Environmental Pollution 253 (2019) 130–140

Contents lists available at ScienceDirect

Environmental Pollution

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

Compilation and spatio-temporal analysis of publicly available total solar and UV irradiance data in the contiguous United States ☆

Ying Zhou [a, 1], Xia Meng [b, 1], Jessica Hartmann Belle [b], Huanxin Zhang [c], Caitlin Kennedy [a], Mohammad Z. Al-Hamdan [d], Jun Wang [c], Yang Liu [b, *]

[a] Environmental Health Tracking Section, Division of Environmental Health Practice and Science, National Center for Environmental Health (NCEH), Centers for Disease Control and Prevention (CDC), Atlanta, GA, USA
[b] Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA
[c] Center for Global and Regional Environmental Research, University of Iowa, Iowa City, IA, USA
[d] Universities Space Research Association, NASA Marshall Space Flight Center, Huntsville, AL 35805G, USA

ARTICLE INFO

Article history:
Received 7 January 2019
Received in revised form
17 June 2019
Accepted 18 June 2019
Available online 26 June 2019

Keywords:
Ultraviolet irradiance
Solar irradiance
Tracking network
Exposure assessment
Skin cancer

ABSTRACT

Skin cancer is the most common type of cancer in the United States, the majority of which is caused by overexposure to ultraviolet (UV) irradiance, which is one component of sunlight. National Environmental Public Health Tracking Program at CDC has collaborated with partners to develop and disseminate county-level daily UV irradiance (2005–2015) and total solar irradiance (1991–2012) data for the contiguous United States. UV irradiance dataset was derived from the Ozone Monitoring Instrument (OMI), and solar irradiance was extracted from National Solar Radiation Data Base (NSRDB) and Solar-Anywhere data. Firstly, we produced daily population-weighted UV and solar irradiance datasets at the county level. Then the spatial distributions and long-term trends of UV irradiance, solar irradiance and the ratio of UV irradiance to solar irradiance were analyzed. The national average values across all years are 4300 Wh/m^2, 2700 J/m^2 and 130 mW/m^2 for global horizontal irradiance (GHI), erythemally weighted daily dose of UV irradiance (EDD) and erythemally weighted UV irradiance at local solar noon time (EDR), respectively. Solar, UV irradiances and the ratio of UV to solar irradiance all increased toward the South and in some areas with high altitude, suggesting that using solar irradiance as indicator of UV irradiance in studies covering large geographic regions may bias the true pattern of UV exposure. National annual average daily solar and UV irradiances increased significantly over the years by about 0.3% and 0.5% per year, respectively. Both datasets are available to the public through CDC's Tracking network. The UV irradiance dataset is currently the only publicly-available, spatially-resolved, and long-term UV irradiance dataset covering the contiguous United States. These datasets help us understand the spatial distributions and temporal trends of solar and UV irradiances, and allow for improved characterization of UV and sunlight exposure in future studies.

1. Introduction

Skin cancer is the most common form of cancer in the United States (U.S.) (CDC, 2018b). The two most common types of skin cancer—basal cell and squamous cell carcinomas—are highly curable, but can be disfiguring and costly to treat. It is estimated that by age 70, nearly one in five non-Hispanic white U.S. residents has had at least one of these two types of skin cancer (Stern, 2010). Melanoma is the third most common skin cancer and causes the most deaths. Melanoma incidence has increased exponentially in the United States (Rigel, 2010). In 2015, 80,442 new cases of Melanomas of the skin were reported, and 8885 people died of Melanomas of the skin in the United States (U.S. Cancer Statistics Working Group, 2018).

Sunlight is a continuous spectrum of electromagnetic radiation that is divided into three major spectra of wavelength: ultraviolet (UV), visible, and infrared. The majority of skin cancers are caused by overexposure to UV radiation. Biological effects of UV radiation vary with wavelength. Most of the UV radiation that reaches Earth's

☆ This paper has been recommended for acceptance by Dr. Sarah Harmon.

* Corresponding author. Department of Environmental Health Rollins School of Public Health Emory University Atlanta, GA 30322, USA.
E-mail addresses: yzhou2@cdc.gov (Y. Zhou), yang.liu@emory.edu (Y. Liu).

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.envpol.2019.06.074
0269-7491/© 2019 Published by Elsevier Ltd.

surface is UVA (320–400 nm), and only 5% is UVB (290–320 nm) (Brenner and Hearing, 2008; Narayanan et al., 2010). Longer wavelength UVA penetrates deeply into the dermis, and efficiently generates reactive oxygen species that can damage DNA via indirect photosensitizing reactions. UVB is almost completely absorbed by the DNA of epidermis cells, which can cause molecular rearrangements such as cyclobutane dimers and other adverse photoproducts. Mutations and cancer can result from these DNA modifications (D'Orazio et al., 2013). Pleasance et al. mapped the complete genome of a melanoma taken from a patient with the disease (Pleasance et al., 2010). The dominant mutational signature reflects DNA damage resulting from UV exposure.

There are several existing data sources for estimating UV exposure. First, the UV index predicts the UV radiation levels on a scale from 1 to 11+. It has been used as a measure of UV exposure in epidemiological studies in the United States and elsewhere (Eide and Weinstock, 2005; Lemus-Deschamps and Makin, 2012; Walls et al., 2013). While the UV Index is easy to interpret, numeric information is lost when a continuous measure of UV radiation exposure is converted to UV index. Second, some studies used solar radiation as a proxy for UV exposure (NCI, 2017; Richards et al., 2011; Tatalovich et al., 2006), such as the National Solar Radiation Database (NSRDB) developed by the National Renewable Energy Laboratory (NREL) (Wilcox, 2012). Tatalovich et al. (2006) compared UVB flux data from seven Surface Radiation Budget Network (SURFRAD) stations with global horizontal irradiance (GHI) data from nearby NSRDB stations. The correlation between UVB and GHI are high with R^2 values greater than 0.9 for the seven station pairs included. However, the spatial variation of correlation and the ratio of UV to solar irradiance over a large domain was not systematically analyzed in the study. Third, erythemal UV irradiance data retrieved by the Total Ozone Mapping Spectrometer (TOMS) have also been used for estimating UV exposure (Hatfield et al., 2009; Moan et al., 2008). Designed primarily to monitor the global distribution of column ozone, TOMS is among the first instruments to measure the backscattered UV radiation from the Earth's atmosphere and has also been used to estimate the surface UV irradiances since its inception in 1979 (Eck et al., 1987; Herman et al., 1996; Herman and Celarier, 1997). The 1979–2000 TOMS monthly mean UV index has been shown to have a 10–30% high bias against the Brewer spectrophotometer measurements in snow-free conditions and the lack of absorbing aerosols in the UV algorithm is one of reasons for this high bias (Fioletov et al., 2004). Over snow-covered surfaces, TOMS UV index has a low bias up to 60% due to the misclassification of snow as clouds (Fioletov et al., 2004).

The Environmental Health Tracking Program of the Centers for Disease Control and Prevention (Tracking Network of CDC) was established to deliver information and data to protect the nation from health issues arising from or directly related to environmental factors (EPHT, 2017). In this study, the Tracking Network and its academic partners generated a long-term, high-resolution UV irradiance dataset for all counties in the contiguous United States using the latest satellite remote sensing data for improved characterization of UV exposure. Given the relatively short duration of the latest UV irradiance data (2005–2015), a solar irradiance dataset with 22 years of data (1991–2012) was also generated. The spatial distributions and long-term temporal trends of solar and UV irradiance estimates based on these two datasets were explored. The newly developed UV and solar irradiance estimates in this analysis also allow us to study the spatial variations of the ratio of UV irradiance to solar irradiance.

2. Data and methods

Irradiance is the density of radiation incident on a given surface, which can be expressed in watts per square meter (W/m^2). For example, total solar irradiance is an instantaneous measure of solar intensity of all wavelengths per unit area. Global horizontal irradiance (GHI) is widely used to represent the total solar irradiance. GHI is the total solar radiation incident on a horizontal surface on the earth, mainly including direct normal irradiance (DNI), which is the solar (beam) radiation directly coming from the sun, and diffuse horizontal irradiance (DHI), which is the scattered solar radiation from the sky dome (Sengupta et al., 2017). Data for UV irradiance, a component of solar irradiance, is presented in terms of the now widely accepted “sunburning” or erythemally-weighted UV radiation. This method is becoming increasingly popular due to previous issues in defining a biologically-relevant UVB spectrum and measurement error in quantifying UVB with current sensors (McKenzie et al., 2004). The framework of this project was summarized in Fig. 1, showing that we produced a UV irradiance dataset from 2005 to 2015 based on UV data from Ozone Monitoring Instrument (OMI), and a solar irradiance dataset from 1991 to 2012 based on solar irradiance data from NSRDB and SolarAnywhere. To understand how well our solar irradiance estimates compare with previous estimates, we compared them with North American Land Data Assimilation System (NLDAS) solar irradiance data. Table 1 summarized the sources, time periods, resolutions of the datasets used, as well as the methods used to assign data in the original datasets to census tracts to produce the final UV and solar irradiance datasets. The brief introductions regarding the algorithms of the original UV and solar irradiance datasets used in this project are in Part 1 of supplementary material.

2.1. UV dataset

The Ozone Monitoring Instrument (OMI) flies on the National Aeronautics and Space Administration (NASA)'s Earth Observing System Aura satellite launched in July 2004 (Levelt et al., 2006). The instrument is a contribution of the Netherlands's Agency for Aerospace Programs (NIVR) in collaboration with the Finnish Meteorological Institute (FMI) to the Aura mission. OMI measures irradiances of UV and a fraction of visible light (270 nm - 500 nm) at spectral resolutions often finer than or comparable to the Brewer spectrophotometer used to measure global UV spectral irradiance on the ground (i.e., 0.55–0.60 nm depending on specific instruments) (Antón et al., 2010; Fioletov et al., 2004). OMI can

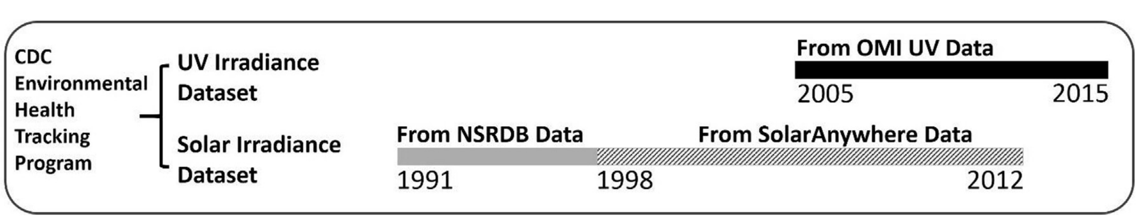

Fig. 1. The time frame and data sources of the UV and solar irradiance datasets.

Table 1
Summary of input data sources and newly produced UV and solar irradiance datasets.

Dataset name	Time period	Parameters	Input dataset name	Spatial resolution of input data	Methods for assigning input data to census tracts
UV irradiance dataset	2005–2015	EDD, EDR, I305, I310, I324 and I380	Level-2 OMI surface UV data (v.003) (Levelt et al., 2006)	13 km × 24 km at nadir	Finding the nearest pixel
Sunlight irradiance dataset	1991–1997	GHI	NSRDB (Wilcox, 2012)	1454 locations, unevenly distributed	Universal kriging with elevation as a predictor
	1998–2012	GHI	SolarAnywhere (Perez et al., 2002; Wilcox, 2012)	10 km × 10 km	Match when the centroid of the census tract falls into the SolarAnywhere grid
Solar irradiance dataset for comparison	1991–2011	Surface downward shortwave radiation	NLDAS solar irradiance data (CDC, 2018a)	1/8th-degree (~14 km)	Aggregate the values of grids with centroids located within the corresponding county's boundary

provide daily coverage of the sunlit portion of the atmosphere.

The Level-2 OMI surface UV data (v.003) (Jari Hovila and Tamminen, 2007) were used in our study. For year 2005–2015, we first extracted the following parameters from OMI surface UV dataset to include in our UV dataset: erythemally weighted daily dose (EDD) (J/m^2), erythemally weighted irradiance at local solar noon time and OMI overpass time (EDR) (mW/m^2), and daily spectral irradiance at local solar noon time at 305, 310, 324, and 380 nm (I305, I310, I324 and I380) (mW/m^2/nm). EDD represents the accumulated amount of UV radiation that can cause sunburn over the course of the day. EDR represents the amount of UV radiation that can cause sunburn around noon when the dose is likely the highest. To estimate EDR at noon, OMI first measures EDR at its overpass time (~13:45 local solar time). Then, EDR at overpass time was scaled to the local solar noon time using different solar zenith angles, while assuming a constant atmospheric profile, e.g., aerosol loadings, clouds. Spectral irradiance at 305 nm and 310 nm is part of UVB irradiance, while spectral irradiance at 324 nm and 380 nm is part of UVA irradiance. The algorithm of OMI UV dataset was summarized in Part 1 of Supplementary material.

OMI surface UV data has the nominal resolution of 13 km × 24 km (along × across) at nadir, and the pixel size grows quickly across the scan, which can be up to 13 km × 128 km at the most outer swath-angle (57°) (http://projects.knmi.nl/omi/research/instrument/characteristics.php?tag=full). After extracting the variables described above for all OMI pixels, we assigned UV irradiance values to U.S. census tracts by finding the nearest OMI pixel within a search radius of 100 km of the census tract. The "row anomaly" occurred due to the technical issues of OMI which produced invalid data in the center-right part of each swath of observations (McPeters et al., 2015) since 2008. Hence census tracts matched with invalid UV irradiance values were assigned "not applicable" for that day. Finally, county level UV irradiance values were calculated using population weighted census tract values in each county based on Equation (1), which puts more weight on census tract with more people.

$$\text{Population} - \text{weighted UV irradiance} = \frac{\sum UV_t \times POP_t}{POP_c} \qquad (1)$$

Where the UV_t is the UV irradiance at tract t, POP_t is population in tract t, and POP_c is the total population of all tracts in county c. Note that census tracts are small, relatively permanent statistical subdivisions of a county or county equivalent in the population census data.

2.2. Solar irradiance dataset

Given the limited number of years in the current UV irradiance data (2005–2015), we also generated a population-weighted solar irradiance dataset which includes daily GHI (Wh/m^2) from 1991 to 2012. There were two data sources of GHI data used in this project—National Solar Radiation Data Base (NSRDB) and SolarAnywhere.

The updated 1991–2010 NSRDB dataset (Wilcox, 2012) holds solar data for 1454 locations (Fig. S1) in the United States and its territories. Nearly all of the solar data in the NSRDB are modeled, and only 40 sites have measured solar data—none of them with complete records. The hourly NSRDB GHI data produced by the METSTAT (Meteorological/Statistical) model were used in this study. The SolarAnywhere data (version 2.4) is gridded data at a 10 × 10 km resolution since 1998. The SolarAnywhere algorithm estimates solar radiation based on Geostationary Operational Environmental Satellite (GOES) imagery (Perez et al., 2002; Perez et al., 2010). The algorithms of NSRDB and SolarAnywhere datasets were summarized in Part 1 of Supplementary material. The hourly GHI data of SolarAnywhere during 1998–2012 were provided to CDC and academic partners via data-use agreement (https://www.solaranywhere.com/).

The original GHI values from both NSRDB and SolarAnywhere were at hourly level. We calculated the daily GHI values by summing the 24 hourly GHI values if hourly data were available for all 24 h within the calendar day; otherwise, the daily GHI was marked as "not applicable" for that day. The final GHI dataset were produced based on NSRDB data from 1991 to 1997 and SolarAnywhere data from 1998 to 2012.

For data from NSRDB, we used universal kriging (UK) with elevation as a predictor to interpolate daily GHI values to the nearest census tracts. Part 2 of supplementary material provided more details on model selection and evaluation. To interpolate GHI values, there were several steps involved. First, we assumed that data within the max search distance have spatial autocorrelations. Therefore, we fitted the semivariogram with the average annual mean GHI values from 1991 to 2010 to find the distance within which GHI values were spatially auto-correlated. Second, the UK model was built with elevation as a predictor at daily level, by setting the maximum distance at 500 km, based on the range from the semivariogram in the first step. Lastly, we predicted the daily GHI values at the centroids of the census tracts using the UK model.

For data from SolarAnywhere, we assigned the gridded GHI value to one census tract if the centroid of the census tract fell into the SolarAnywhere grid. After assigning the NSRDB or SolarAnywhere GHI data to census tracts, we calculated county level GHI values using population weighted census tract values, using a similar approach for UV irradiance data as shown in Equation (1).

For years 1998–2005, data are available from both NSRDB and SolarAnywhere. For these overlapping years, we chose to use SolarAnywhere in the solar irradiance dataset because: 1) the agreement between measured and modeled GHI was better (regression $R^2 = 0.88$) for SolarAnywhere than that for NSRDB

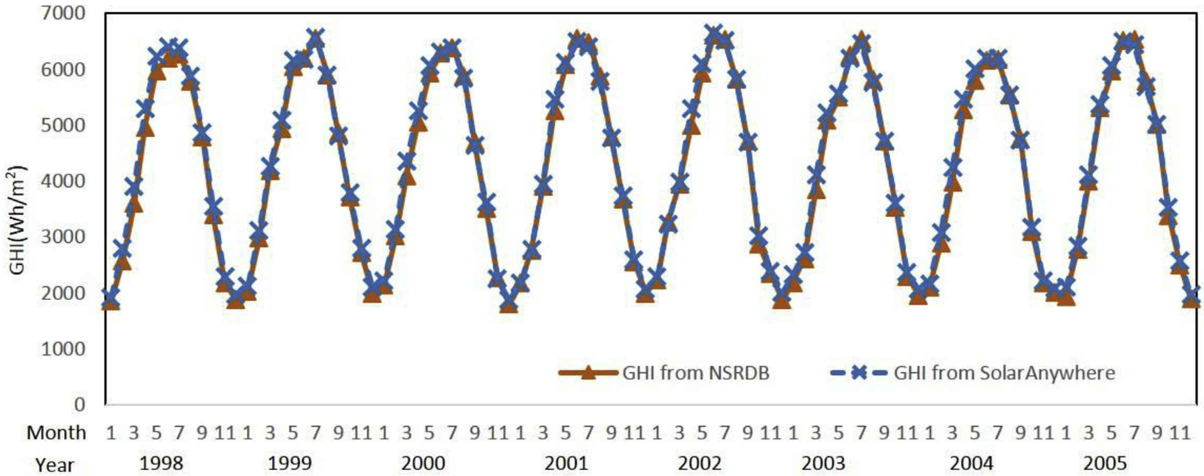

Fig. 2. Comparison of monthly national means of daily global horizontal irradiance (GHI), from NSRDB and SolarAnywhere in 1998–2005.

(regression $R^2 = 0.83$) (Myers et al., 2005; Wilcox, 2012), 2) the NSRDB was developed based on data from 29 ground monitoring sites while the SolarAnywhere algorithm runs on satellite images with comprehensive spatial coverage, and 3) the NSRDB data at 1454 sites were distributed unevenly in space, while SolarAnywhere provided gridded data at 0.1° resolution nationwide. We used these overlapping years to evaluate the consistency of the two data sources before merging NSRDB data (1991–1997) with SolarAnywhere data (1998–2012) into the final solar irradiance dataset. Fig. 2 shows the time series of the monthly national means of population-weighted total solar irradiance levels calculated from NSRDB and SolarAnywhere products for 1998 to 2005. They are highly consistent.

The linear regression model (Equation (2)) was fitted based on the monthly means of population weighted solar irradiance at that state level using GHI values from NSRDB and SolarAnywhere between 1998 and 2005. The scatterplot is shown in Fig. 3.

$$\text{SolarAnywhere\_GHI} = 131.66 + 0.99 \times \text{NSRDB\_GHI} \quad (2)$$

Fig. 3 and regression analysis results show that the NSRDB irradiance is highly correlated with SolarAnywhere irradiance when considering both temporal (month) and spatial (state) distributions. The R^2 is 0.99 and slope is 0.99 in the regression, with solar irradiance levels from SolarAnywhere being slightly higher than those from NSRDB. To ensure the consistency of the combined GHI estimates, we calibrated county level daily GHI values from NSRDB for 1991–1997 using Equation (2). We then combined the calibrated estimates from 1991 to 1997 with the estimates from 1998 to 2012 based on SolarAnywhere and created the dataset of daily solar irradiance from 1991 to 2012 at the county level.

2.2.1. Comparison with North American Land Data Assimilation System (NLDAS) solar irradiance data

We compared our solar irradiance estimates with North American Land Data Assimilation System (NLDAS) solar irradiance data (1/8th-degree (~14 km) resolution) to further validate the long-term trends of our solar irradiance estimates. NLDAS data is also modeled data and is available through CDC WONDER online database (CDC, 2018a). The county-level data were computed by averaging the values of all the grid cells whose centroids were located within the corresponding county's boundary. In the cases where the county was so small that no grid cell's centroid is located in it (i.e., the county is smaller than one grid cell), the value of the grid cell that covers the largest portion of the county's area was assigned

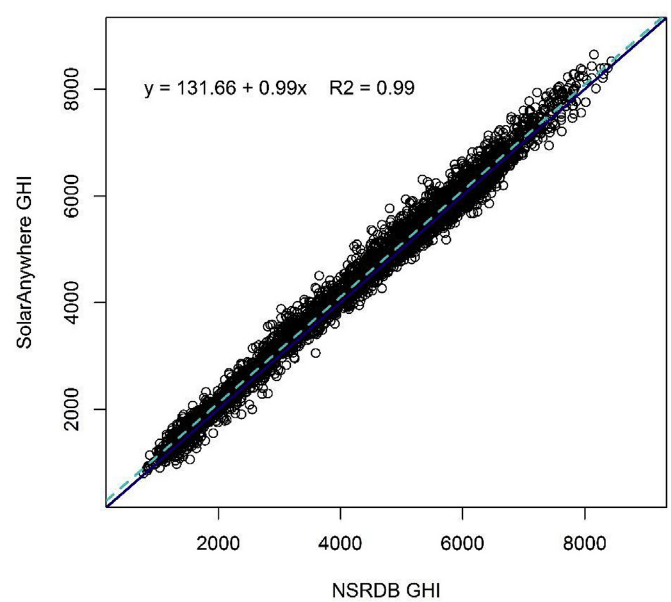

Fig. 3. Comparison between SolarAnywhere GHI and NSRDB GHI used for dataset calibration between 1998 and 2005. The solid line is the 1:1 line and the dashed line represents the regression of monthly mean GHI data at state level between NSRDB and SolarAnywhere between 1998 and 2005. The unit for X-axis and Y-axis are both Wh/m^2.

to such county.

To compare the two datasets, we first divided WONDER GHI data by 3.6 to convert the unit from kJ/m^2 to Wh/m^2, which is the unit for NSRDB and SolarAnywhere GHI data. Then we compared the annual means between the two datasets at national and county levels.

2.3. Spatial distribution and temporal trend analysis

Summary statistics were calculated for each environmental measure at the national and state level, respectively. In addition, for the years in which both UV and solar irradiance data are available (2005–2012), we calculated the ratio of EDD to GHI by county to study the spatial variations of the ratio of UV irradiance to solar irradiance.

To determine the long-term temporal trends of solar irradiance, we conducted linear regression using "year" as independent variable, and annual averages GHI as dependent variable, at the national and state level, respectively in SAS 9.4 (SAS Institute Inc, Cary, NC, USA). The regression coefficient for the independent variable "year" shows the average change in GHI per year. Similarly, we conducted linear regression using annual average EDD and EDR as dependent variable respectively, and "year" as independent variable to explore the long-term temporal trends of UV irradiance. In addition, as there are natural fluctuations in solar and UV irradiances, linear regression may not always provide the best fit when determining long-term trend. Therefore, as a sensitivity analysis, we used Joinpoint Trend Analysis Software (NCI, 2018) to test for an

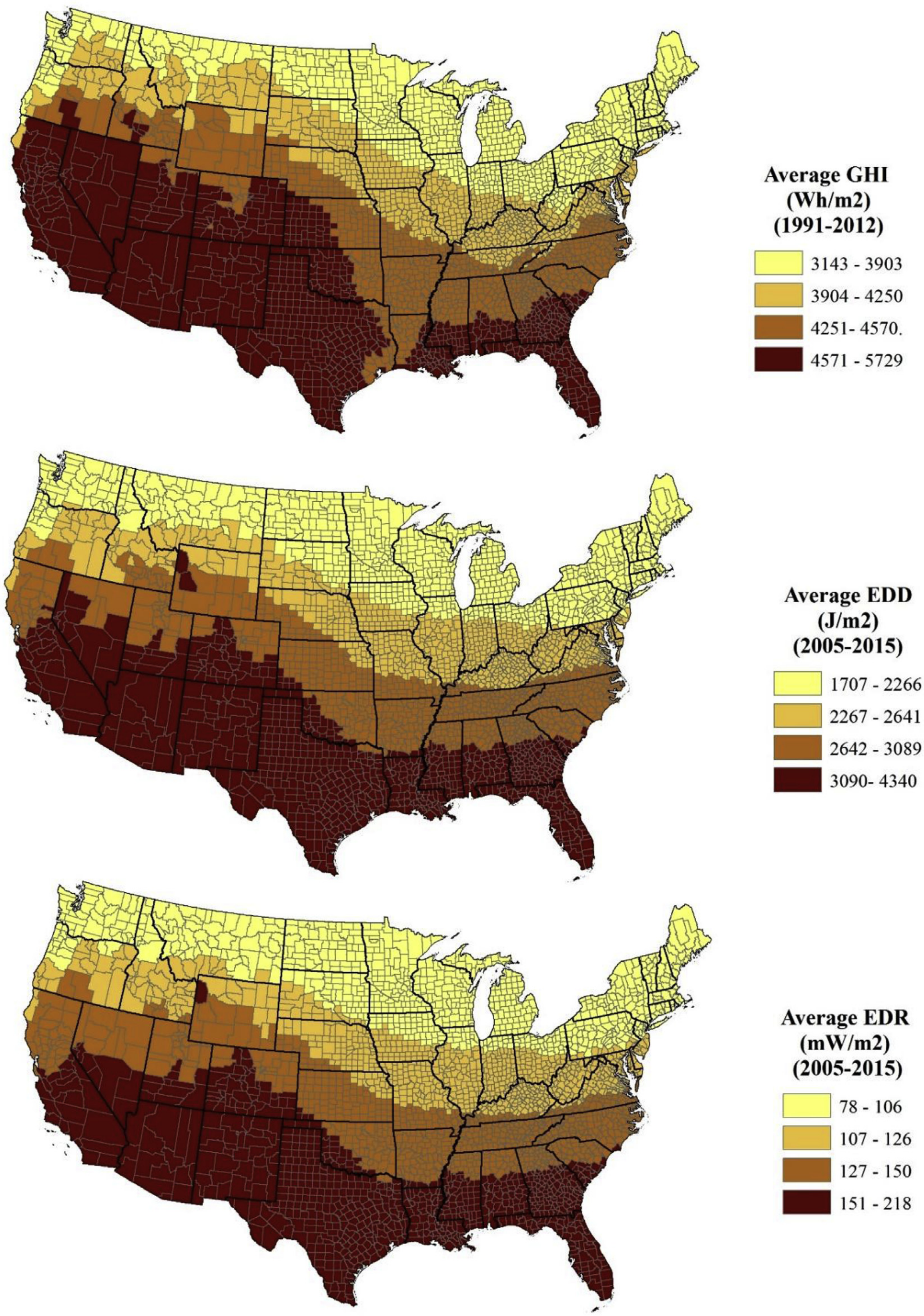

Fig. 4. Spatial distribution of average daily global horizontal irradiance (GHI), erythemally weighted daily dose (EDD) of UV, and erythemally weighted irradiance (EDR) of UV. Intervals are based on quartiles. County boundary for year 2015 was used.

apparent change in trend that is statistically significant. Joinpoint is a statistical software for the analysis of trends which allows several different lines to be connected together at the “joinpoints”. We allowed a minimum of 0 joinpoint (which means a straight line) and a maximum of 4 joinpoints for solar irradiance and 1 joinpoint for UV irradiance. Note that the maximum number of joinpoints are determined based on the number of data points in these datasets (NCI, 2018).

3. Results

3.1. Summary statistics

Table S1 in supplementary material shows the summary statistics of GHI for years 1991–2012, and EDD and EDR for years 2005–2015 on the state level. The national averages across all years available for each variable are 4300 Wh/m^2, 2700 J/m^2 and 130 mW/m^2 for daily GHI, EDD and EDR, respectively. Spectral irradiances at 305, 310, 324, and 380 nm are on average 35, 64, 264, and 491 mW/(m^2·nm) in the contiguous United States from 2005 to 2015. Average daily dose of UV irradiance (EDD) accounts for an average of 0.016 percent of average daily total solar irradiance (GHI).

3.2. Spatial distributions

Fig. 4 shows the spatial distributions of average daily GHI during 1991–2012 and average EDD and EDR during 2005–2015. In summary, the UV and solar irradiances both increased toward the South and the West (likely due to higher altitude). The highest state averages for GHI, EDD and EDR all occurred in Arizona. The lowest averages occurred in Vermont, Maine, and North Dakota for GHI, EDD and EDR, respectively.

The correlation coefficient between average annual mean EDD and GHI at county level was 0.87. Fig. 5 illustrates the spatial distribution of the ratios of EDD to GHI, which shows the percentage of UV irradiance in total solar irradiance. Fig. 5 shows that the ratios mainly increased toward the South and in some areas with high altitude, e.g. Colorado, New Mexico and west Wyoming. Even though UV and solar irradiances both increased toward the South and the West (as shown in Fig. 4), and they are highly correlated, the spatial pattern of the ratio of UV to solar irradiance in Fig. 5 means that the increase in UV irradiance was faster than solar irradiance in the South and in areas with high elevation. This suggests that UV exposure may be underestimated when solar irradiance is used as proxy in areas with high UV to solar irradiance ratio and vice versa.

3.3. Temporal trends

Fig. 6 suggests that temporal patterns of solar irradiance data we derived based on NSRDB and SolarAnywhere data are mostly consistent with NLDAS solar irradiance data at the national level. The average difference is 5% between our data and NLDAS data at county level. The average difference between year 1991 and 1997 is 6.8%, which is slightly larger than the average difference of 4.0% between 1998 and 2011, when we switched from using NSRDB data to SolarAnywhere data as input. The largest difference (9.4%) is in 1992 while the smallest difference (2.2%) is in 2002. To investigate whether combining NSRDB and SolarAnywhere data caused the larger difference between NLDAS and our solar irradiance data before 1998, we also calculated the annual mean values of GHI from all NSRDB stations with valid GHI values between 1991 and 2005. Fig. S2 in supplementary material shows that the difference in GHI values between NLDAS data and NSRDB stations data is also larger before 1998, suggesting that combining the NSRDB and SolarAnywhere data is unlikely the main cause of the larger difference between NLDAS and our irradiance data before 1998.

Trend analysis based on newly produced solar irradiance data in this study shows that there is a statistically significant increase in national average daily GHI of 14 Wh/m^2 (0.3%) per year from 1991 to 2012 (p-value <0.01) (Fig. S3a). Linear regressions at state level suggest that total solar irradiance increased significantly in all states except for Maine, Massachusetts, Rhode Island, Vermont, North Dakota, and Montana (Fig. 7). The national annual average UV irradiances also have increased significantly over the years.

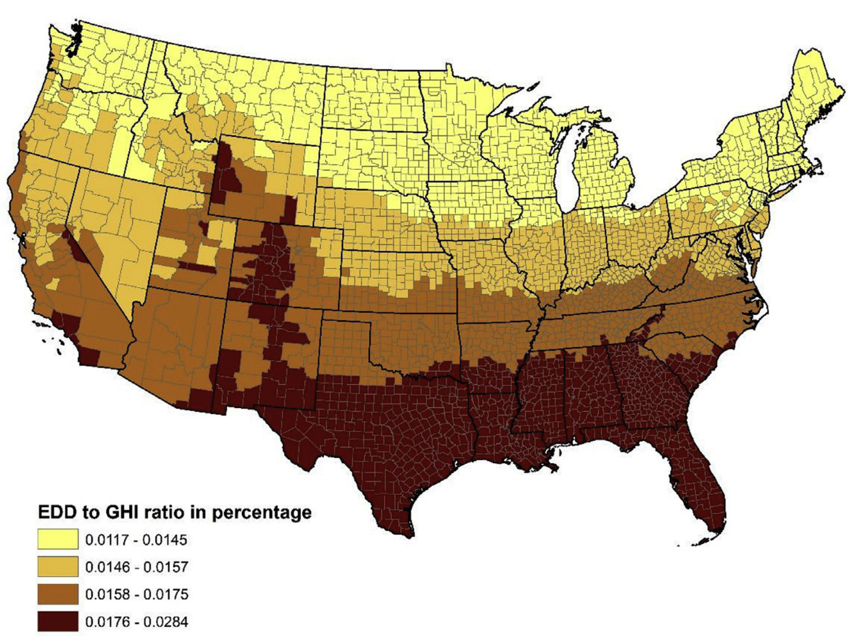

Fig. 5. Spatial distribution of the ratio of UV (EDD) to solar irradiance (GHI) (in percentage) between 2005 and 2012. Intervals are based on quartile.

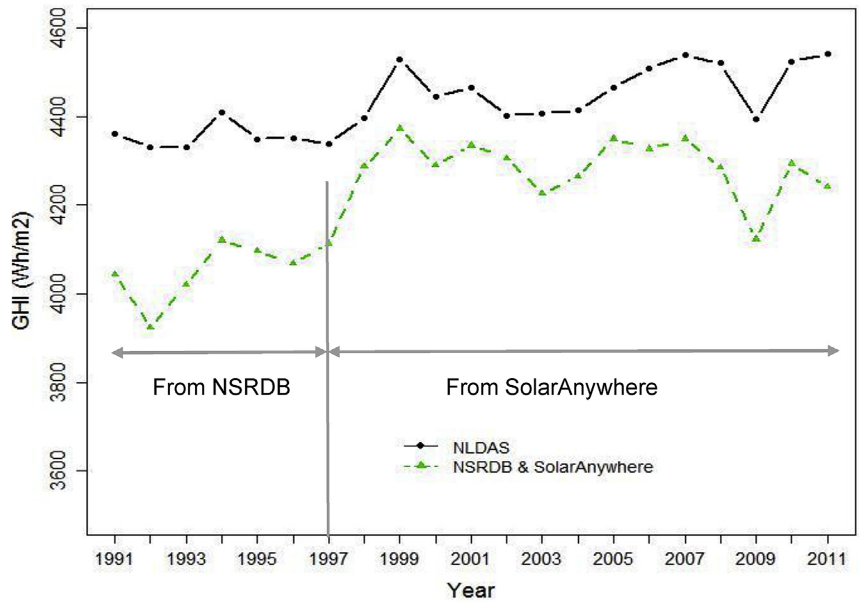

Fig. 6. Comparison of county annual average of daily global horizontal irradiance (GHI) from NLDAS data with GHI estimates based on combined NSRDB and SolarAnywhere data.

There is a statistically significant increase of 13 J/m^2 (0.5%) per year in EDD (p-value <0.01) (Fig. S3b) and a statistically significant increase of 0.52 (mW/m^2) (0.4%) per year in EDR (p-value = 0.02) (Fig. S3c) on average at the national level. For the two measures of UV irradiance, 25 states had statistically significant increasing trend in EDD from 2005 to 2015, while 18 states had statistically significant increasing trend in EDR (Fig. 7). We calculated national average EDD and EDR by averaging the county estimates. Therefore, states with more counties are likely to have more impact on national level estimates. The 5 states with most counties are Texas, Georgia, Virginia, Kentucky, and Missouri. In addition, the states which increased enough to impact the national level increase are likely those with the biggest regression coefficients in the trend analysis. For EDD, the 5 states with the biggest regression coefficients are: New Mexico, Kentucky, California, South Carolina and New Hampshire. For EDR, the 5 states with the largest regression coefficients are: California, Kentucky, New Hampshire, New Mexico and Massachusetts.

3.4. Sensitivity analysis

In the sensitivity analysis for solar irradiance, when we allowed joinpoints, results for the national average and 33 out of 48 contiguous states and the District of Columbia remained the same, i.e., there is no change in trend and linear regression provides the best fit. Fig. S4a is an example of the output from Joinpoint analysis for Alabama, which shows no change in trend. The remaining 16 states had statistically significant change in the trend. For example, Fig. S4b shows the result from Joinpoint analysis for GHI in Colorado. A joinpoint was identified in 2002. Before 2002, there was a significant increase in annual average daily GHI of 37 Wh/m^2 per year. After 2002, no significant trend was detected. The majority of the 16 states had one significant change in trend which occurred between the years of 1998–2006 (See supplemental material, Table S2). The general temporal trend pattern in these 16 states is similar: a significant increase in GHI for the years before the first join point (i.e., slope significantly greater than 0), followed by a change in slope, e.g. leveling off (Fig. S4b), or decrease. Four states had 2 joinpoints (Arkansas, Kansas, Missouri, South Dakota) and Mississippi has 4 joinpoints, though none of these slopes after the first joinpoint is significantly different from 0. Fig. S4c is an example of Kansas which had two joinpoints. For EDD, when we allowed join points, one state (New Mexico) had a significant change in trend: there was a significant increase in EDD between 2005 and 2011, and no significant change after 2011. For EDR, two states (Arizona and New Mexico) had one joinpoint in 2011 following the same pattern as EDD.

4. Discussion

4.1. Spatial and temporal trends of the UV and solar irradiances data

In this study, we produced daily county level UV (2005–2015) and solar irradiance estimates (1991–2012) in the contiguous United States. Solar irradiance and UV irradiance both generally increased toward the South and the altitude, suggesting that latitude and altitude increase UV irradiance as they increase. The ratio of UV to solar irradiance also mainly increased toward the South and in some areas with high altitude. The national annual average solar and UV irradiances have increased significantly over the years included in this analysis. UV irradiance can be impacted by a series of complex factors, including solar activity, geographical parameters (e.g. latitude, elevation and features of the receiving terrain), atmospheric absorption (e.g., absorption by ozone depending on wavelength range), scattering (e.g., the Rayleigh scattering of air molecules and the Mie scattering of clouds and aerosol particles), and surface reflection (Chang et al., 2012). For spatial distribution, the increase to the West is likely due to the higher elevation in the Western half of the United States. Earlier studies reported similar findings. For example, using measurements from 29 observation sites in the United States and its borders, Gao et al. found that latitude and altitude were the principles factors that regulate average daily UV dosage (Gao et al., 2007). The study reported that statistically significant nonlinear relationships can be established between averaged daily UV dose and latitude and altitude, with latitude having a more significant effect than altitude. In comparison, longitude was not statistically significant in predicting UV

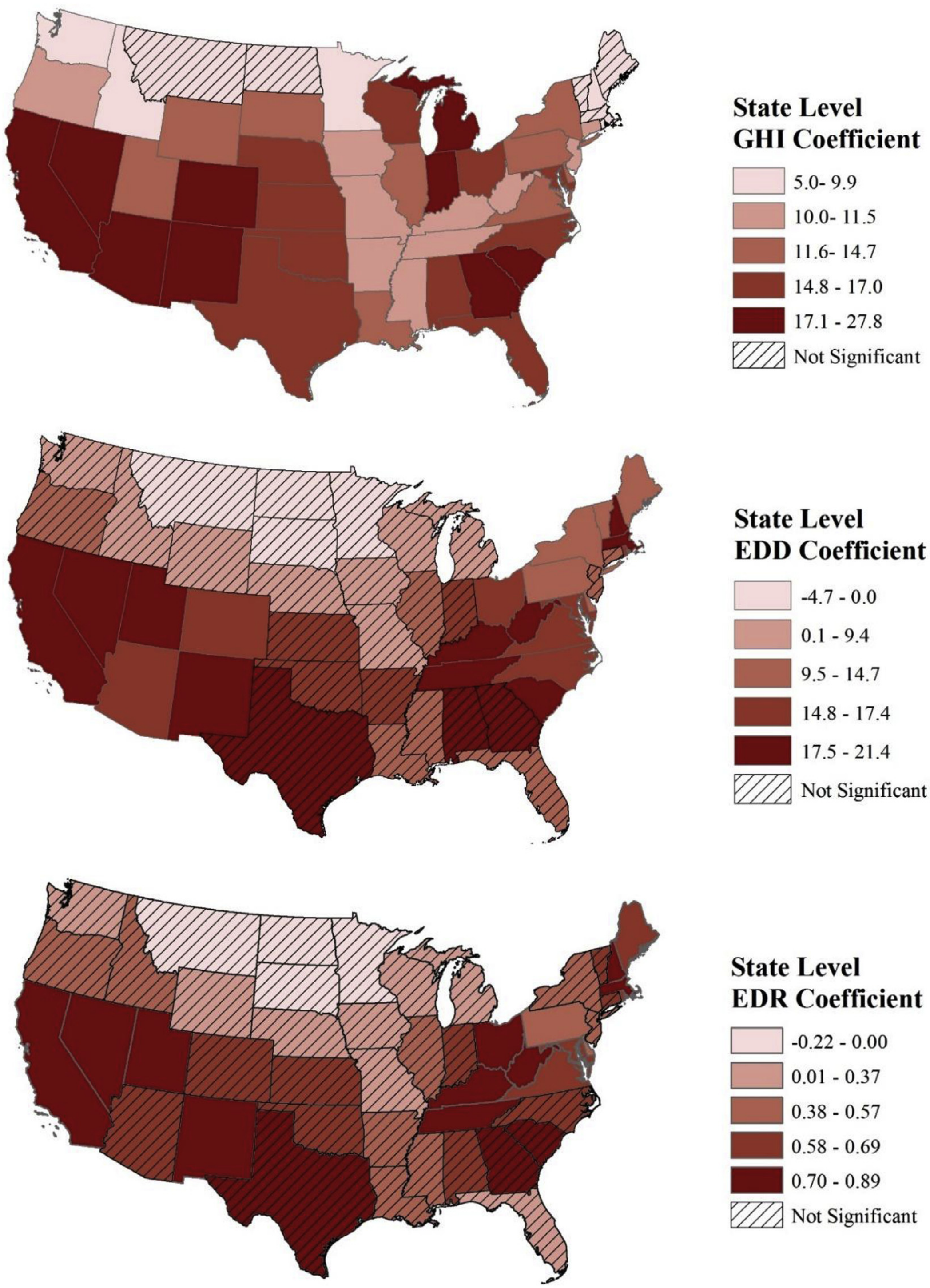

Fig. 7. Regression coefficients in trend analysis for daily horizontal irradiance (GHI), erythemally weighted daily dose (EDD) of UV, and erythemally weighted irradiance at local solar noon time (EDR) of UV by state. These regression coefficients show the average change per year of daily GHI (in Wh/m^2 per year), EDD (in J/m^2 per year), and EDR (in mW/m^2 per year), respectively.

irradiance. The Colorado and Indiana sites in that study have similar latitude, however, UV irradiances in the Colorado site, which is 3000 m higher in altitude, was significantly higher than the Indiana site (Gao et al., 2007). Another study found that daily erythemal UV doses on the Tibetan Plateau, the highest plateau in the world, were about twice those observed at geographically close, lower-altitude locations, such as Chengdu in China and New Delhi in India (Norsang et al., 2011). Several factors could contribute to the altitude effect. The higher the altitude, the thinner, dryer and cleaner the air, and therefore there is less atmospheric scattering and absorption of UV (e.g., by tropospheric ozone), resulting in higher UV radiation levels at the surface (Chang et al., 2012; Norsang et al., 2011).

For temporal trend, fewer states showed significant increase in

UV irradiance than solar irradiance. This could be because UV irradiance has increased less significantly than solar irradiance, or because there are fewer years of UV irradiance data available than solar irradiance data. We also plan to expand the dataset to cover more years after 2015 as OMI data become available. Several factors have been linked with the variations of UV and solar irradiance data. In situ surface observations showed that downward surface solar radiation in two surface observation sites in Illinois and Mississippi increased by 0.58 and 1.0 W/m^2 per year respectively between 2000 and 2014 (Cusworth et al., 2017). Their modeling results showed that decreasing aerosols may have driven clear-sky downward surface solar radiation in these two surface observation sites (Cusworth et al., 2017). One study using data generated from TOMS (predecessor of OMI) found that the UVB trend over the United States between 1980 and 2002 was mainly driven by the effects of total ozone variations combined with some auxiliary factors such as aerosols and total cloud amount (Chang et al., 2012). Another study assessed the influence of synoptic weather patterns on solar irradiance variability over time in northern Europe (Parding et al., 2016). It concluded that while large-scale atmospheric circulation can explain some variability in shortwave irradiances at the Earth's surface, other factors (e.g., decreasing aerosol emissions) also play an important role in northern Europe. We explored the temporal trend of cloud amount data (Fig. S5) from International Satellite Cloud Climatology Project (ISCCP, https://www.ncdc.noaa.gov/isccp) and found a negative correlation between cloud amount and solar irradiance. For example, the cloud amount dropped dramatically from 1997 to 1998, when the solar irradiance increased; the cloud amount increased in 2009, when the solar irradiance decreased in that year. Taking the change between 2008 and 2009 for example, the differences of cloud amount between 2008 and 2009 was highly negatively correlated with the differences of solar irradiance at state level ($R^2 = 0.7$, Fig. S6). The spatial distributions of the increase in cloud amount and the decrease in solar irradiance are highly consistent (Fig. S7), suggesting that the changes of cloud amount might be one of the factors leading to the change in solar irradiance in year 2009. Future study is needed on the potential mechanisms that result in the variations of solar and UV irradiances.

4.2. Comparison with previous GHI interpolations

Tatalovich et al. (2006) compared the model performances based on universal kriging (UK) method and ANUPLIN package, which was based on thin plate smoothing splines (TPS), with 30-year (1961–1990) averaged GHI values from 239 NSRDB stations and found that the root mean square error (RMSE) of UK was significantly higher than that from ANUPLIN. Limited by the access of ANUPLIN package, we compared the model performance of UK and TPS for our data with R packages ('gstat' and 'fields', respectively). We found that they were comparable in model performance at daily level (Part 2 in supplementary material). The reasons for the different results between our study and Tatalovich et al. (2006) might be the differences in study periods, temporal resolution (30-year mean vs. daily), number and spatial distributions of NSRDB stations involved, and packages used for conducting the TPS procedure. For example, Tatalovich et al. (2006) interpolated 239 station-based GHI values to grids at much finer resolution (1 km × 1 km). In our study, we interpolated GHI values from 1454 locations to the centroids of census tracts. The difference in spatial resolution (i.e., from 239 stations to 1 km × 1 km grids vs. from 1454 locations to census tracts) may have contributed to the difference in performance of the different interpolation methods.

4.3. Strengths

The county level daily UV irradiance dataset developed in this project is to our knowledge currently the only publicly available spatially resolved long-term UV irradiance dataset covering the contiguous United States. Along with the solar irradiance data, our products provide both direct and indirect estimates of UV exposure for future epidemiological studies. There are several strengths in our study.

First, when UV Index is used, UV exposure with the same UV Index value is assumed to be the same. Our product provides UV irradiance estimates as a continuous value which allows for more accuracy and flexibility in quantitative analysis. For example, on one hand, exposure-response relationship in epidemiology studies can be calculated based on UV irradiance estimate as a continuous variable. On the other hand, UV irradiance values can be easily converted to UV Index if there is a need to compare with previous studies that used UV Index to estimate UV exposure. Second, comparing with previous TOMS UV data, OMI UV data we used as input in this study has higher spatial resolution, finer spectral resolution, and lower bias. The spatial resolution for OMI surface UV dataset is 13 km × 24 km at nadir and its spectral resolution is about 0.5 nm from 270 nm to 500 nm, whereas the spatial resolution of TOMS is 50 × 50 km^2 at nadir and it only has six 1-nm wide spectral bands in the UV range. Additionally, previous research found a good agreement between OMI UV estimates and ground measurements, as shown by strong correlation with the ground measurements ($r = 0.88$) and low bias. Overall, OMI data underestimates overpass EDR by ~4%, solar noon time EDR by ~8% (Zhang et al., 2019). Third, our products are designed especially for health researchers who are not familiar with processing satellite and raw solar irradiance data. Our UV and solar irradiance data can be linked with health data at county level directly by the county Federal Information Processing Standard (FIPS) code. Our population-weighted estimates give more weight to more populous census tracts, and hence are more relevant for studying the corresponding population health effects due to UV exposure. Fourth, our results show that the UV proportion of solar irradiance mainly increases toward the South and in high altitude areas, suggesting that using solar irradiance as indicator of UV irradiance may bias the true pattern of UV exposure in studies covering large geographic regions. We provide both UV and solar irradiance data for 11-overlapping years, which allow researchers to calculate the UV proportion of solar irradiance and use it to calibrate the solar irradiance to better estimate the UV exposure when UV data are not available. Finally, our UV and solar irradiance datasets are at the daily level and could be aggregated easily to longer time intervals. The high temporal resolution and flexibility could support health effects studies of both long-term and short-term UV exposures. There are two ways to access these datasets through the Tracking Network. They can be directly downloaded at https://ephtracking.cdc.gov/download. Currently, county-level daily Global Horizontal Irradiance (GHI) data are available daily from January 1991 to December 2012, and county-level daily OMI UV irradiance data are available daily from October 2004 to December 2015. Alternatively, based on these datasets, we developed indicators and measures under the content area "Sunlight and UV." Users can access these measures through the Data Explorer on Tracking Network (EPHT, 2017).

4.4. Limitations

There are several limitations of our study. First, the current OMI surface UV algorithm assumes that aerosol loadings, the total column ozone, and the cloud optical depth remain the same throughout the day. This could cause errors when estimating EDD or scaling OMI overpass time EDR to local noontime EDR. Future geostationary satellite such as TEMPO, Sentinel-4 and GEMS would be able to help resolve the issue. Second, the current OMI surface UV algorithm uses a monthly aerosol climatology to correct the absorbing aerosol effects, which may not account for the daily and diurnal variations. Note that in spite of this assumption, the comparisons between ground observational data and EDR estimates based on OMI surface UV algorithm at OMI overpass time showed good agreement with ground observations, with high correlation and low bias. Third, we found good agreement between our solar irradiance estimates and NLDAS data, but there was still a small systematic difference between them, which was larger before 1998. This difference might be because we calculated our county-level GHI based on population-weighted census tract values, while NLDAS did not take population into consideration. In addition, different data sources and methods for estimating solar irradiances were used. For example, the native spatial resolutions are 10 km $\times$ 10 km for SolarAnywhere, 14 km $\times$ 14 km for NLDAS, while 1991 to 2010 NSRDB contains data for 1454 sites, which are not evenly distributed (Wilcox, 2012). Additionally, the interpolation from NSRDB stations to census tracts could also introduce errors. However, since NLDAS is also modeling data (instead of real measurement data), this comparison is mainly intended to help us understand the variability of solar data from different data sources. Fourth, slope is another factor affecting the solar energy incident (Dubayah and Rich, 1995), as the south facing slopes receive more heat than the north facing slopes in the Northern Hemisphere. Tatalovich et al. (2006)'s analysis reported that 23 percent of counties in the continental USA have predominantly north- or south-facing slopes, which could introduce some bias to our solar irradiance estimates in those counties between 1991 and 1997. Therefore, the exposure uncertainty should be paid attention in future epidemiological studies, especially in areas where have predominantly north- or south-facing slopes.

5. Conclusion

In this study, we generated and made publicly available two spatially resolved long-term datasets for UV and solar irradiance data covering the contiguous United States. The UV irradiance dataset is currently the only data source for surface UV irradiances for all counties in the contiguous United States after 2005. Spatially, solar and UV irradiances both increased toward the South and the West (likely due to higher altitude). The ratio of UV to solar irradiance mainly increased toward the South and in some areas with high altitude, suggesting that simply using solar irradiance as indicator in studies covering large geographic regions may bias the true pattern of UV exposure, especially at large spatial scale. Our datasets make it possible to quantitatively characterize the bias associated with using solar irradiance as a proxy for UV exposure. Temporally, the national annual average solar and UV irradiances both increased significantly over the years included in this analysis. These newly generated datasets allow for improved characterization of solar and UV exposure in future skin cancer epidemiology studies and for designing targeted public health interventions.

Declaration of interests

All authors hereby state that a) our study does not involve human subjects; b) the manuscript is not being considered for publication elsewhere; c) there's no financial/personal interest or belief that could affect our objectivity; d) all authors have read the manuscript and are in agreement that the work is ready for submission to Environmental Pollution and that we accept the responsibility for the manuscript contents.

Acknowledgement

The work of J. Belle, X. Meng, J. Wang, H. Zhang and Y. Liu was partially supported by the National Aeronautics and Space Administration (NASA) of the United States Applied Sciences Program, managed by J. Haynes (grant no. NNX14AG01G and NNX16AQ28G, PI: Y. Liu). We thank Dr. W. Dana Flanders for his advice on the trend analysis. The findings and conclusions in this paper are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention. The authors declare they have no actual or potential competing financial interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envpol.2019.06.074.

References

Antón, M., Cachorro, V.E., Vilaplana, J.M., Toledano, C., Krotkov, N.A., Arola, A., Serrano, A., de la Morena, B., 2010. Comparison of UV irradiances from Aura/ozone monitoring instrument (OMI) with Brewer measurements at El Arenosillo (Spain) – Part 1: analysis of parameter influence. Atmos. Chem. Phys. 10, 5979–5989.
Brenner, M., Hearing, V.J., 2008. The protective role of melanin against UV damage in human skin. Photochem. Photobiol. 84, 539–549.
CDC, 2018a. North America Land Data Assimilation System (NLDAS) Daily Sunlight (KJ/m2) (1979-2011). CDC WONDER Online Database.
CDC, 2018b. Skin Cancer.
Chang, N.-B., Gao, W., Gao, Z., 2012. Spatial statistical analyses of global trends of ultraviolet B fluxes in the continental United States. GIsci. Rem. Sens. 49, 735–754.
Cusworth, D.H., Mickley, L.J., Leibensperger, E.M., Iacono, M.J., 2017. Aerosol trends as a potential driver of regional climate in the central United States: evidence from observations. Atmos. Chem. Phys. 17, 13559–13572.
D'Orazio, J., Jarrett, S., Amaro-Ortiz, A., Scott, T., 2013. UV radiation and the skin. Int. J. Mol. Sci. 14, 12222–12248.
Dubayah, R., Rich, P.M., 1995. Topographic solar-radiation models for GIS. Int. J. Geogr. Inf. Syst. 9, 405–419.
Eck, T.F., Bhartia, P.K., Hwang, P.H., Stowe, L.L., 1987. Reflectivity of Earth's surface and clouds in ultraviolet from satellite observations. J. Geophys. Res. Atmos. 92, 4287–4296.
Eide, M.J., Weinstock, M.A., 2005. Association of UV index, latitude, and melanoma incidence in nonwhite populations - US surveillance, epidemiology, and end results (SEER) program, 1992 to 2001. Arch. Dermatol. 141, 477–481.
EPHT, 2017. National Environmental Public Health Tracking Network (EPHTN).
Fioletov, V.E., Kimlin, M.G., Krotkov, N., McArthur, L.J.B., Kerr, J.B., Wardle, D.I., Herman, J.R., Meltzer, R., Mathews, T.W., Kaurola, J., 2004. UV index climatology over the United States and Canada from ground-based and satellite estimates. J. Geophys. Res. Atmos. 109, 1–13.
Gao, W., Wang, X., Ustin, S.L., Gao, W., Davis, J., Olson, B., Janson, G., Slusser, J., 2007. Dependence of Erythemally Weighted UV Radiation on Geographical Parameters in the United States, Remote Sensing and Modeling of Ecosystems for Sustainability IV.
Hatfield, L.A., Hoffbeck, R.W., Alexander, B.H., Carlin, B.P., 2009. Spatiotemporal and spatial threshold models for relating UV exposures and skin cancer in the central United States. Comput. Stat. Data Anal. 53, 3001–3015.
Herman, J.R., Bhartia, P.K., Ziemke, J., Ahmad, Z., Larko, D., 1996. UV-B increases (1979–1992) from decreases in total ozone. Geophys. Res. Lett. 23, 2117–2120.
Herman, J.R., Celarier, E.A., 1997. Earth surface reflectivity climatology at 340 – 380 nm from TOMS data. J. Geophys. Res. Atmos. 102, 28003–28011.
Jari Hovila, A.A., Tamminen, Johanna, 2007. OMI/Aura Surface UV Irradiance 1-orbit L2 Swath 13x24 Km V003, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC). Accessed: April 2016, 10.5067/Aura/OMI/DATA2027.
Lemus-Deschamps, L., Makin, J.K., 2012. Fifty years of changes in UV Index and implications for skin cancer in Australia. Int. J. Biometeorol. 56, 727–735.
Levelt, P.F., Van den Oord, G.H.J., Dobber, M.R., Malkki, A., Huib, V., de Vries, J., Stammes, P., Lundell, J.O.V., Saari, H., 2006. The ozone monitoring instrument. IEEE Trans. Geosci. Remote Sens. 44, 1093–1101.
McKenzie, R., Smale, D., Kotkamp, M., 2004. Relationship between UVB and erythemally weighted radiation. Photochem. Photobiol. Sci. 3, 252–256.
McPeters, R.D., Frith, S., Labow, G.J., 2015. OMI total column ozone: extending the long-term data record. Atmos. Measur. Technol. 8, 4845–4850.
Moan, J., Porojnicu, A., Dahlback, A., Setlow, R., 2008. Addressing the health benefits

and risks, involving vitamin D or skin cancer, of increased sun exposure. Proc. Natl. Acad. Sci. U.S.A. 105, 668–673.

Myers, D.W.S., Marion, W., George, R., Anderberg, M., 2005. Broadband model performance for an updated national solar radiation database in the United States of America. In: Proc. Solar World Congress. International Solar Energy Society.

Narayanan, D.L., Saladi, R.N., Fox, J.L., 2010. Ultraviolet radiation and skin cancer. Int. J. Dermatol. 49, 978–986.

NCI, 2017. County Level UV Exposure Data for the Continental United States. National Cancer Institute. Geographic Information Systems and Science for Cancer Control.

NCI, 2018. Joinpoint regression program, 4.6.0.0. In: Statistical Research and Applications Branch. National Cancer Institute.

Norsang, G., Kocbach, L., Stamnes, J., Tsoja, W., Pincuo, N., 2011. Spatial distribution and temporal variation of solar UV radiation over the Tibetan Plateau. Appl. Phys. Res. 3.

Parding, K.M., Liepert, B.G., Hinkelman, L.M., Ackerman, T.P., Dagestad, K.-F., Olseth, J.A., 2016. Influence of synoptic weather patterns on solar irradiance variability in Northern Europe. J. Clim. 29, 4229–4250.

Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R., Vignola, F., 2002. A new operational model for satellite-derived irradiances: description and validation. Sol. Energy 73, 307–317.

Perez, R., Kivalov, S., Schlemmer, J., Hemker, K., Renne, D., Hoff, T.E., 2010. Validation of short and medium term operational solar radiation forecasts in the US. Sol. Energy 84, 2161–2172.

Pleasance, E.D., Cheetham, R.K., Stephens, P.J., McBride, D.J., Humphray, S.J., Greenman, C.D., Varela, I., Lin, M.-L., Ordonez, G.R., Bignell, G.R., Ye, K., Alipaz, J., Bauer, M.J., Beare, D., Butler, A., Carter, R.J., Chen, L., Cox, A.J., Edkins, S., Kokko-Gonzales, P.I., Gormley, N.A., Grocock, R.J., Haudenschild, C.D., Hims, M.M., James, T., Jia, M., Kingsbury, Z., Leroy, C., Marshall, J., Menzies, A., Mudie, L.J., Ning, Z., Royce, T., Schulz-Trieglaff, O.B., Spiridou, A., Stebbings, L.A., Szajkowski, L., Teague, J., Williamson, D., Chin, L., Ross, M.T., Campbell, P.J., Bentley, D.R., Futreal, P.A., Stratton, M.R., 2010. A comprehensive catalogue of somatic mutations from a human cancer genome. Nature 463, 191–196.

Richards, T.B., Johnson, C.J., Tatalovich, Z., Cockburn, M., Eide, M.J., Henry, K.A., Lai, S.M., Cherala, S.S., Huang, Y.J., Ajani, U.A., 2011. Association between cutaneous melanoma incidence rates among white US residents and county-level estimates of solar ultraviolet exposure. J. Am. Acad. Dermatol. 65, S50–S57.

Rigel, D.S., 2010. Epidemiology of melanoma. Semin. Cutan. Med. Surg. 29, 204–209.

Sengupta, M.H., Aron, Gueymard, Christian, Wilbert, Stefan, Renne, Dave, 2017. Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications, second ed. National Renewable Energy Laboratory (NREL) (NREL/TP-5D00-68886). www.nrel.gov/publications.

Stern, R., 2010. Prevalence of a history of skin cancer in 2007. Arch. Dermatol. 146, 279–282.

Tatalovich, Z., Wilson, J.P., Cockburn, M., 2006. A comparison of thiessen polygon, kriging and spline models of potential UV exposure. Cartogr. Geogr. Inf. Sci. 33, 217–231.

U.S. Cancer Statistics Working Group, 2018. U.S. Cancer Statistics Data Visualizations Tool, Based on November 2017 Submission Data (1999-2015). U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute.

Walls, A.C., Han, J., Li, T., Qureshi, A.A., 2013. Host Risk Factors, Ultraviolet index of residence, and incident malignant melanoma in situ among U.S. women and men. Am. J. Epidemiol. 177, 997–1005.

Wilcox, S., 2012. National Solar Radiation Database 1991–2010 Update. User's Manual National Renewable Energy Laboratory (NREL).

Zhang, H.X., Wang, J., Garcia, L.C., Zeng, J., Dennhardt, C., Liu, Y., Krotkov, N.A., 2019. Surface erythemal UV irradiance in the continental United States derived from ground-based and OMI observations: quality assessment, trend analysis and sampling issues. Atmos. Chem. Phys. 19, 2165–2181.