Downscaling of solar global irradiation in R

F. Antonanzas-Torres[a,*], F. J. Martínez-de-Pisón[a], J. Antonanzas[a], O. Perpiñán-Lamigueiro[b,c]

[a] *EDMANS Group, Department of Mechanical Engineering, University of La Rioja, Logroño, Spain.*
[b] *Electrical Engineering Department, EUITI-UPM, Ronda de Valencia 3, 28012 Madrid, Spain.*
[c] *Instituto de Energía Solar, Ciudad Universitaria s/n, Madrid, Spain*

Abstract

A methodology to downscale solar irradiation from satellite-derived databases is described using R software. Different packages such as `raster`, `parallel`, `solaR`, `gstat`, `sp` and `rasterVis` are considered in this study to improve solar resource estimation in areas with complex topography, in which the downscaling suppose a very useful tool to reduce inherent deviations of satellite-derived irradiation databases, which globally lack of high spatial resolution. A topographical analysis of horizon blocking and sky-view is developed with a digital elevation model to determine which fraction of hourly solar irradiation reaches the Earth surface. Eventually, *kriging with external drift* is applied for a better estimation of solar irradiation throughout the region analyzed. This methodology has been implemented as an example within La Rioja region in northern Spain, highlighting a mean absolute error 25.5% lower than with the original database.

Keywords: Solar irradiation, R, `raster`, `solaR`, digital elevation model, shade analysis, downscaling

1. Introduction

During the last few years the development of photovoltaic energy has flourished in developing countries with both multi-megawatt power plants and micro installations. However, the scarcity of long and reliable solar irradiation data from pyranometers in many of these countries makes necessary the estimation of solar irradiation from other meteorological variables or satellite photographies (Schulz et al., 2009). In the case of estimating irradiation from other measured meteorological variables, models need to be validated with nearby pyranometer registers, since these models lack of spatial generalization. Thus, in some regions without any nearby pyranometer these models become not an option and therefore, irradiation data must be obtained from satellite estimates. Although satellite-derived irradiation databases such as the NASA Surface meteorology and Solar Energy (SSE)[1], the National Renewable Energy Laboratory (NREL)[2], INPE[3], SODA[4] and the Climate Monitoring Satellite Application Facility (CM

*Corresponding author
Email address: `antonanzas.fernando@gmail.com` (F. Antonanzas-Torres)

[1] `http://maps.nrel.gov/SWERA`
[2] `http://www.nrel.gov/gis/solar.html`
[3] `http://www.inpe.br`
[4] `http://www.soda-is.com/eng/index.html`

SAF)[5] provide wide spatial coverage, only NASA and some CM SAF climate data sets present global coverage, while diminishing its spatial resolution (Table 1).

The spatial resolutions of satellite estimates are generally in the range of kilometers, averaging solar irradiation and omitting the impact of topography within each cell. As a result, intra-cell variations can be very significative in areas with local micro-climatic characteristics and also in areas with complex topography, which in many occasions coincide. Thus, under these events these irradiation data might not be of sufficient accuracy to project a photovoltaic installation. Perez et al. (1994) analyzed the spatial behavior of solar irradiation and concluded that the break-even distance from satellite estimates against pyranometers is in the order of 7 km and that variations are harsh to be estimated for distances further than 40 km. Antonanzas-Torres et al. (2013) rejected ordinary kriging as a spatial interpolation method of solar irradiation in Spain with stations separated more than 50 km in mountainous regions, as a result of its high spatial variability in these areas. The NASA-SSE and CM SAF SIS Climate Data Set (GHI) provide global coverage with resolutions of 1x1° and 0.25x0.25° (Table 1), which in most latitudes implies a grosser resolution than the previously mentioned 40-50 km. One of the alternatives to obtain higher spatial resolution of solar irradiation is the downscaling of satellite-estimates. Irradiation downscaling can be based on interpolation kriging techniques when pyranometers registers are available, implementing related-to-irradiation continuous variables such as elevation, sky-view-factor and other meteorological variables as external drifts (Alsamamra et al., 2009; Batlles et al., 2008). The downscaling is generally based on digital elevation models (DEM) with satellite-derived irradiation data to account for the effect of complex topography and has been previously applied in mountainous areas such as the Mont Blanc Massif (France) (Corripio, 2003) and Sierra Nevada (Spain) (Bosch et al., 2010; Ruiz-Arias et al., 2010) with image resolutions of 3.5x3.5 km. However, the NASA-SSE and CM SAF SIS Climate Data Set are based on much lower resolutions and suppose the unique irradiation datasets in numerous countries recently interested in solar energy. In this paper, a downscaling methodology of global solar irradiation is explained by means of R software and studied over La Rioja region (a very mountainous region in northern Spain). Data from the CM SAF with 0.03x0.03° resolution is considered and then downscaled to higher resolution (200x200m). On a second step, *kriging with external drift*, also denoted as *universal kriging*, is applied to interpolate data from 6 on-ground pyranometers within the region considering this downscaled CM SAF data as explanatory variable. Eventually, a downscaled map of annual global solar radiation is obtained throughout this region.

2. Data

The CM SAF was funded in 1992 as a joint venture of several European meteorological institutes, with the collaboration of the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) to retrieve, archive and distribute climate data to be used for climate monitoring and climate analysis (Posselt et al., 2012). Two kind of categories are provided: operational products and climate data. The operational products are built on data validated with on-ground stations and provided in near-to-present time and the climate data are long-term series to evaluate inter-annual variability. This study is built on hourly surface incoming solar radiation and direct irradiation climate data, denoted as SIS and SID by CM SAF respectively, for the year 2005. These climate data have been derived from Meteosat first generation satellites (Meteosat 2 to 7, 1982-2005) and validated using on-ground registers from the Baseline Surface

[5] http://www.cmsaf.eu

Radiation Network (BSRN) as a reference. The target accuracy of SIS and SID in hourly means is 15 W/m^2 (Posselt et al., 2011), providing a maximum spatial resolution of 0.03x0.03°. In the study, SIS and SID data are selected with spatial resolution of 0.03x0.03°. Data is freely accesible via FTP through the CM SAF website.

Hourly GHI registers from SOS Rioja[6] in 6 meteorological stations (shown in Figure 1 and Table 2) during 2005 serve as complementary measurements for the downscaling within the region studied.

These stations integrate First Class pyranometers (according to ISO 9060) with uncertainty of 5% in daily totals. These data have been filtered from spurious assuming in those cases the average between the previous and following hourly measurements.

The digital elevation model (DEM) is also freely obtained from the product MDT-200 of the ©Spanish Institute of Geography[7] with spatial resolution of 200x200m.

3. Method

This section describes the methodology proposed. Figure 2 displays the method diagram using red ellipses and lines for data sources, blue ellipses and lines for derived rasters (results), and black rectangles and lines for operations.

3.1. Irradiation decomposition

Initially, diffuse horizontal irradiation (DHI) is obtained throughout the difference between global horizontal irradiation (GHI) and beam horizontal irradiation (BHI) rasters, previously obtained from CM SAF. DHI and BHI are firstly disaggregated from the original gross resolution (0.03x0.03°) into the DEM resolution (200x200m), leading to remain similar values in disaggregated pixels to the original gross resolution pixel. On a second step, DHI can be divided in two components: isotropic diffuse irradiation (DHI_{iso}) and anisotropic diffuse irradiation (DHI_{ani}) according to the Hay & Mckay model (Hay and Mckay, 1985) (Equation 1). This model is based on the anisotropy index (k_1) defined as the relation between the beam irradiance ($B(0)$) and the extra-terrestrial irradiance ($B_0(0)$), as shown in Equation 2. High k_1 values are typical in clear sky atmospheres, while low k_1 values are frequent in overcasted and atmospheres with high aerosol density.

$$DHI = DHI_{iso} + DHI_{ani} \tag{1}$$

$$k_1 = \frac{B(0)}{B_0(0)} \tag{2}$$

The DHI_{iso} accounts for the incoming diffuse irradiation portion from an isotropic sky, being more significant in very cloudy days (Equation 3).

$$DHI_{iso} = DHI \cdot (1 - k_1) \tag{3}$$

DHI_{ani}, also denoted as circumsolar diffuse irradiation, considers the incoming portion from the circumsolar disk and can be analyzed as beam irradiation (Perpiñán-Lamigueiro, 2013) (Equation 4).

$$DHI_{ani} = DHI \cdot k_1 \tag{4}$$

[6] http://www.larioja.org/npRioja/default/defaultpage.jsp?idtab=442821
[7] http://www.ign.es

3.2. Sky view factor and horizon blocking

The topographical analysis is performed accounting for the visible sky sphere (sky view) and also the horizon blocking. The DHI_{iso} is directly dependent on the sky-view factor (SVF), which computes the proportion of visible sky related to a flat horizon. The sky-view factor has been previously considered in other irradiation assessments (Corripio, 2003; Ruiz-Arias et al., 2010). The sky view factor is calculated in each DEM pixel considering 72 vectors (separated 5° each) and evaluating the maximum horizon angle (θ_{hor}) along 20km in each vector (Equation 5). The θ_{hor} stands for the maximum angle between the altitude of a location and the elevation of the group of points along each vector, related to a horizontal plane on the location. Locations without horizon blocking achieve SVF close to 1, which stands for a whole visible sky semi-sphere.

$$SVF = 1 - \int_0^{2\pi} sin^2\theta_{hor} d\theta \tag{5}$$

Eventually, the downscaled DHI_{iso} ($DHI_{iso,down}$) is computed with Equation 6.

$$DHI_{iso,down} = DHI_{iso} \cdot SVF \tag{6}$$

The horizon blocking is analyzed evaluating the solar geometry in 15 minutes samples, particularly the solar elevation (γ_s) and the solar azimuth (ψ_s). Secondly, the mean hourly γ_s and ψ_s (from those 15 minutes rasters) are calculated and then disaggregated as previously explained for DHI and BHI rasters. The decision of solving the solar geometry with low resolution rasters permits reducing significantly computational time without penalizing results. The θ_{hor} corresponding to each ψ_s is compared with the θ_{zs}. As a consequence, if the θ_{zs} is greater than the θ_{hor} implies horizon blocking on the surface analyzed and therefore, BHI and DHI_{ani} blocked. Eventually, the sum of $DHI_{ani,down}$, $DHI_{iso,down}$ and $BHI_{iso,donw}$ constitutes the downscaled global horizontal irradiation GHI_{down}.

3.3. Post-processing: kriging with external drift

The fact that the afore-explained downscaling accounts for the irradiation loss due to the horizon blocking and sky-view factor leads to introduce a trend in estimates (lowering them) when compared with the original data (gross resolution data). However, satellite-derived irradiation data is implicitly considering shades, as a consequence of the lower albedo registered in these zones, although it is latter averaged throughout the pixel. GHI_{down} can be considered as useful bias of the behavior of solar irradiation within the region studied. The *universal kriging* or *kriging with external drift* (KED) includes the information from exhaustively-sampled explanatory variables in the interpolation. As a result, GHI_{down} is considered the explanatory variable to interpolate measured irradiation data from some on-ground calibrated pyranometers which is denoted as *post-processing*. GHI_{down} is correlated with the DEM as a consequence of the high influence of the horizon blocking with topography and therefore estimations can be derived separating the deterministic ($\hat{m}(\mathbf{s}_\theta)$) and stochastic components ($\hat{\epsilon}(\mathbf{s}_\theta)$) (Equations 7 and 8).

$$\hat{z}(\mathbf{s}_\theta) = \hat{m}(\mathbf{s}_\theta) + \hat{\epsilon}(\mathbf{s}_\theta) \tag{7}$$

$$\hat{z}(\mathbf{s}_\theta) = \sum_{k=0}^{p} \hat{\beta}_k q_k(\mathbf{s}_\theta) + \sum_{i=1}^{n} \lambda_i \epsilon(\mathbf{s}_i) \tag{8}$$

where $\hat{\beta}_k$ are the estimated coefficients of the deterministic model, $q_k(\mathbf{s}_\theta)$ are the auxiliary predictors obtained from the fitted values of the explanatory variable at the new location, λ_i are

the kriging weights determined by the spatial dependence structure of the residual, and $\epsilon(\mathbf{s}_i)$ are the residual at location $\mathbf{s}_i$ (Antonanzas-Torres et al., 2013). The semivariogram is a function defined as Equation 9 based on a constant variance of ϵ and also on the assumption that spatial correlation of ϵ depends on the distance amongst instances ($\mathbf{h}$) instead of their position (Pebesma, 2004).

$$\gamma(\mathbf{h}) = \frac{1}{2}\mathrm{E}(\epsilon(\mathbf{s}) - \epsilon(\mathbf{s}+\mathbf{h}))^2 \tag{9}$$

Given that the afore-explained sample variogram only collates estimates from observed points, a fitting model of this variogram is generally considered to extrapolate the spatial behavior of observed points to the area studied. In the literature different variogram functions are commonly defined such as the exponential, gaussian or spherical models. In this line, different parameters such as the sill, range and nugget of the model must be adjusted to best fit the sample variogram (Hengl, 2009). The nugget effect, generally associated to intrinsec micro-variability and measurement error, models the discontinuity of the variogram at the origin. It must be highlighted that when the nugget effect is recorded, kriging differs from a regular interpolation and as a result estimates are different from measured values. The variogram model of solar horizontal irradiation has been evaluated in Spain, concluding that a pure nugget fitting behaves best, implying no spatial auto-correlation on residuals (Antonanzas-Torres et al., 2013).

4. Implementation

The methodology proposed is applied in La Rioja region (north of Spain). Figure 3 shows the correspondent annual global horizontal irradiation from CM SAF with resolution 0.03x0.03°.

4.1. Packages

The downscaling described in this paper has been implemented using the free software environment R (R Development Core Team, 2013) and several contributed packages:

- `raster` (Hijmans and van Etten, 2013) for spatial data manipulation and analysis.
- `solaR` (Perpiñán-Lamigueiro, 2012) for the solar geometry.
- `gstat` (Pebesma and Graeler, 2013) and `sp` (Pebesma et al., 2013) for the geostatistical analysis.
- `parallel` for multi-core parallelization.
- `rasterVis` (Perpiñán-Lamiguiero and Hijmans, 2013) for spatial data visualization methods.

```
R> library(sp)
R> library(raster)
R> rasterOptions(todisk=FALSE)
R> rasterOptions(chunksize = 1e+06, maxmemory = 1e+07)
R> library(maptools)
R> library(gstat)
R> library(lattice)
R> library(latticeExtra)
R> library(rasterVis)
R> library(solaR)
R> library(parallel)
```

4.2. Data

Satellite data can be freely downloaded previous registration from CM SAF[8] in the *data access* space selecting *web user interface* and *climate data sets* and then choosing hourly climate data sets named *SIS* (Global Horizontal Irradiation)) and *SID* (Beam Horizontal Irradiation) for 2005. Both rasters are projected to the UTM projection for compatibility with the Digital Elevation Model.

```
R> projUTM <- CRS('+proj=utm␣+zone=30')
R> projLonLat <- CRS('␣+proj=longlat␣+ellps=WGS84')

R> listFich <- dir(pattern='SIShm2005')
R> stackSIS <- stack(listFich)
R> stackSIS <- projectRaster(stackSIS,crs=projUTM)

R> listFich <- dir(pattern='SIDhm2005')
R> stackSID <- stack(listFich)
R> stackSID <- projectRaster(stackSID, crs=projUTM)
```

We compute the annual global irradiation which will be used as the reference for future steps.

```
R> SISa2005 <- calc(stackSIS, sum, na.rm=TRUE)
```

The Spanish Digital Elevation Model is obtained previous registration from the ©Spanish Institute of Geography[9] in the *free download of digital geographic information for non-commercial use* space, and then crop to the region analyzed (La Rioja). As stated above, this DEM uses the UTM projection.

```
R> elevSpain <- raster('elevSpain.grd')
R> elev <- crop(elevSpain, extent(479600, 616200, 4639600, 4728400))
R> names(elev)<-'elev'
```

4.3. Sun geometry

The first step is to compute the sun angles (height and azimuth) and the extraterrestial solar irradiation for every cell of the CMSAF rasters. The function `calcSol` from the `solaR` package calculates the daily and intradaily sun geometry. This function and the `overlay` method from the `raster` package produce three multilayer `Raster` objects with the sun geometry needed for the next steps. For the sake of brevity we show only the procedure for the extraterrestial solar irradiation. The calculation of sun geometry is performed with the resolution of CM SAF.

As a first step, we define a function to extract the hour for aggregation, choose the annual irradiation raster as reference, and define a raster with longitude and latitude coordinates.

```
R> hour <- function(tt)as.POSIXct(trunc(tt, 'hours'))

R> r <- SISa2005

R> latlon <- stack(init(r, v='y'), init(r, v='x'))
R> names(latlon) <- c('lat', 'lon')
```

[8] www.cmsaf.eu
[9] http://www.ign.es

The extraterrestrial irradiation is calculated with 5-min samples. Every point is a column of the `data.frame` `locs`. Its columns are traversed with `lapply` so for every point of the `Raster` object a time series of extraterrestial solar irradiation is computed. The result, `B05min`, is a `RasterBrick` object with a layer for each element of the time index `BTi`, which is aggregated to an hourly raster with `zApply` and transformed to the UTM projection.

```
R> BTi <- seq(as.POSIXct('2005-01-01␣00:00:00'),
+ as.POSIXct('2005-12-31␣23:55:00'), by='5␣min')

R> B05min <- overlay(latlon, fun=function(lat, lon){
+ locs <- as.data.frame(rbind(lat, lon))
+ b <- lapply(locs, function(p){
+
+ hh <- local2Solar(BTi, p[2])
+ sol <- calcSol(p[1], BTi=hh)
+ Bo0 <- as.data.frameI(sol)$Bo0
+ Bo0 })
+ res <- do.call(rbind, b)})

R> B05min <- setZ(B05min, BTi)
R> names(B05min) <- as.character(BTi)

R> B0h <- zApply(B05min, by=hour, fun=mean)
R> projectRaster(B0h,crs=projUTM)
```

4.4. Irradiation components

The CMSAF rasters must be transformed to the higher resolution of the DEM (UTM 200mx200m). As a consequence of the different pixel geometry between DEM (square) and irradiation rasters (rectangle) the process is performed in two steps.

The first step increases the spatial resolution of the irradiation rasters to a similar and also larger pixel size than the DEM with `disaggregate`, where `sf` is the scale factor. The second step post-processes the previous step by means of a bilineal interpolation which resamples the raster layer and achieves same DEM resolution (`resample`). This two-step disaggregation avoids loosing original values of the gross resolution raster that would be directly interpolated with the one-step disaggregation.

```
R> sf <- res(stackSID)/res(elev)

R> SIDd <- disaggregate(stackSID, sf)
R> SIDdr <- resample(SIDd, elev)

R> SISd <- disaggregate(stackSIS, sf)
R> SISdr <- resample(SISd, elev)
```

On the other hand, the diffuse irradiation is obtained from the global and beam irradiation rasters. The two components of the diffuse irradiation, isotropic and anisotropic, can be separated with the anisotropy index, computed as the ratio between beam and extraterrestial irradiation.

```
R> Difdr <- SISdr - SIDdr

R> B0hd <- disaggregate(B0h, sf)
R> B0hdr <- resample(B0hd, elev)
```

```
R> k1 <- SIDdr/B0hdr

R> Difiso <- (1-k1) * Difdr
R> Difani <- k1 * Difdr
```

4.5. Sky view factor and horizon blocking

4.5.1. Horizon angle

The maximum horizon angle required for the horizon blocking analysis and also to derive the SVF is obtained with the next code.

The `alfa` vector is visited with `mclapply` (using parallel computing). For each direction angle (elements of this vector) the maximum horizon angle is calculated for a set of points across that direction from each of the locations defined in `xyelev` (derived from the DEM raster and transformed in the matrix `locs` visited by rows)

```
R> xyelev <- stack(init(elev, v='x'),
+ init(elev, v='y'),
+ elev)
R> names(xyelev) <- c('x', 'y','elev')


R> inc <- pi/36
R> alfa <- seq(-0.5*pi,(1.5*pi-inc), inc)

R> locs <- as.matrix(xyelev)
```

Separations between the origin locations and points along each direction are defined by `resD`, the maximum resolution of the DEM, `d`, maximum distance to visit, and consequently in the vector `seps`.

```
R> resD <- max(res(elev))

R> d <- 20000
R> seps <- seq(resD, d, by=resD)
```

The elevation (`z1`) of each point in `xyelev` is converted into the horizon angle: the largest of these angles is the horizon angle for that direction. The result of each `apply` step is a matrix which is used to fill in a RasterLayer (`r`). The result of `mclapply` is a list, `hor`, of `RasterLayer` which can be converted into a `RasterStack` with `stack`. Each layer of this `RasterStack` corresponds to a different direction.

```
R> hor <- mclapply(alfa, function(ang){
+ h <- apply(locs, 1, function(p){
+ x1 <- p[1]+cos(ang)*seps
+ y1 <- p[2]+sin(ang)*seps
+ p1 <- cbind(x1,y1)
+ z1 <- elevSpain[cellFromXY(elevSpain,p1)]
+ hor <- r2d(atan2(z1-p[3], seps))
+ maxHor <- max(hor[which.max(hor)], 0)
+ })
+ r <- raster(elev)
+ r[] <- matrix(h, nrow=nrow(r), byrow=TRUE)
+ r}, mc.cores=8)
R> horizon <- stack(hor)
```

This operation is very time-demanding as a result of working with high resolution files. Computational time can be lowered by increasing sampling space (200m), sectorial angle (5°) or reducing the maximum distance (20km).

4.5.2. Horizon blocking

The horizon blocking is analyzed evaluating the solar geometry in 15 minutes samples, particularly the solar elevation and azimuth angles throughout the original irradiation raster. Secondly, the hourly averages ared calculated, disaggregated and post-processed as previously explained for the irradiation rasters. The decision of solving the solar geometry with low resolution rasters allows for the significant reduction of computational time without penalizing results.

First, the azimuth raster is cut in different classes according to the alfa vector (directions). The values of the `horizon` raster corresponding to each angle class are extracted using `stackSelect`.

```
R> idxAngle <- cut(AzShr, breaks=r2d(alfa))
R> AngAlt <- stackSelect(horizon, idxAngle)
```

The number of layers of `AngAlt` is the same as `idxAngle` and, therefore can be used for comparison with with the solar height angle, `AlShr`. If `AngAlt` is greater, there is horizon blocking (`dilogical=0`).

```
R> dilogical <- ((AngAlt-AlShr) < 0)
```

With this binary raster, beam irradiation and diffuse anisotropic irradiation can be corrected with horizon blocking.

```
R> Dirh <- SIDdr * dilogical
R> Difani <- Difani * dilogical
```

4.5.3. Sky view factor

The Sky View Factor can be easily computed from the `horizon` object with the equation proposed above. This factor corrects the isotropic component of the diffuse irradiation.

```
R> SVFRuizArias <- calc(horizon, function(x) sin(d2r(x))^2)
R> SVF <- 1 - mean(SVFRuizArias)

R> Difiso <- Difiso * SVF
```

Finally, the global irradiation is the sum of the three corrected components, beam and anisotropic diffuse irradiation including horizon blocking, and isotropic diffuse irradiation with the sky view factor.

```
R> GHIh <- Difanis + Difiso + Dirh
R> GHI2005a <- calc(GHIh, fun=sum)
```

4.6. Kriging with external drift

The downscaled irradiation rasters can be improved using kriging with external drift. Irradiation data from on-ground meteorological stations is interpolated with the downscaled irradiation raster as explanatory variable. To define the variogram here we use the results previously published in (Antonanzas-Torres et al., 2013).

```
R> load('Stations.RData')
R> UTM <- SpatialPointsDataFrame(Stations[,c(2,3)], Stations[,-c(2,3)],
+ proj4string=CRS('+proj=utm␣+zone=30␣+ellps=WGS84'))


R> vgmCMSAF <- variogram(GHImed ~ GHIcmsaf, UTM)
R> fitvgmCMSAF <- fit.variogram(vgmCMSAF, vgm(model='Nug'))

R> gModel <- gstat(NULL, id='G0yKrig',
+ formula= GHImed ~ GHIcmsaf,
+ locations=UTM, model=fitvgmCMSAF)

R> names(GHI2005a) <- 'GHIcmsaf'
R> G0yKrig <- interpolate(GHI2005a, gModel, xyOnly=FALSE)
```

4.7. Analysis of the results

Figures 3 shows the annual GHI as per CM SAF with the gross resolution analyzed (0.03x0.03°) and Figures 4 and 5 show the downscaled maps (200x200m) without and with the kriging with external drift.

4.7.1. Model performance

In order to evaluate the performance of the methodology proposed, relative differences evaluated with station measurements are shown in Figure 6. As it can be deduced from this Figure, relative differences are reduced in *downscaling with KED* when compared with CM SAF or *downscaling without KED*, remaining in a $\pm$ 15%.

The mean absolute error (MAE) and root mean square error ($RMSE$), described in Equations 10 and 11, are used as indicators of model performance.

$$MAE = \frac{\sum_{i=1}^{n} |x_{est} - x_{meas}|}{n} \tag{10}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (x_{est} - x_{meas})^2}{n}} \tag{11}$$

where n is number of stations and x_{est} and x_{meas} the annual estimated and measured irradiation, respectively.

Table 3 shows the MAE and RMSE obtained with CM SAF and with the methodology proposed before and after the *KED*. The *KED* leads to a meaningful improvement in estimates, 25.5% and 27.4% lower MAE and RMSE when compared with CM SAF.

The higher MAE recorded in station locations within CM SAF and *downscaling without KED* is also explained in the irradiation maps shown in Figures 3 and 4. The GHI_{annual} is excesively lowered in certain regions of the area studied with *downscaling without KED* compared to $GHI_{down,ked}$, which is also shown in Figure 6 .

4.7.2. Zonal variability

The intrapixel variability due to the downscaling procedure is indicative of the importance of the topography as attenuator of solar irrradiation. As a result, this zonal variability will be higher in pixels with complex topography and the downscaling more useful. Figure 7 shows the relative difference between the downscaling with KED and CM SAF. As it might be deduced, CM SAF over-estimates GHI in this region in the range of 11-22%. Figures 8 and 9 display

the standard deviations of the downscaled maps within each cell of the original CM SAF raster (0.03x0.03°). The `zonal` function from `raster` library permits this calculation, explaining the intrinsecal variability of solar radiation within gross resolution pixels. Consequently, in those pixels with higher standard deviation a higher variability will be present. Figure 9 shows how the KED method smooth the deviation within pixels and also the range of solar irradiation in the region (Figures 4 and 5).

5. Concluding comments

A methodology to downscale solar irradiation is described and presented using R software. This applied methodology is useful to increase accuracy and spatial resolution of gross resolution satellite-estimates of solar irradiation.

It has been proved that areas with complex topography present higher differences with original gross resolution data as a consequence of horizon blocking and lower sky-view factors and therefore, downscaling is highly recommended in these areas.

Kriging with external drift with `gstat` package has been proved very useful in the downscaling of solar irradiation when some on-ground registers are available and a explanatory variable is provided.

This methodology has been implemented as an example within La Rioja region in northern Spain, highlighting the reduction of 25.5% and 27.4% in MAE and RMSE when compared with the original gross resolution database. The high repeatibility of this methodology and the reduction in errors obtained might be also very useful in the downscaling of other meteorological variables apart from solar irradiation.

Session information

The results discussed in this paper were obtained in a session with these characteristics:

- R version 2.15.2 (2012-10-26), `x86_64-apple-darwin9.8.0`
- Locale: `es_ES.UTF-8/es_ES.UTF-8/es_ES.UTF-8/C/es_ES.UTF-8/es_ES.UTF-8`
- Base packages: `base`, `datasets`, `graphics`, `grDevices`, `grid`, `methods`, `parallel`, `stats`,`utils`
- Other packages: `foreign` 0.8-51, `gstat` 1.0-16, `hexbin` 1.26.0, `lattice` 0.20-15, `latticeExtra` 0.6-19, `maptools` 0.8-14, `raster` 2.1-16, `rasterVis` 0.20-01, `RColorBrewer` 1.0-5, `rgdal` 0.8-01, `solaR` 0.33, sp 1.0-8, `zoo` 1.7-9
- Loaded via a namespace (and not attached): `intervals` 0.14.0, `spacetime` 1.0-4, `tools` 2.15.2, `xts` 0.9-3

Acknowledgements

We are indebted to the University of La Rioja (fellowship FPI 2012) and the Research Institute of La Rioja (*IER*) for funding parts of this research.

References

Alsamamra, H., Ruiz-Arias, J. A., Pozo-Vázquez, D., Tovar-Pescador, J., 2009. A comparative study of ordinary and residual kriging techniques for mapping global solar radiation over southern spain. Agricultural and Forest Meteorology 149 (8), 1343 – 1357.

Antonanzas-Torres, F., Cañizares, F., Perpiñán, O., 2013. Comparative assessment of global irradiation from a satellite estimate model (cm saf) and on-ground measurements (siar): a spanish case study. Renewable and Sustainable Energy Reviews 21, 248–261.

Batlles, F., Bosch, J., Tovar-Pescador, J., Martínez-Durbán, M., Ortega, R., Miralles, I., 2008. Determination of atmospheric parameters to estimate global radiation in areas of complex topography: Generation of global irradiation map. Energy Conversion and Management 49 (2), 336 – 345.

Bosch, J., Batlles, F., Zarzalejo, L., López, G., 2010. Solar resources estimation combining digital terrain models and satellite images techniques. Renewable Energy 35 (12), 2853 – 2861.

Corripio, J., 2003. Vectorial algebra algorithms for calculating terrain parameters from dems and solar radiation modelling in mountainous terrain. International Journal of Geographical Information Science 17, 1–23.

Hay, E., Mckay, D., 1985. Estimating solar irradiance on inclined surfaces: a review and assessment of methodologies. International Journal of Solar Energy 3, 203–240.

Hengl, T., 2009. A Practical Guide to Geostatistical Mapping.
URL http://spatial-analyst.net/book/

Hijmans, R. J., van Etten, J., 2013. `raster`: Geographic Data Analysis and Modeling. R package version 2.1-25.
URL http://CRAN.R-project.org/package=raster

Pebesma, E., Bivand, R., Rowlingson, B., Gomez-Rubio, V., 2013. `sp`: Classes and Methods for Spatial Data. R package version 1.0-9.
URL http://CRAN.R-project.org/package=sp

Pebesma, E., Graeler, B., 2013. `gstat`: Spatial and Spatio-Temporal Geostatistical Modelling, Prediction and Simulation. R package version 1.0-16.
URL http://CRAN.R-project.org/package=gstat

Pebesma, E. J., 2004. Multivariable geostatistics in s: the `gstat` package. Computers & Geosciences 30, 683–691.

Perez, R., Seals, R., Stewart, R., Zelenka, A., Estrada-Cajigal, V., 1994. Using satellite-derived insolation data for the site/time specific simulation of solar energy systems. Solar Energy 53 (6), 491 – 495.

Perpiñán-Lamigueiro, O., 8 2012. `solaR`: Solar radiation and photovoltaic systems with r. Journal of Statistical Software 50 (9), 1–32.
URL http://www.jstatsoft.org/v50/i09

Perpiñán-Lamigueiro, O., 2013. Energía Solar Fotovoltaica.
URL http://procomun.wordpress.com/documentos/libroesf/

Perpiñán-Lamiguiero, O., Hijmans, R. J., 2013. `rasterVis`: Visualization Methods for the Raster Package. R package version 0.20-07.
URL http://CRAN.R-project.org/package=rasterVis

Posselt, R., Mueller, R., Stöckli, R., Trentmann, J., 2012. Remote sensing of solar surface radiation for climate monitoring — the cm-saf retrieval in international comparison. Remote Sensing of Environment 118 (0), 186 – 198.

Posselt, R., Muller, R., Trentmann, J., Stockli, R., 2011. Meteosat (mviri) solar surface irradiance and effective cloud albedo climate data sets. the cm saf validation report. Tech. rep., The EUMETSAT Network of Satellite Application Facilities.

R Development Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.
URL http://www.R-project.org

Ruiz-Arias, J. A., Cebecauer, T., Tovar-Pescador, J., Šúri, M., 2010. Spatial disaggregation of satellite-derived irradiance using a high-resolution digital elevation model. Solar Energy 84 (9), 1644 – 1657.

Schulz, J., Albert, P., Behr, H.-D., Caprion, D., Deneke, H., Dewitte, S., Dürr, B., Fuchs, P., Gratzki, A., Hechler, P., Hollmann, R., Johnston, S., Karlsson, K.-G., Manninen, T., Müller, R., Reuter, M., Riihelä, A., Roebeling, R., Selbach, N., Tetzlaff, A., Thomas, W., Werscheck, M., Wolters, E., Zelenka, A., 2009. Operational climate monitoring from space: the eumetsat satellite application facility on climate monitoring (cm-saf). Atmospheric Chemistry and Physics 9 (5), 1687–1709.

Tables and Figures

Database	Product	Spatial coverage	Spatial resolution	Temporal coverage	Temporal resolution
CM SAF	SIS Climate Data Set (GHI)	Global	0.25x0.25°	1982-2009	Daily means
CM SAF	SIS Climate Data Set (GHI)	70S-70N, 70W-70E	0.03x0.03°	1983-2005	Hourly means
CM SAF	SID Climate Data Set (BHI)	70S-70N, 70W-70E	0.03x0.03°	1983-2005	Hourly means
SODA	Helioclim 3 v2 and v3 (GHI)	66S-66N,66W-66E	5km	2005	15 minutes
SODA	Helioclim 3 v2 and v3 (GHI)	66S-66N,66W-66E	5km	2005	15 minutes
NREL	GHI Moderate resolution	Central and South America, Africa, India, East Asia	40x40km	1985-1991	Monthly means of daily GHI
NASA	SSE	Global	1x1°	1983-2005	Daily means

Table 1: Summary of solar irradiation databases

#	Name	Net.	Lat.(º)	Long.(º)	Alt.	GHI_a
1	Ezcaray	SOS	42.33	-3.00	1000	1479
2	Logroño	SOS	42.45	-2.74	408	1504
3	Moncalvillo	SOS	42.32	-2.61	1495	1329
4	San Roman	SOS	42.23	-2.45	1094	1504
5	Ventrosa	SOS	42.17	-2.84	1565	1277
6	Yerga	SOS	42.14	-1.97	1235	1448

Table 2: Summary of the meteorological stations selected.

	CM SAF	without KED	with KED
MAE	101.35	175.63	75.54
RMSE	118.65	196.53	86.18

Table 3: Summary of errors obtained in kWh/m^2.

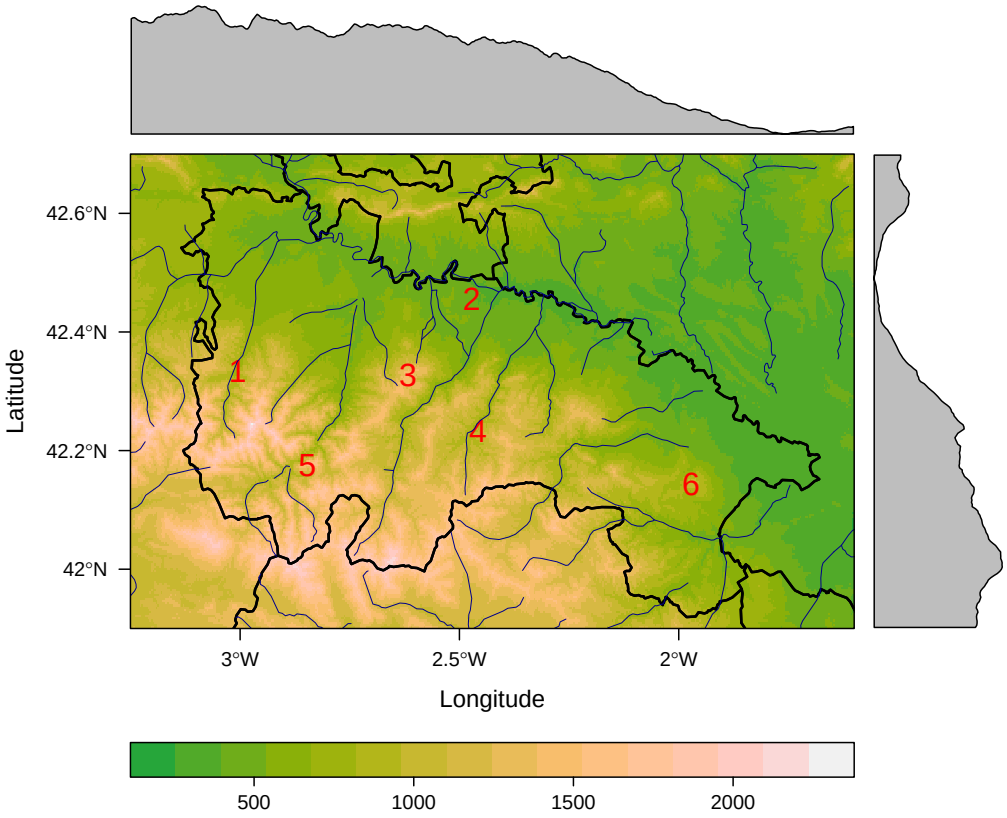

Figure 1: Region analyzed and meteorological stations considered

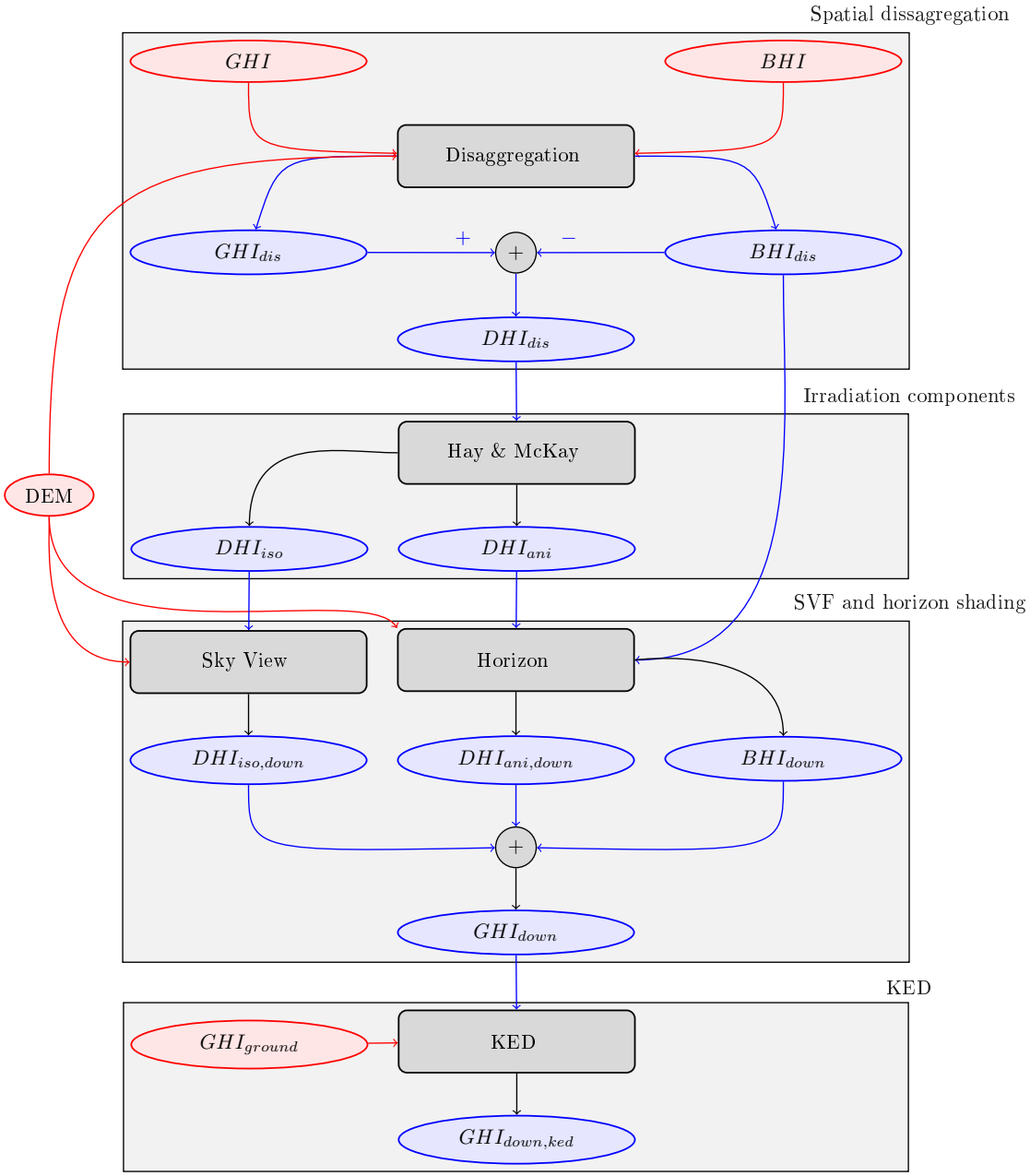

Figure 2: Methodology of downscaling: this figure uses red ellipses and lines for data sources, blue ellipses and lines for derived rasters (results), and black rectangles and lines for operations.

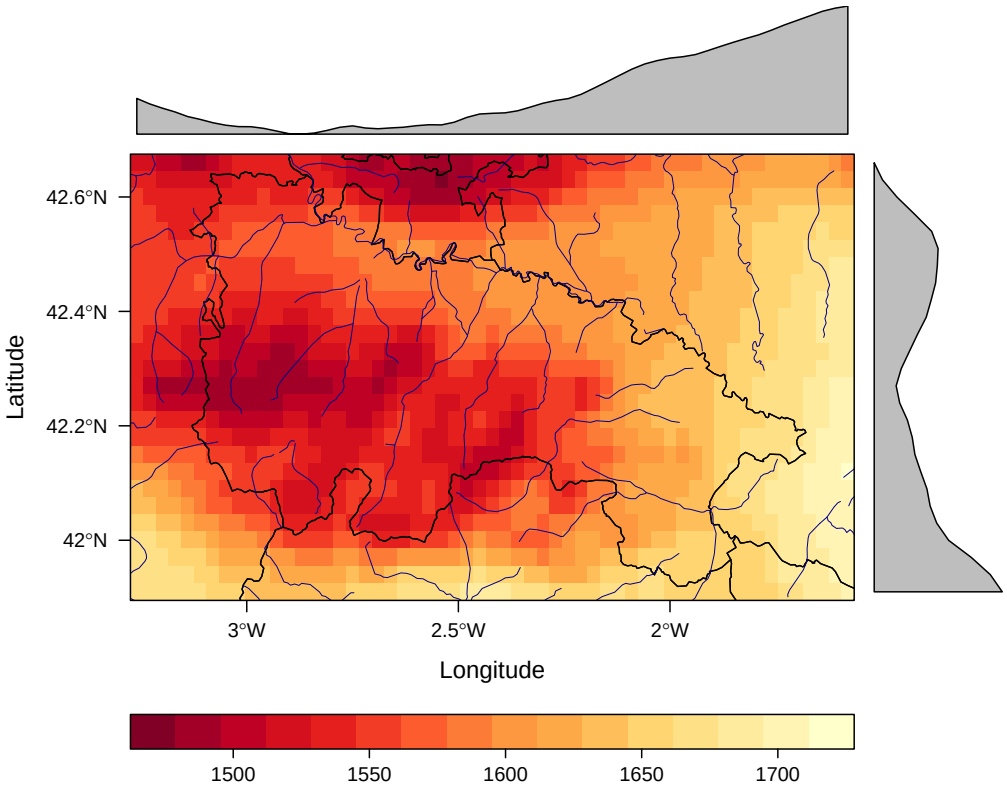

Figure 3: Annual GHI of 2005 from CM SAF estimates (0.03x0.03°) in La Rioja

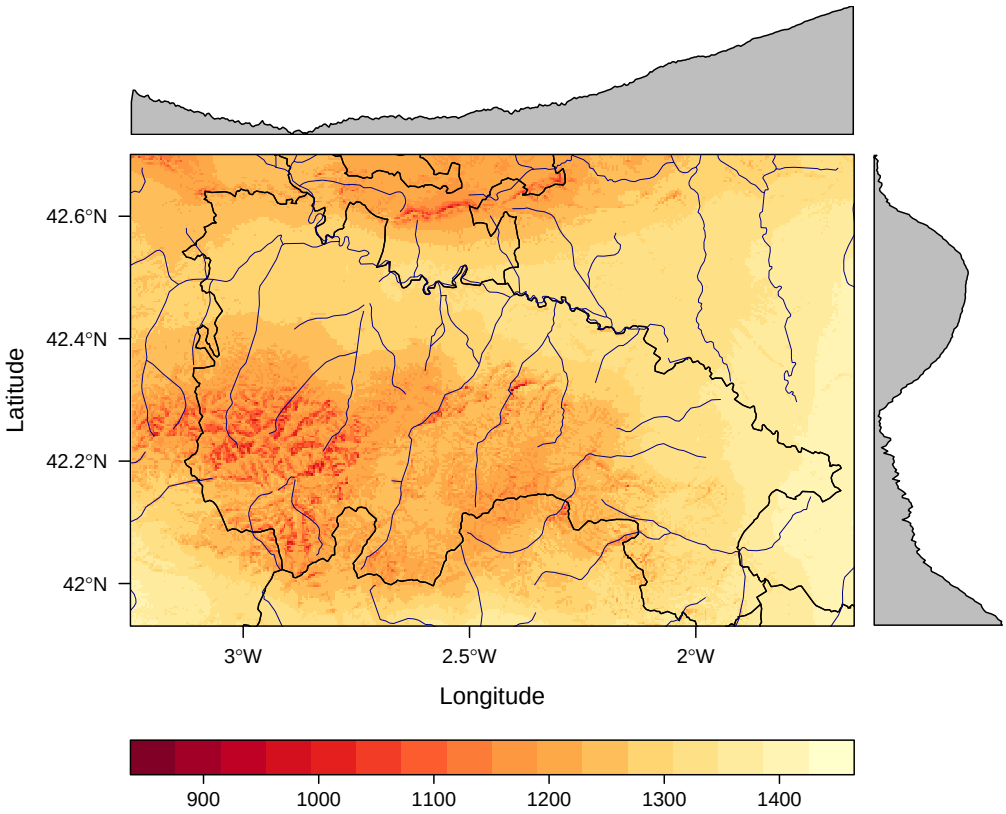

Figure 4: Annual GHI of 2005 downscaled without KED (0.03x0.03°) in La Rioja

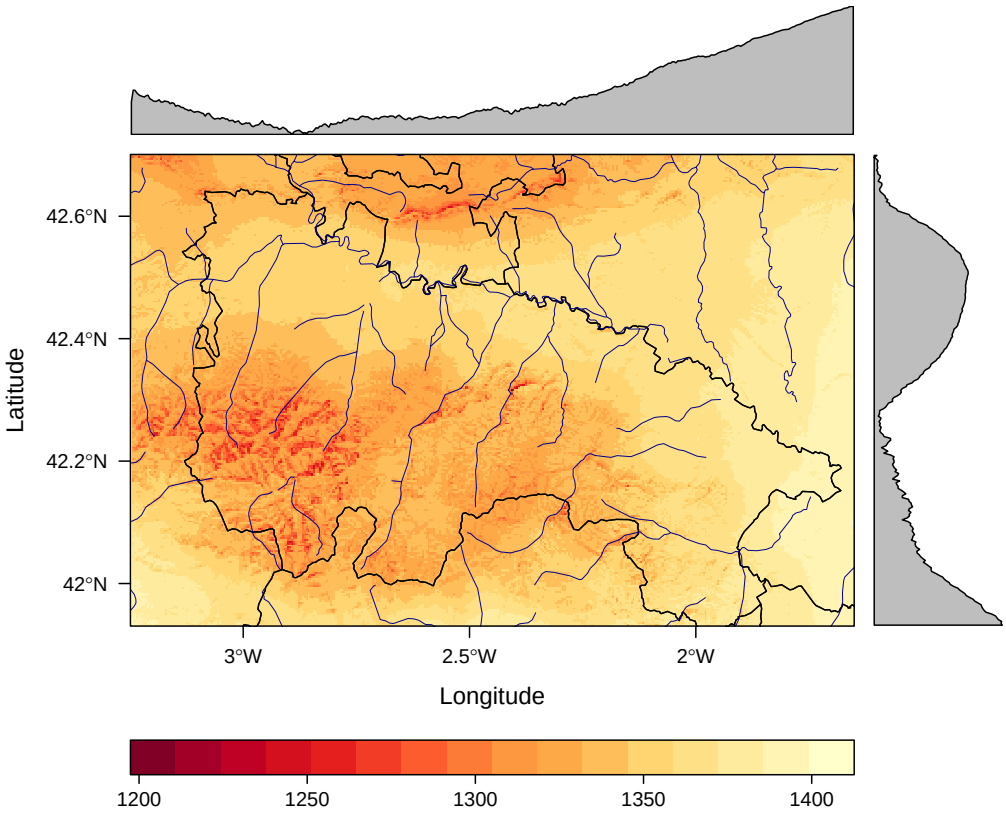

Figure 5: Annual GHI of 2005 downscaled with KED (0.03x0.03°) in La Rioja

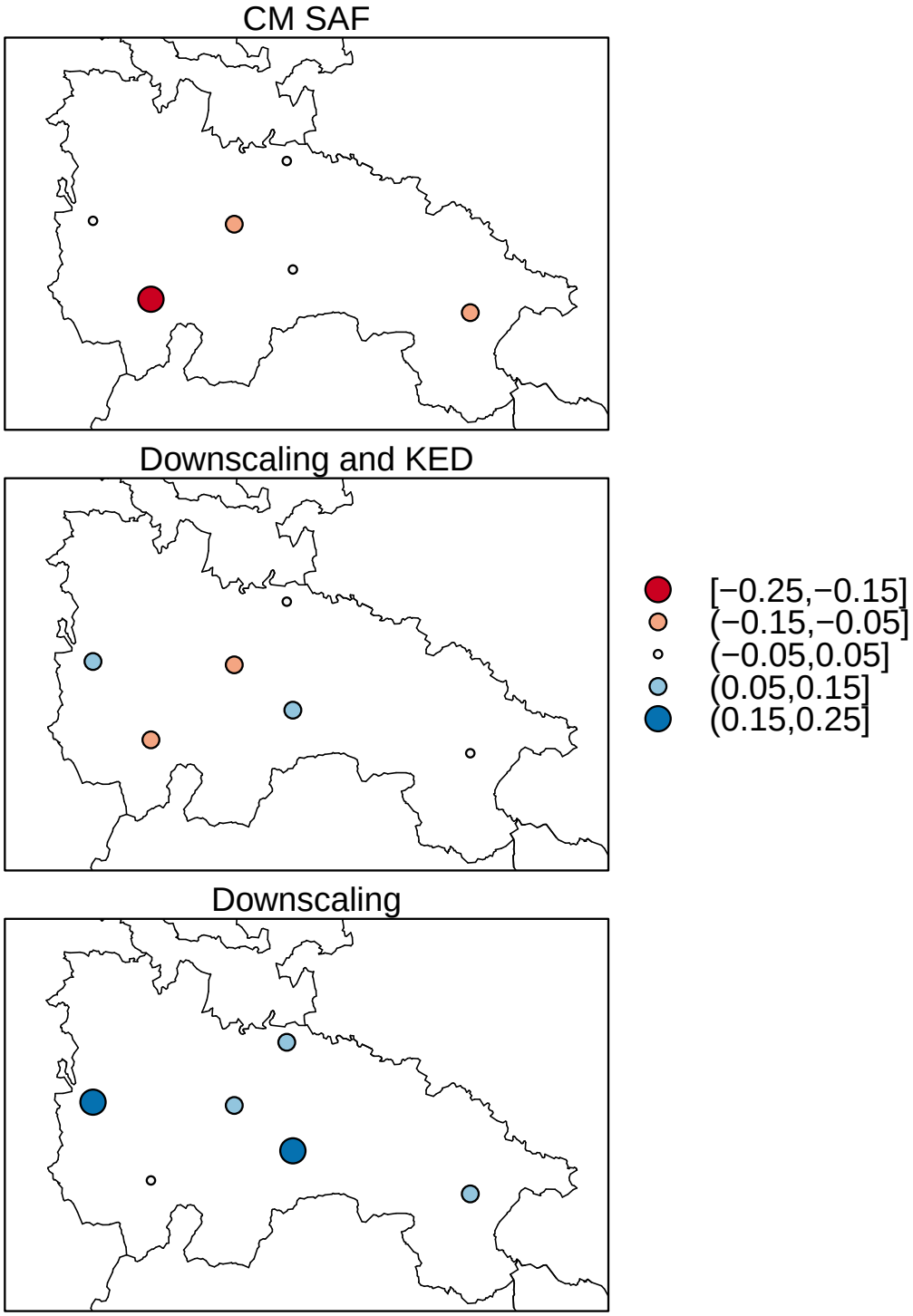

Figure 6: Annual relative differences evaluated with station measurements.

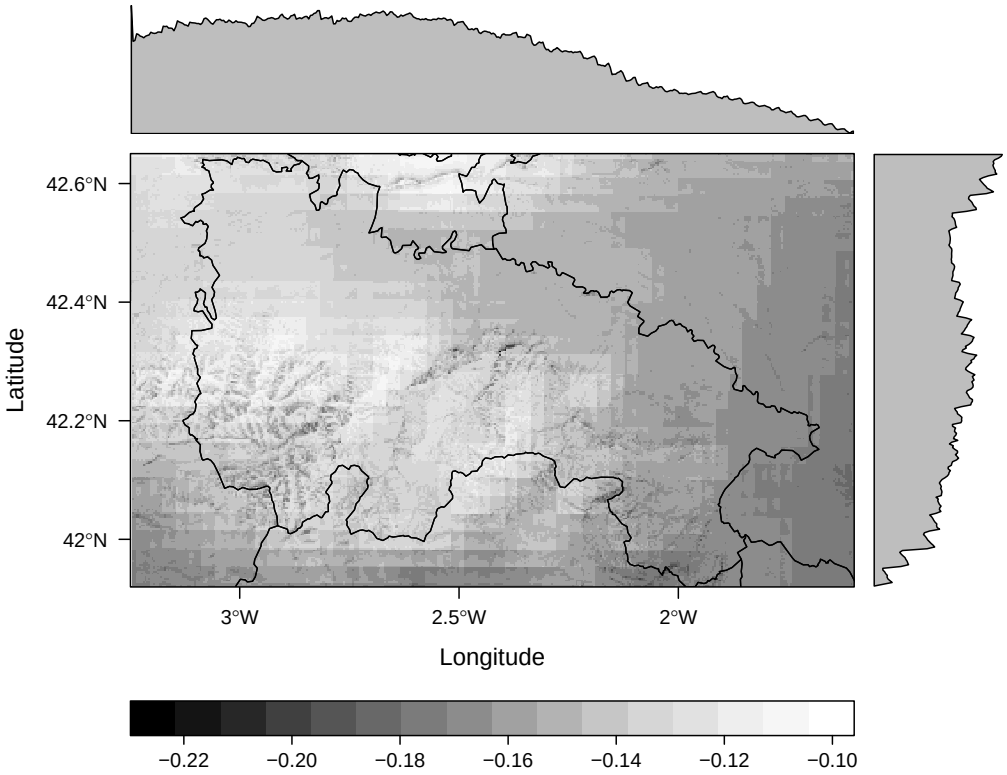

Figure 7: Relative difference of GHI_{KED} and $GHI_{CMSAF,down}$ related to $GHI_{CMSAF,down}$

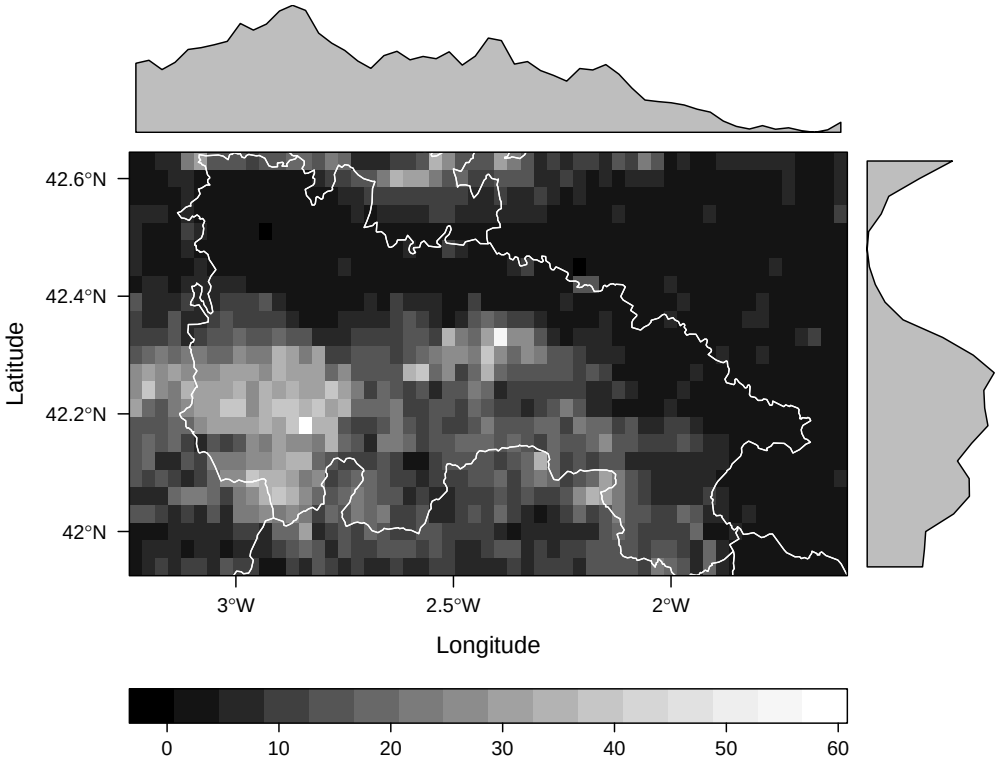

Figure 8: Difference of zonal standard deviations (kWh/m^2) between downscaling without KED and with KED.

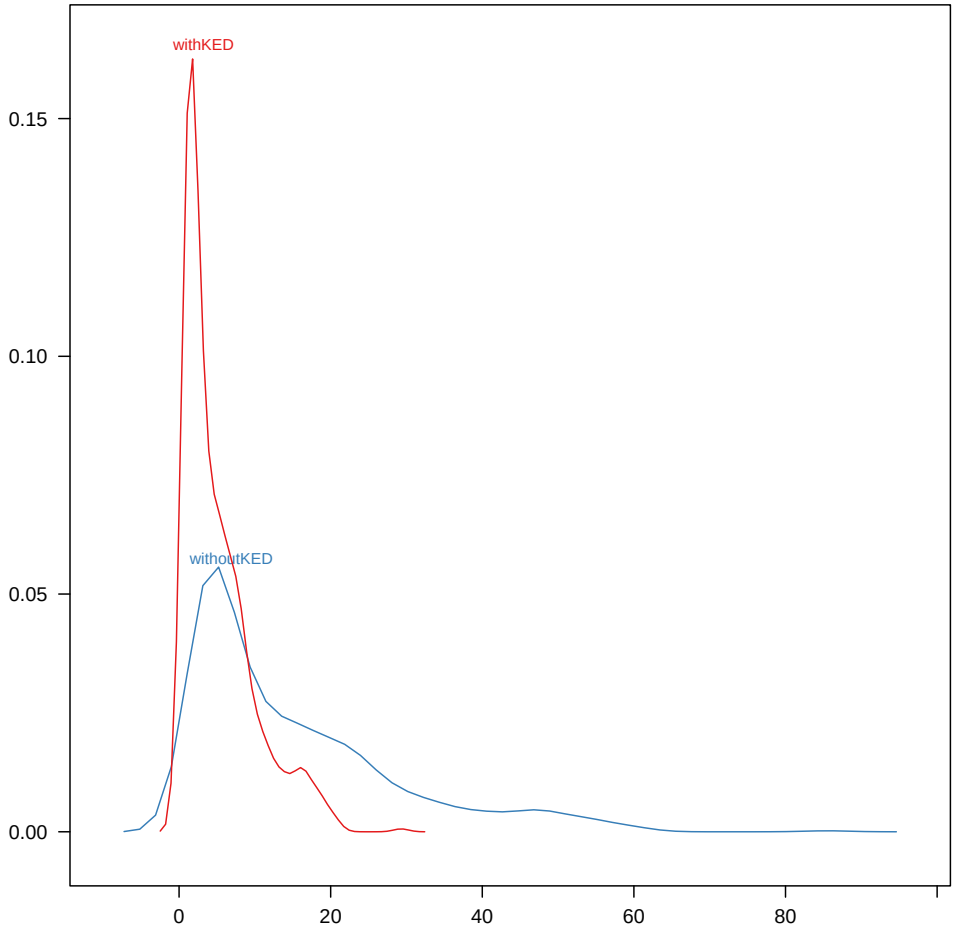

Figure 9: Density plot of zonal standard deviations between CM SAF and downscaling.