclassifications could be expected due to the global focus and coarser resolution of the KG [55], or because misclassifications in mountainous regions in South America could not be corrected due to lack of data [56]. Therefore, the KG classification might not be suitable for the analysis over smaller areas, since solar irradiance can vary significantly over relatively small distances due to microclimate effects of topography and the general behavior of regional weather patterns [3,7,20].

Secondly, unlike previous GHI regionalization approaches that applied clustering algorithms after reducing the dimensionality of the satellite-derived datasets [9–14], the sFDA method classifies the dataset based on the dissimilarities between the curves from each grid point weighted by the dissimilarities throughout space. By considering the existing spatial correlation, the method is able to find groups of curves that are spatially homogeneous, since the weighting function increases the dissimilarities among distant points [57]. In addition, the sFDA method provides tools such as functional boxplots in which the dispersion of the curves within each subregion can be explored in both spatial and temporal dimensions [16].

Thirdly, the regionalization results from our case study indicate that the sFDA method captures the GHI variability due to microclimate effects. This ability is reflected for example in the spatial classification of the subregions along the Andes (section 4.1.1.2). This outcome is contrary to that of Ballari et al. [16], who applied the same regionalization approach for precipitation data but local variabilities along the eastern Andes were not captured. The authors attributed this limitation to the low resolution of the input satellite-derived dataset, which is confirmed by our findings, since we use data at higher resolution (3×3 km grid resolution instead of 27×27 km used in Ref. [16]).

Finally, the variabilities due to the topographic effects are also captured by the sFDA. The interannual variability results of our case study are in line with those of Habte et al. [8], who found high spatial variabilities over western South America. The authors attributed this finding to the complex topography and large ranges of elevation change over small areas. Our results support this association; however, we further identify that the variabilities not only account for the change of elevation, but also the microclimate effects in the different subregions. In addition, the high variabilities in the eastern Andes may also be attributed to the satellite image artifact (Appendix A).

5.1. Practical contributions

This work contributes to the better understanding of the availability and variability of GHI in Ecuador. Previous studies have only provided information regarding the magnitude of the solar resource in Ecuador [21,58,59], but not a comprehensive analysis about its spatio-temporal variability. Ordoñez et al. [21] built an updated solar atlas using the typical meteorological year (TMY) dataset from the NSRDB and highlighted those provinces with good solar potential. However, TMY datasets represent typical rather

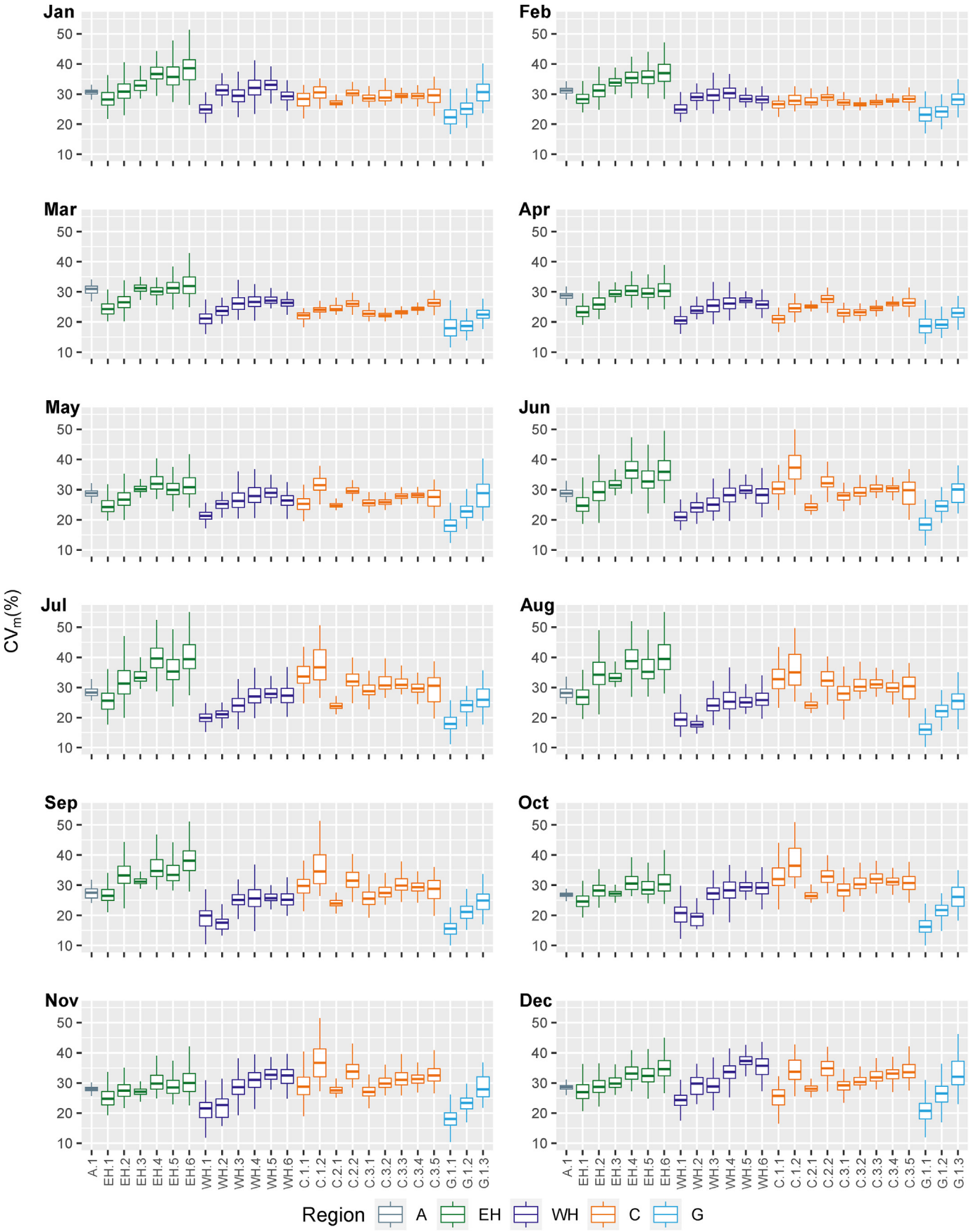

Fig. 13. Boxplots of the monthly coefficient of variation CV_m (%) of mean daily total GHI during 1998–2018 per subregion and grouped by region: Amazon (A) Eastern highlands (EH), Western highlands (WH), Coast (C), and Galapagos (G). The outliers are excluded for clarity.

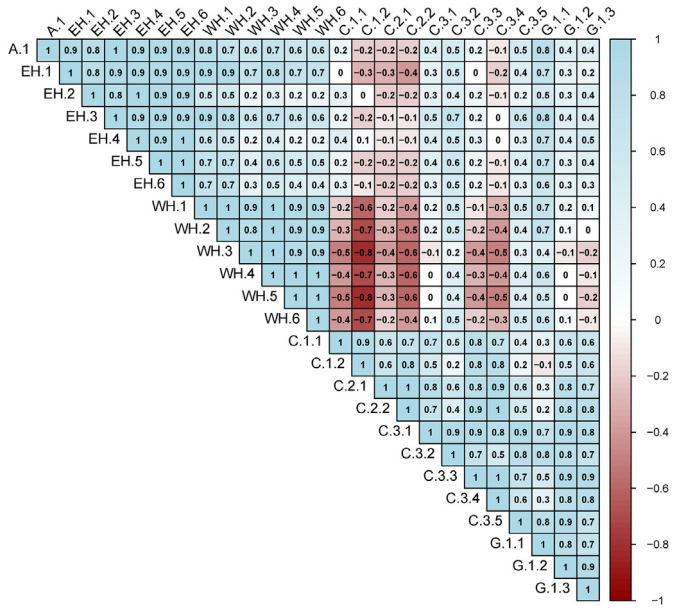

Fig. 14. Correlation matrix of monthly mean daily total GHI time series of the median curves of each subregion. Negative correlation coefficients (r) are highlighted in red. Complementarity levels: strong: $-0.9 < r \leq -0.6$, moderate: $-0.6 < r \leq -0.3$, weak $-0.3 < r < 0.0$.

than extreme conditions, therefore they are not suitable to completely understand the resource variability [7]. In contrast, our study uses a 21-year period dataset from the NSRDB that enables both the identification of spatial areas with the best potential, and the evaluation of the variability of GHI over time, which is highly relevant for the design and financial feasibility of solar energy projects [3].

Another important finding is that the subregions on the coast and the western Andean highlands show significant seasonal complementarities. This aspect raises the possibility for exploring planning and operation strategies, known as geographical smoothing effect [12,60–62], to smooth out the intra-annual GHI variability and consequently the solar production in these areas. Nevertheless, according to the statistical analysis of the interannual variability, high variabilities from May to December are found in C.1.2 and C.2.2. Since these subregions show strong complementarity to the WH region, further investigation is needed for assessing the complementarities during ENSO events, which could have an effect on the solar resource, as previously investigated in Colombia [51,63].

The validation of the NSRDB dataset over the study area was beyond the scope of this study. In that respect, Ordoñez et al. [21] found a good agreement between the satellite estimates and ground-measured solar radiation data from 53 stations in Ecuador's mainland. However, due to the sparse monitoring systems across the country, no measurements from the eastern and southern Ecuadorian Amazon, nor from the Galapagos Islands were used in the mentioned study. Therefore, it is recommended to perform further validations supported by the findings from our study. The geographical location of the median curve of each subregion, which is considered the most representative temporal pattern of the subregion, could be used for the placement of ground-based sensors and thus establish an optimal measurement network. The subregions with low interannual variability would require shorter measurement periods to characterize the solar resource [3].

The regionalization of GHI in our case study is performed using monthly averaged data to understand the intra-annual variabilities and seasonal complementarities. Nevertheless, the proposed methodology can be used at different temporal scales (e.g., daily or hourly), as well as to analyze other renewable resources, or energy production data. We use the Fourier basis system as the smoothing method assuming that the monthly time series are periodic, however, for nonperiodic data (e.g., hourly wind speed data), the spline basis would be recommended (further details can be found in Ref. [64]). A practical limitation for using the sFDA regionalization method might be, however, the high computation cost in the case of a larger number of grid points or data at a high temporal resolution.

6. Conclusions

In this work we present a methodology to characterize the spatio-temporal variability and complementarity of long-term gridded satellite GHI data using spatial functional data analysis. The methodology is implemented over Ecuador's mainland and the Galapagos Islands to demonstrate its applicability over a region characterized by complex climate and terrain. The results indicate that the regionalization method identifies coherent areas with similar intra-annual variability patterns and effectively captures the GHI variations influenced by microclimate and topographic effects. Further, the resulting subregions provide an appropriate basis for the regional analysis of the interannual variability and complementarity. The proposed methodology can be applied to other regions or countries and adapted to analyze the spatio-temporal variability of other renewable resources or energy yield.

The findings from this study provide the first comprehensive spatio-temporal characterization of GHI in Ecuador, which is of particular relevance to support the optimal planning and deployment of solar power systems in the country. The regionalization map and the variability statistics provide explicit information to identify optimal locations for the deployment of irradiance measurement networks and solar power plants. In the continental area, WH.1 is one of the most suitable subregions, since it shows the highest annual mean values (5.4 $kWhm^{-2}d^{-1}$) and the lowest annual coefficient of variation (3.4%). Further areas with high potential ($\geq$4.5 $kWhm^{-2}d^{-1}$) are located in the Inter-Andean valleys, the Amazon lowlands and along the coastline. In Galapagos, all islands show great potential ($\geq$4.8 $kWhm^{-2}d^{-1}$), especially the subregion G.1.1, which has the highest annual mean values (6.0 $kWhm^{-2}d^{-1}$) and the lowest annual coefficient of variation (3.9%).

Moreover, the newly identified seasonal complementary behavior between the coast and the Andean Highlands lays the foundation for planning geographically-dispersed solar power plants, with the goal of smoothing the solar resource variability. Future research should explore the variability of other renewable resources and the complementarity between them. This analysis will support the development of strategies to diversify the power mix in order to compensate for the seasonal variability of hydrological resources and thus reduce dependencies on fossil fuel thermal power in Ecuador. The impact of ENSO on energy resources should also be investigated to support the design of long-term strategies for the Ecuadorian power sector.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

The NSRDB dataset used in this study can be downloaded from

https://nsrdb.nrel.gov.

CRediT authorship contribution statement

Mariela Tapia: Conceptualization, Methodology, Software, Validation, Formal analysis, Visualization, Writing – original draft. **Detlev Heinemann:** Writing – review & editing. **Daniela Ballari:** Software, Writing – review & editing. **Edwin Zondervan:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the High-Performance Computing Team from the University of Oldenburg for their computing facilities. Thanks to Arnim von Gleich and Pablo Thier for their useful comments. The first author dedicates this work in memory of Prof. Dr. Stefan Gößling-Reisemann.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.renene.2022.03.049.

Appendix A

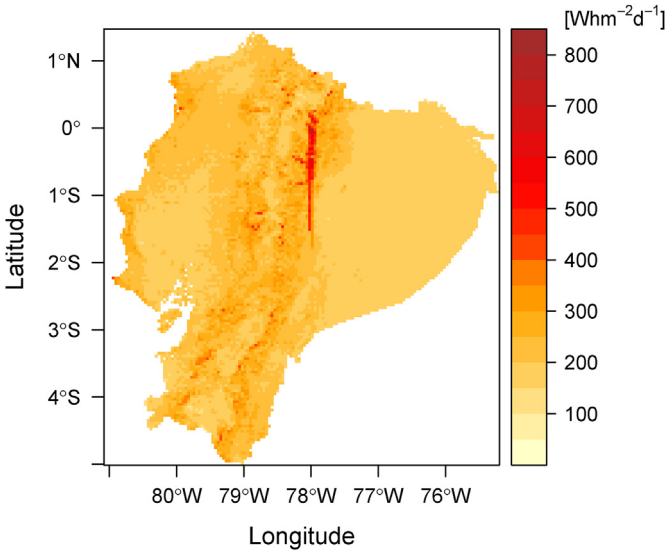

Fig. A.1. Standard deviation of the annual mean daily total GHI during 1998–2018 at 4 × 4 km spatial resolution over Ecuador's mainland. Pixels at longitudes 78.02°W and 77.98°W between latitudes 0.13°N and 1.71°S showed anomalies during the period 2008/11/24 to 2017/12/31. Source: Own representation, data from the NSRDB.

Appendix B

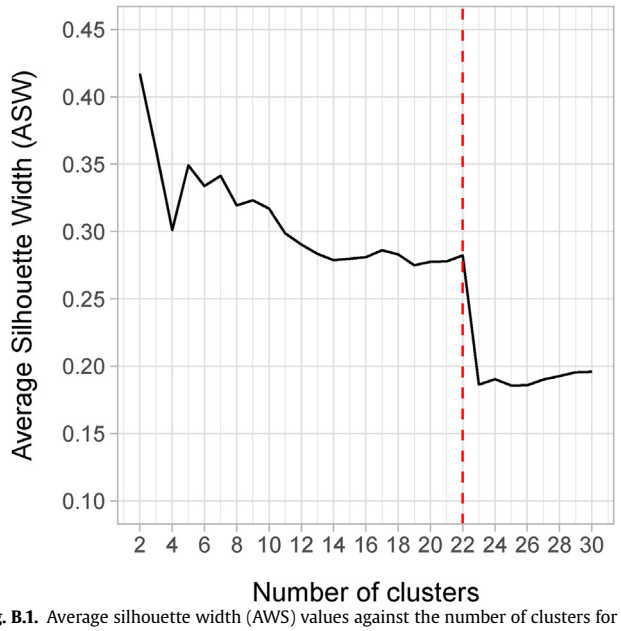

Fig. B.1. Average silhouette width (AWS) values against the number of clusters for the regionalization of GHI in Ecuador's mainland. 22 clusters (vertical dotted line) are selected as optimal partition because the AWS value decreases significantly for higher number of partitions. In addition, as seen in Fig. C.1, the functional boxplots of the 22 subregions display a uniform dispersion of the functional curves with a minimum number of outliers curves, which suggests a coherent spatial classification.

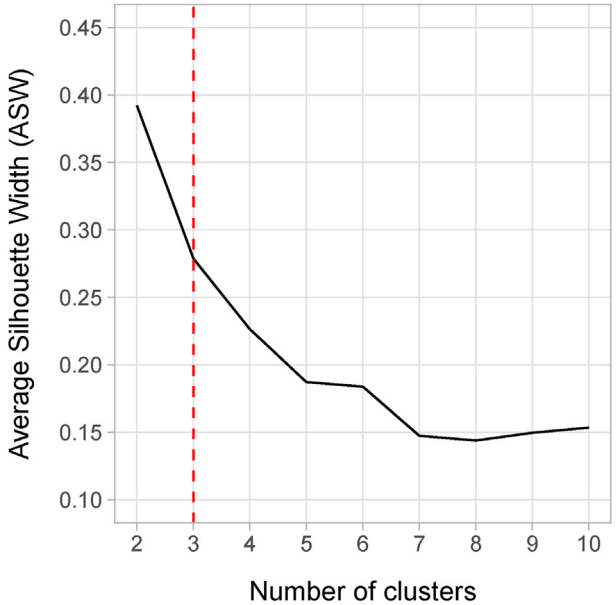

Fig. B.2. Average silhouette width (AWS) values against the number of clusters for the regionalization of GHI in the Galapagos Islands. Three clusters (vertical dotted line) are selected as optimal partition. Although the AWS at 2 is higher, 3 clusters are selected because the spatial distribution of the clusters is aligned to the climatic zones described in Ref. [47]. In addition, as seen in Fig. C.1, the functional boxplots of the 3 subregions display a uniform dispersion of the functional curves without outlier curves, which suggests a coherent spatial classification.

Appendix C

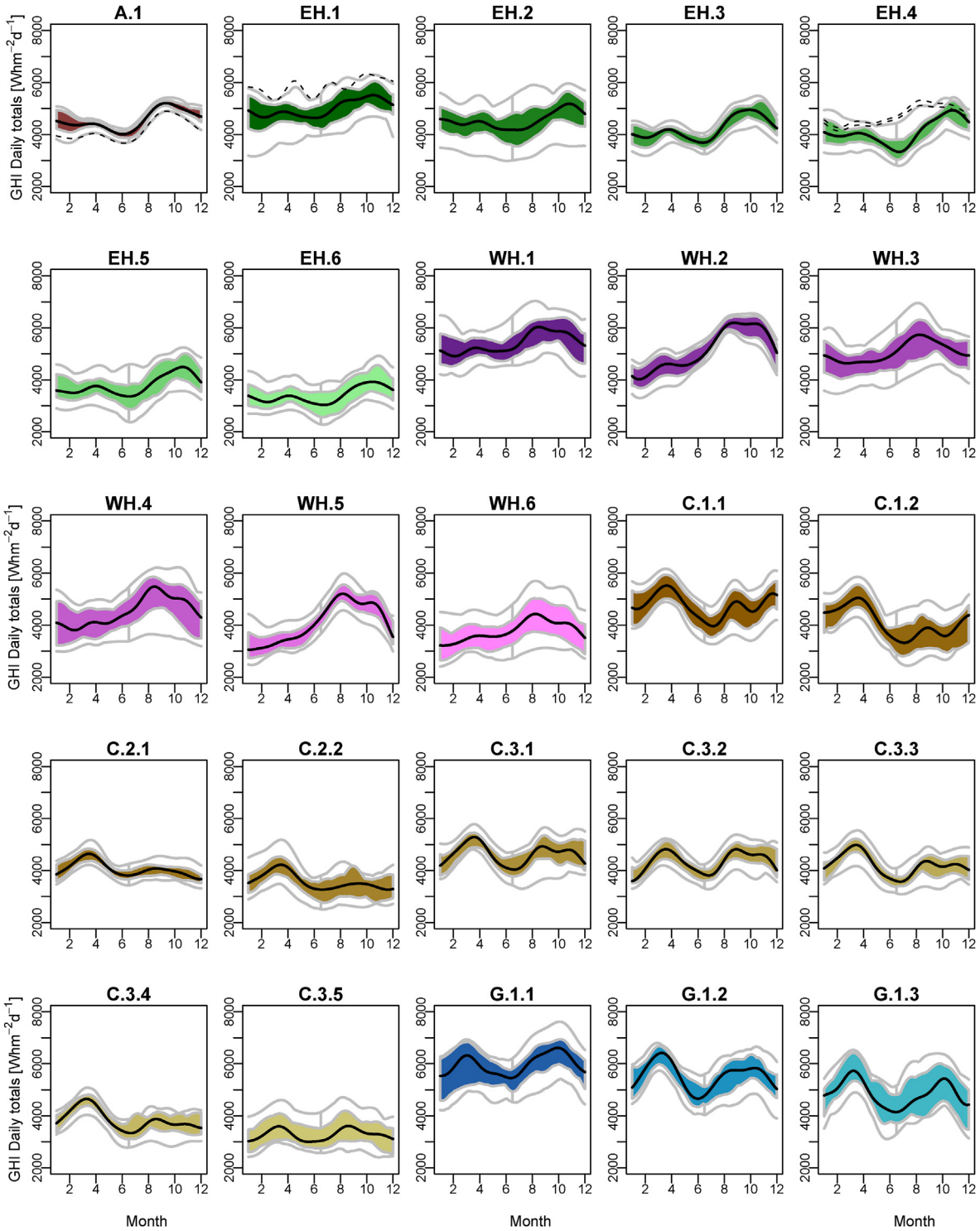

Fig. C.1. Functional boxplots that show the dispersion of the GHI functional curves per subregion in Ecuador's mainland and the Galapagos Islands, named according to their corresponding region: Amazon (A), Eastern highlands (EH), Western highlands (WH), Coast (C), and Galapagos (G). The colored shapes represent the interquartile range, the gray external lines depict the minimum and maximum curves, the black lines are the median curves interpreted as the representative temporal pattern of each subregion. The black dotted lines in A.1, EH.1, EH.4 represent the outlier curves. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

References

[1] IRENA, Renewable energy capacity highlights. https://www.irena.org/publications/2021/March/Renewable-Capacity-Statistics-2021, 2021. (Accessed 26 January 2022).

[2] C.M. Fernández Peruchena, L. Ramírez, M.A. Silva-Pérez, V. Lara, D. Bermejo, M. Gastón, S. Moreno-Tejera, J. Pulgar, J. Liria, S. Macías, R. Gonzalez, A. Bernardos, N. Castillo, B. Bolinaga, R.X. Valenzuela, L.F. Zarzalejo, A statistical characterization of the long-term solar resource: towards risk assessment for solar power projects, Sol. Energy 123 (2016) 29–39, https://doi.org/10.1016/j.solener.2015.10.051.

[3] C.A. Gueymard, S.M. Wilcox, Assessment of spatial and temporal variability in the US solar resource from radiometric measurements and predictions from models using ground-based or satellite data, Sol. Energy 85 (2011) 1068–1084, https://doi.org/10.1016/j.solener.2011.02.030.

[4] S. Lohmann, C. Schillings, B. Mayer, R. Meyer, Long-term variability of solar direct and global radiation derived from ISCCP data and comparison with reanalysis data, Sol. Energy 80 (2006) 1390–1401, https://doi.org/10.1016/j.solener.2006.03.004.

[5] F.J. Rodríguez-Benítez, C. Arbizu-Barrena, F.J. Santos-Alamillos, J. Tovar-Pescador, D. Pozo-Vázquez, Analysis of the intra-day solar resource variability in the Iberian Peninsula, Sol. Energy 171 (2018) 374–387, https://doi.org/10.1016/j.solener.2018.06.060.

[6] M. Sengupta, Y. Xie, A. Lopez, A. Habte, G. Maclaurin, J. Shelby, The National solar radiation data base (NSRDB), Renew. Sustain. Energy Rev. 89 (2018) 51–60, https://doi.org/10.1016/J.RSER.2018.03.003.

[7] F. Vignola, C. Grover, N. Lemon, A. McMahan, Building a bankable solar radiation dataset, Sol. Energy 86 (2012) 2218–2229, https://doi.org/10.1016/j.solener.2012.05.013.

[8] A. Habte, M. Sengupta, C. Gueymard, A. Golnas, Y. Xie, Long-term spatial and temporal solar resource variability over America using the NSRDB version 3 (1998–2017), Renew. Sustain. Energy Rev. 134 (2020), https://doi.org/10.1016/j.rser.2020.110285.

[9] C. Gutiérrez, M.Á. Gaertner, O. Perpiñán, C. Gallardo, E. Sánchez, A multi-step scheme for spatial analysis of solar and photovoltaic production variability and complementarity, Sol. Energy 158 (2017) 100–116, https://doi.org/10.1016/j.solener.2017.09.037.

[10] A. Laguarda, R. Alonso-Suárez, R. Terra, Solar irradiation regionalization in Uruguay: understanding the interannual variability and its relation to El Niño climatic phenomena, Renew. Energy 158 (2020) 444–452, https://doi.org/10.1016/j.renene.2020.05.083.

[11] J.M. Vindel, R.X. Valenzuela, A.A. Navarro, J. Polo, Temporal and spatial variability analysis of the solar radiation in a region affected by the intertropical convergence zone, Meteorol, Appl 27 (2020) 1–10, https://doi.org/10.1002/met.1824.

[12] A. Zagouras, H.T.C. Pedro, C.F.M. Coimbra, Clustering the solar resource for grid management in island mode, Sol. Energy 110 (2014) 507–518, https://doi.org/10.1016/j.solener.2014.10.002.

[13] A. Zagouras, R.H. Inman, C.F.M.M. Coimbra, On the determination of coherent solar microclimates for utility planning and operations, Sol. Energy 102 (2014) 173–188, https://doi.org/10.1016/j.solener.2014.01.021.

[14] A. Zagouras, A. Kazantzidis, E. Nikitidou, A.A. Argiriou, Determination of measuring sites for solar irradiance, based on cluster analysis of satellite-derived cloud estimations, Sol. Energy 97 (2013) 1–11, https://doi.org/10.1016/j.solener.2013.08.005.

[15] J. Jacques, C. Preda, Functional data clustering: a survey, Adv. Data Anal. Classif. 8 (2014) 231–255, https://doi.org/10.1007/s11634-013-0158-y.

[16] D. Ballari, R. Giraldo, L. Campozano, E. Samaniego, Spatial functional data analysis for regionalizing precipitation seasonality and intensity in a sparsely monitored region: unveiling the spatio-temporal dependencies of precipitation in Ecuador, Int. J. Climatol. 38 (2018) 3337–3354, https://doi.org/10.1002/joc.5504.

[17] R. Giraldo, S. Dabo-Niang, S. Martínez, Statistical modeling of spatial big data: an approach from a functional data analysis perspective, Stat. Probab. Lett. 136 (2018) 126–129, https://doi.org/10.1016/j.spl.2018.02.025.

[18] P. Pourrut (Ed.), El Agua en el Ecuador: Clima, precipitaciones, escorrentía, Corporación Editora Nacional, Colegio de Geógrafos del Ecuador, 1995. Quito, https://horizon.documentation.ird.fr/exl-doc/pleins_textes/pleins_textes_7/divers2/010014823.pdf.

[19] E. Morán-Tejeda, J. Bazo, J.I. López-Moreno, E. Aguilar, C. Azorín-Molina, A. Sanchez-Lorenzo, R. Martínez, J.J. Nieto, R. Mejía, N. Martín-Hernández, S.M. Vicente-Serrano, Climate trends and variability in Ecuador (1966–2011), Int. J. Climatol. 36 (2016) 3839–3855, https://doi.org/10.1002/joc.4597.

[20] M. Sengupta, A. Habte, S. Kurtz, A. Dobos, S. Wilbert, E. Lorenz, T. Stoffel, D. Renné, C. Gueymard, D. Myers, S. Wilcox, P. Blanc, R. Perez, Best practices handbook for the collection and use of solar resource data for solar energy applications. https://doi.org/10.18777/ieashc-task46-2015-0001, 2015.

[21] F. Ordoñez, D. Vaca-Revelo, J. Lopez-Villada, Assessment of the solar resource in andean regions by comparison between satellite estimation and ground measurements: study case of Ecuador, J. Sustain. Dev. 12 (2019) 62–75, https://doi.org/10.5539/jsd.v12n4p62.

[22] ARCERNNR, Estadística anual y multianual del sector eléctrico ecuatoriano 2019. https://www.controlrecursosyenergia.gob.ec/wp-content/uploads/downloads/2020/11/Estad%C3%ADstica-2019-nuevologo-small.pdf, 2020. (Accessed 26 January 2022).

[23] D. Watkins, 30-Meter SRTM tile downloader. https://dwtkns.com/srtm30m/, 2021. (Accessed 26 January 2022).

[24] Instituto Geográfico Militar, Geoportal IGM. http://www.geoportaligm.gob.ec, 2021. (Accessed 26 January 2022).

[25] A. Habte, M. Sengupta, A. Lopez, Y. Xie, G. MacLaurin, Assessment of the National solar radiation Database (NSRDB 1998-2016), in: 2018 IEEE 7th World Conf. Photovolt. Energy Conversion, WCPEC 2018 - A Jt. Conf. 45th IEEE PVSC, 28th PVSEC 34th EU PVSEC, 2018, pp. 2305–2308, https://doi.org/10.1109/PVSC.2018.8547589.

[26] R. Giraldo, P. Delicado, J. Mateu, Hierarchical clustering of spatially correlated functional data, Stat. Neerl. 66 (2012) 403–421, https://doi.org/10.1111/j.1467-9574.2012.00522.x.

[27] R. Giraldo, P. Delicado, J. Mateu, Ordinary kriging for function-valued spatial data, Environ. Ecol. Stat. 18 (2011) 411–426, https://doi.org/10.1007/s10651-010-0143-y.

[28] J.O. Ramsay, B.W. Silverman, Functional Data Analysis, Second, Springer Science + Business Media, New York, 2005. https://link.springer.com/book/10.1007/b98888.

[29] M.A. Oliver, R. Webster, Basic Steps in Geostatistics: the Variogram and Kriging, Springer International Publishing, Cham, 2015, https://doi.org/10.1007/978-3-319-15865-5.

[30] J.H. Ward, Hierarchical grouping to optimize an objective function, J. Am. Stat. Assoc. 58 (1963) 236–244, https://doi.org/10.2307/2282967.

[31] P.J. Rousseeuw, Silhouettes, A graphical aid to the interpretation and validation of cluster analysis, J. Comput. Appl. Math. 20 (1987) 53–65, https://doi.org/10.1016/0377-0427(87)90125-7.

[32] M.P. Cantão, M.R. Bessa, R. Bettega, D.H.M. Detzel, J.M. Lima, Evaluation of hydro-wind complementarity in the Brazilian territory by means of correlation maps, Renew. Energy 101 (2017) 1215–1225, https://doi.org/10.1016/J.RENENE.2016.10.012.

[33] R Core Team, R: a language and environment for statistical computing. www.R-project.org, 2020.

[34] J.O. Ramsay, S. Graves, G. Hooker, Fda: functional data analysis. https://cran.r-project.org/package=fda, 2020.

[35] R. Giraldo, J. Mateu, P. Delicado, geofd: an R package for function-valued geostatistical prediction, Rev. Colomb. Estadística 35 (2012) 383–405. https://www.redalyc.org/articulo.oa?id=89925367002.

[36] M. Coltorti, C.D. Ollier, Geomorphic and tectonic evolution of the Ecuadorian Andes, Geomorphology 32 (2000) 1–19, https://doi.org/10.1016/S0169-555X(99)00036-7.

[37] R. Rollenbeck, J. Bendix, Rainfall distribution in the Andes of southern Ecuador derived from blending weather radar data and meteorological field observations, Atmos. Res. 99 (2011) 277–289, https://doi.org/10.1016/j.atmosres.2010.10.018.

[38] R. Celleri, P. Willems, W. Buytaert, J. Feyen, Space–time rainfall variability in the Paute basin, Ecuadorian Andes, Hydrol. Process. 21 (2007) 3316–3327, https://doi.org/10.1002/hyp.6575.

[39] A. Laraque, J. Ronchail, G. Cochonneau, R. Pombosa, J.L. Guyot, Heterogeneous distribution of rainfall and discharge regimes in the Ecuadorian Amazon basin, J. Hydrometeorol. 8 (2007) 1364–1381, https://doi.org/10.1175/2007JHM784.1.

[40] J. Bendix, R. Rollenbeck, D. Göttlicher, J. Cermak, Cloud occurrence and cloud properties in Ecuador, Clim. Res. 30 (2006) 133–147, https://doi.org/10.3354/cr030133.

[41] M. Vuille, R.S. Bradley, F. Keimig, Climate variability in the Andes of Ecuador and its relation to tropical pacific and atlantic sea surface temperature anomalies, J. Clim. 13 (2000) 2520–2535, https://doi.org/10.1175/1520-0442(2000)013<2520:CVITAO>2.0.CO;2.

[42] W. Buytaert, R. Celleri, P. Willems, B. De Bièvre, G. Wyseure, Spatial and temporal rainfall variability in mountainous areas: a case study from the south Ecuadorian Andes, J. Hydrol. 329 (2006) 413–421, https://doi.org/10.1016/j.jhydrol.2006.02.031.

[43] P. Emck, A Climatology of South Ecuador - with Special Focus on the Major Andean Ridge as Atlantic-Pacific Climate Divide, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), 2007. https://opus4.kobv.de/opus4-fau/frontdoor/index/index/docId/477.

[44] R. Moya, Climas del Ecuador, CLIMAS DEL ECUADOR 2016.pdf, http://www.serviciometeorologico.gob.ec/gisweb/TIPO_DE_CLIMAS/PDF/, 2006. (Accessed 12 November 2021), 1-14.

[45] M. Ilbay-Yupa, W. Lavado-Casimiro, P. Rau, R. Zubieta, F. Castillón, Updating regionalization of precipitation in Ecuador, Theor. Appl. Climatol. 143 (2021) 1513–1528, https://doi.org/10.1007/s00704-020-03476-x.

[46] B. Erazo, L. Bourrel, F. Frappart, O. Chimborazo, D. Labat, L. Dominguez-Granda, D. Matamoros, R. Mejia, Validation of satellite estimates (tropical rainfall measuring mission, TRMM) for rainfall variability over the pacific slope and coast of Ecuador, Water 10 (2018), https://doi.org/10.3390/w10020213.

[47] M. Trueman, N. D'Ozouville, Characterizing the Galapagos terrestrial climate in the face of global climate change, Galapagos Res 67 (2010) 26–37. https://www.darwinfoundation.org/en/publications/galapagos-research/galapagos-research-67.

[48] K.E. Trenberth, El Niño southern oscillation (ENSO), in: Encycl. Ocean Sci., Elsevier, 2019, pp. 420–432, https://doi.org/10.1016/B978-0-12-409548-9.04082-3.

[49] A. Bendix, J. Bendix, Heavy rainfall episodes in Ecuador during El Niño events and associated regional atmospheric circulation and SST patterns, Adv. Geosci. 6 (2006) 43–49, https://doi.org/10.5194/adgeo-6-43-2006.
[50] J.P. Sachs, S.N. Ladd, Climate and oceanography of the Galapagos in the 21st Century: expected changes and research needs, Galápagos Res. 67 (2002) 50–54. https://www.darwinfoundation.org/en/publications/galapagos-research/galapagos-research-67.
[51] F. Henao, J.P. Viteri, Y. Rodríguez, J. Gómez, I. Dyner, Annual and interannual complementarities of renewable energy sources in Colombia, Renew. Sustain. Energy Rev. 134 (2020), https://doi.org/10.1016/j.rser.2020.110318.
[52] M. Kottek, J. Grieser, C. Beck, B. Rudolf, F. Rubel, World map of the Köppen-Geiger climate classification updated, Meteorol. Z. 15 (2006) 259–263, https://doi.org/10.1127/0941-2948/2006/0130.
[53] J. Ascencio-Vásquez, K. Brecl, M. Topič, Methodology of Köppen-Geiger-Photovoltaic climate classification and implications to worldwide mapping of PV system performance, Sol. Energy 191 (2019) 672–685, https://doi.org/10.1016/j.solener.2019.08.072.
[54] X. Sun, J.M. Bright, C.A. Gueymard, B. Acord, P. Wang, N.A. Engerer, Worldwide performance assessment of 75 global clear-sky irradiance models using Principal Component Analysis, Renew. Sustain. Energy Rev. 111 (2019) 550–570, https://doi.org/10.1016/j.rser.2019.04.006.
[55] A. Giler-Ormaza, A province with several climates: what knowledge of hydrology can be useful? A literature review, Maskana 9 (2018) 41–51, https://doi.org/10.18537/mskn.09.02.06.
[56] M.C. Peel, B.L. Finlayson, T.A. McMahon, Updated world map of the Köppen-Geiger climate classification, Hydrol. Earth Syst. Sci. 11 (2007) 1633–1644, https://doi.org/10.5194/hess-11-1633-2007.
[57] E. Romano, J. Mateu, R. Giraldo, On the performance of two clustering methods for spatial functional data, AStA Adv. Stat. Anal. 99 (2015) 467–492, https://doi.org/10.1007/s10182-015-0253-9.
[58] J. Cevallos-Sierra, J. Ramos-Martin, Spatial assessment of the potential of renewable energy: the case of Ecuador, Renew. Sustain. Energy Rev. 81 (2018) 1154–1165, https://doi.org/10.1016/j.rser.2017.08.015.
[59] R.C. Echegaray-Aveiga, M. Masabanda, F. Rodriguez, T. Toulkeridis, F. Mato, Solar energy potential in Ecuador, in: 2018 5th Int. Conf. EDemocracy EGovernment, ICEDEG 2018, 2018, pp. 46–51, https://doi.org/10.1109/ICEDEG.2018.8372318.
[60] M. Lave, J. Kleissl, E. Arias-Castro, High-frequency irradiance fluctuations and geographic smoothing, Sol. Energy 86 (2012) 2190–2199, https://doi.org/10.1016/j.solener.2011.06.031.
[61] J. Marcos, L. Marroyo, E. Lorenzo, M. García, Smoothing of PV power fluctuations by geographical dispersion, Prog. Photovoltaics Res. Appl. 20 (2012) 226–237, https://doi.org/10.1002/pip.1127.
[62] A. Mills, M. Ahlstrom, M. Brower, A. Ellis, R. George, T. Hoff, B. Kroposki, C. Lenox, N. Miller, J. Stein, Y. Wan, Understanding variability and uncertainty of photovoltaics for integration with the electric power system, United States, https://www.osti.gov/biblio/979812, 2009.
[63] L. Parra, S. Gómez, C. Montoya, F. Henao, Assessing the complementarities of Colombia's renewable power plants, Front. Energy Res. 8 (2020) 1–12, https://doi.org/10.3389/fenrg.2020.575240.
[64] J.O. Ramsay, G. Hooker, S. Graves, Functional Data Analysis with R and MATLAB, Springer-Verlag New York, New York, 2009, https://doi.org/10.1007/978-0-387-98185-7.