Agricultural and Forest Meteorology 176 (2013) 1–9

Contents lists available at SciVerse ScienceDirect

Agricultural and Forest Meteorology

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

Calibration of solar radiation models for Europe using Meteosat Second Generation and weather station data

Jędrzej S. Bojanowski[a,*], Anton Vrieling[b], Andrew K. Skidmore[b]

[a] European Commission, Joint Research Centre, Institute for Environment and Sustainability, Via E. Fermi 2749, I-21027 Ispra, VA, Italy
[b] Department of Natural Resources, Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

ARTICLE INFO

Article history:
Received 20 July 2012
Received in revised form 28 February 2013
Accepted 5 March 2013

Keywords:
Solar radiation
Model calibration
MSG-SEVIRI
MARS Crop Yield Forecasting System
Remote sensing
Spatial interpolation

ABSTRACT

Solar radiation is a key input variable for crop growth models. However, direct measurement of solar radiation is performed operationally for only a limited number of weather stations. Instead of direct measurements, empirical solar radiation models are used that link solar radiation to more commonly measured meteorological variables. Coefficients for these models are site-dependent and therefore generally interpolated from the few locations where solar radiation is measured. In this study, three solar radiation models were calibrated (Ångström–Prescott, Supit–Van Kappel, and Hargreaves) using a daily solar radiation product derived from Meteosat Second Generation data. This satellite-based calibration of model coefficients led to a higher accuracy when estimating daily solar radiation, as compared to the use of interpolated ground-based model coefficients. The average relative root mean square error for Meteosat Second Generation-based calibrated models was 1.9% lower for the Supit–Van Kappel model ($p<0.001$, $n=137$), and 1.8% lower for the Hargreaves model ($p<0.001$, $n=222$). There was no significant improvement using the Ångström–Prescott model. The Meteosat Second Generation-based model coefficients were interpolated to create continuous coefficient maps for Europe. From these maps it is possible to estimate solar radiation from the sunshine duration, cloud coverage and air temperature range for every location in Europe without prior calibration. We conclude that Meteosat Second Generation-based calibration of model coefficients improves the accuracy of solar radiation estimates.

1. Introduction

Continuous and accurate crop growth monitoring and yield forecasting at national and continental scales are important for early warning systems in developing countries, as well as for global food security and economic forecasting. The Monitoring Agricultural Resources Unit (MARS) of the European Commission's Joint Research Centre (JRC) provides crop yield forecasts for Europe throughout the crop season through the use of the MARS Crop Yield Forecasting System (Boogaard et al., 2002; Baruth et al., 2007). The MARS Crop Yield Forecasting System (MCYFS) collects and processes daily meteorological data together with crop management and soil data in order to model crop growth. The WOFOST crop growth model (van Diepen et al., 1989) is at the core of the MCYFS. WOFOST is adapted to the European scale based on a 25 km × 25 km grid. Crop growth indicators modelled by WOFOST, including water-limited storage organs and leaf area index, are used as predictors to issue short-term forecasts of European crop production from the beginning of the growing season up to harvesting. Key meteorological parameters used in the MCYFS at a daily time-step include: minimum, maximum, and mean air temperatures, total precipitation, mean water vapour pressure, mean wind speed, solar radiation and snow depth. These parameters are measured or modelled at weather station level and then interpolated to the 25 km × 25 km grid.

Of the required meteorological input variables to crop growth models, solar radiation is the most difficult parameter to obtain, due to the limited number of weather stations at which this parameter is measured (Roerink et al., 2012). Throughout the paper, we use the term 'solar radiation' to refer to the total of direct solar radiation and diffuse sky radiation received by a horizontal surface during a specific day. In other studies this is also referred to as 'global radiation'. In most European countries, the network of solar radiation measuring stations is too sparse for reliable interpolation of direct measurements. Accordingly, various approaches have been developed to estimate solar radiation. The MCYFS employs empirical solar radiation models which estimate solar radiation based on more commonly measured meteorological variables. For each weather station, the solar radiation model used depends on the parameters measured at that station. If solar radiation is directly measured, this measurement is used. Otherwise if a sunshine

* Corresponding author at: Federal Office of Meteorology and Climatology (MeteoSwiss), Zürich, Switzerland. Tel.: +41 44 256 9356; fax: +41 44 256 92 55.
E-mail address: jedrzej.bojanowski@meteoswiss.ch (J.S. Bojanowski).

0168-1923/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.agrformet.2013.03.005

duration estimate is available, solar radiation is modelled using the Ångström–Prescott model (Ångström, 1924; Prescott, 1940). If daily cloud cover, and minimum and maximum air temperatures are available, solar radiation is calculated using the Supit–Van Kappel model (Supit and van Kappel, 1998). If only minimum and maximum air temperatures are available then solar radiation is derived using the Hargreaves model (Hargreaves et al., 1985).

One of the key limitations of the empirical solar radiation models is their site-dependent coefficients (Abraha and Savage, 2008). Because of the lack of measured solar radiation data, these coefficients cannot be accurately determined for each weather station. The MCYFS uses the coefficients derived by Supit and van Kappel (1998) and van Kappel and Supit (1998). The first step of their approach was to identify a set of so called *reference stations* for which model coefficients were estimated, based on solar radiation measurements. Secondly, kriging was used to interpolate coefficients to all other European weather stations. However, the number of the reference stations used was limited due to the small number of stations reporting reliable solar radiation measurements. As a consequence, 256 reference stations were used to interpolate the coefficients for 7,126 weather stations over Europe. Those coefficients are currently used in the MCYFS.

Remotely sensed satellite observations are expected to significantly improve solar radiation estimation by their spatial and temporal continuity. Roerink et al. (2012) validated solar radiation estimates retrieved from Meteosat Second Generation data. The down-welling surface shortwave radiation flux (DSSF) is the name of the product operationally generated by the EUMETSAT's Land Surface Satellite Application Facility (LSA SAF) with a temporal frequency of 30 min at the full spatial resolution of the MSG/SEVIRI instrument (3 km). It was shown that the down-welling surface shortwave radiation flux represented a major improvement over the current approach for deriving solar radiation implemented in the MCYFS (Roerink et al., 2012). Nevertheless, Roerink et al. (2012) concluded that operational implementation of the down-welling surface shortwave radiation-flux product in the MCYFS was not yet possible because the time-series are still too short (since 2005 when Meteosat Second Generation was launched). The MCYFS relies on the regression between historic simulated and reported crop yields for which a consistent time-series of at least 10–15 years is needed (Roerink et al., 2012).

Besides the direct use of the satellite-derived solar radiation product by the MCYFS, we assess whether the product may be used indirectly as solar radiation reference data, thus improving the estimation of solar radiation model coefficients. This is achieved by calibrating the radiation models for weather stations where previously no calibration could be carried out. The increased number of locations with estimated model coefficients potentially improves the interpolation of model coefficients for the whole of Europe.

The objectives of this paper are: (1) to calibrate the solar radiation models (Ångström–Prescott, Supit–Van Kappel, and Hargreaves) used in the MCYFS by means of satellite-based down-welling surface shortwave radiation flux data as a reference; (2) to compare the performance of these newly calibrated model coefficients with those currently used in the MCYFS; and (3) to estimate the error in solar radiation estimation caused by the interpolation of model coefficients.

2. Materials

2.1. Surface solar radiation from Meteosat Second Generation

To fulfil the first objective of the paper, we used 6 years (2005–2010) of the down-welling surface shortwave radiation flux (DSSF) 30-min product derived from Meteosat Second Generation satellite data by the Land Surface Analysis Satellite Applications Facility (LSA SAF) and aggregated to daily values by the Flemish Institute for Technological Research (VITO) on behalf of JRC-MARS.

In the retrieval scheme applied, the down-welling surface shortwave radiation flux, F (W m^{-2}), is approximated as:

$$F = F_0 dT \cos\theta \qquad (1)$$

where F_0 (W m^{-2}) is the solar constant, T (unitless) the effective transmittance of the atmosphere and θ_s (rad) the solar zenith angle (Geiger et al., 2008). The factor d (unitless) accounts for a varying distance of the Sun as a function of the day of the year. To calculate the effective transmittance, two different approaches are used depending on whether a given pixel is classified as clear or cloudy. The cloud mask is generated by means of software components developed by the Satellite Application Facility through a project "Support to Nowcasting and Very Short Range Forecasting" (Geiger et al., 2008). The approach used for the cloud-free pixels does not employ the satellite signal of the Meteosat Second Generation. The effective transmittance is estimated as a function of the atmospheric composition following Frouin et al. (1989) employing the modelled water vapour column density taken from the European Centre for Medium-Range Weather Forecast (ECMWF), and the amount of ozone retrieved from the Total Ozone Mapping Spectrometer (TOMS). The methodology applied for the cloudy pixels mainly follows Brisson et al. (1999) and employs the data sensed by three channels of the Spinning Enhanced Visible and Infrared Imager (SEVIRI) sensor centred at 0.6 μm, 0.8 μm and 1.6 μm. The method relies on a simplified physical description of the radiation transfer in the cloud-atmosphere-surface system, assuming that the whole pixel is covered by a homogenous cloud layer (Ineichen et al., 2009).

2.2. Selection of weather stations

Meteorological data were obtained from the JRC-MARS database. It contains data that are collected for around seven thousand weather stations in Europe, of which approximately five thousand are currently still active. In the frame of the MCYFS project, before entering the database, the data were checked for consistency and errors by comparison of each observation with the corresponding observations of surrounding stations. Finally, the data were aggregated to a daily time step.

For each solar radiation model we used a different subset of stations for the calibration, depending on the meteorological input needed by the models and meteorological variables measured at each station. We calibrated the Hargreaves model using remotely sensed solar radiation data (DSSF) for 2153 stations for which air temperature range data were available for the period 2005–2010, i.e. a period with complete DSSF product availability. For 1540 stations, the cloud coverage was also measured, which was used to calibrate the Supit–Van Kappel model. The Ångström–Prescott model was calibrated for the 820 stations at which sunshine duration was measured.

To analyse the performance of the solar radiation models, we extracted data from all stations which reported both solar radiation and the required meteorological variables (per model) for a period longer than 4 years and including at least one year within the period 2005–2010. The maximum length of the remaining station time series was equal to 40 years, while the average length was 10 years. We removed the *reference stations* from our validation set of stations. In the current MCYFS, these reference stations are the locations for which model coefficients were calculated based on the ground measurements and then interpolated. In our validation set of stations, we retained 39 stations to evaluate the

Ångström–Prescott model, 134 to evaluate the Supit–Van Kappel model, and 222 to evaluate the Hargreaves model.

3. Methods

3.1. Estimating solar radiation from sunshine hours

If weather stations do not report direct solar radiation measurements, the first choice in the MCYFS is to model solar radiation from sunshine hours. The method proposed by Ångström (1924) and modified by Prescott (1940) models the surface solar radiation (I_s) as a linear relationship between the daily atmospheric transmissivity and the fraction of daily sunshine duration (n) and day length (N), both in hours per day. This relation is expressed as:

$$I_s = I_x \left(A_a + A_b \frac{n}{N} \right) \tag{2}$$

where A_a and A_b are site-specific empirical coefficients and I_x (MJ m^{-2} day^{-1}) extraterrestrial solar radiation. Although the coefficients have a physical meaning (e.g. the sum of A_a and A_b can be treated as clear sky transmissivity), they are usually derived empirically using measured reference solar radiation data.

3.2. Estimating solar radiation from cloud coverage and air temperature

Sunshine hours are reported for only a limited number of stations. However recordings of daytime cloud cover and daily maximum and minimum air temperatures are more common, as these can be measured without specialist instruments. The portion of potential solar radiation reaching the earth's surface during the day is reduced due to cloud cover. The relation between solar radiation and daily air temperature range is based on the assumption that during clear-sky days the temperature range is greater; in other words, during the day the air temperature is greater as the incoming solar radiation is not filtered by clouds, while during the night the infrared emission from the soil surface is rapidly lost in the atmosphere and the air temperature decreases more quickly. Supit and van Kappel (1998) combined a cloud coverage model (Wörner, 1967) and an air temperature model (Hargreaves et al., 1985) as follows:

$$I_s = I_x \left[S_a \sqrt{T_x - T_n} + S_b \sqrt{1 - \frac{C_w}{8}} \right] + S_c \tag{3}$$

where S_a, S_b and S_c are the site-specific empirical coefficients, T_x and T_n are the maximum and minimum daily air temperatures (°C), and C_w is the mean total cloud cover during daytime observations (in oktas).

3.3. Estimating solar radiation from air temperature

Daily maximum and minimum daily air temperatures are reported for practically all stations, while cloud cover is not. A simple method of relating I_s to daily air temperature range was proposed by Hargreaves et al. (1985):

$$I_s = I_x H_a \sqrt{T_x - T_n} + H_b \tag{4}$$

where H_a and H_b are site-specific empirical coefficients. Because the equation is simple and most weather stations collect air temperature data, the model can be easily used for most stations. The model is based on the assumption that the site is not significantly affected by advection. In many cases, this assumption can result in temperature-based model estimates of low accuracy (Winslow et al., 2001).

3.4. Calibration of the models

To calculate coefficients for all three models and stations used in this study, we applied the least square regression technique as implemented in the R-package *sirad* (Bojanowski, 2012; R Development Core Team, 2012). The method determines the model coefficients that provide an optimal linear fit between modelled and reference solar radiation data. In the least square regression technique, optimal refers to the solution with the lowest root mean square error (RMSE).

First, we estimated model coefficients for the Ångström–Prescott, Supit–Van Kappel, and Hargreaves models for the 820, 1540 and 2153 stations respectively using the Meteosat Second Generation-based DSSF product from years 2005–2010 as a reference. In this article, we refer to this calibration as '*Meteosat Second Generation-based*'.

We applied the same calibration approach using ground-based solar radiation measurements. This resulted in a reference dataset of model coefficients that was subsequently used to assess the accuracy of the Meteosat Second Generation-based calibration. We split each station's time series into two equal parts (for calibration and validation) and we used the second part (more recent period) for this so called '*ground*' calibration. *Ground-based* coefficients for the Ångström–Prescott model were calculated for 39 stations, for the Supit–Van Kappel model for 134 stations, and for the Hargreaves model for 222 stations.

3.5. Spatial interpolation of model coefficients

We interpolated the *Meteosat Second Generation-based* coefficients using kriging to obtain a model-coefficient map for Europe. The kriging procedure models a semivariogram from the spatial correlation structure between location values and the distance between the locations (Cressie, 1991). In our case, the location values are the calibrated model coefficients. The semivariance structure is estimated by the experimental semivariogram expressed as:

$$\gamma(h) = \frac{1}{2} \sum (z(S_i) - z(S_i + h))^2 \tag{5}$$

where h is a lag distance, $z(S_i)$ is the value of a model coefficient at some sample location and $z(S_i + h)$ is the value of the neighbour at distance $S_i + h$. Using the R package *gstat* (Pebesma, 2004; R Development Core Team, 2012), the obtained empirical semivariograms were fitted by using theoretical semivariogram models such as spherical for H_a and S_c, Matern (Stein, 1999) for H_b, S_a, S_b, and Gaussian for A_a, A_b. Ordinary kriging was then used to spatially interpolate the model coefficients.

3.6. Evaluation of solar radiation model coefficients

Using the procedures set out above for each station and each model, we obtained three sets of model coefficients: (1) ground-based, (2) Meteosat Second Generation-based, and (3) Meteosat Second Generation-based interpolated. As a fourth set, we also evaluated the model coefficients that are currently used in the MCYFS. For each dataset and station, the model coefficients were applied to the first half of the time-series of available meteorological observations (which had not been used before for the calibration), resulting in four sets of solar radiation time series per station. Additionally, for evaluating the *Meteosat Second Generation-based* and *Meteosat Second Generation-based interpolated* coefficients, the period 2005–2010 which had been previously used for calibration, was removed. As noted above, for each model a different subset of weather stations was used, depending on the model's required parameters

Table 1
Performance statistics for European weather stations of the solar radiation models depending on the calibration method used (ground measurements, Meteosat Second Generation [MSG], currently used in the MCYFS).

	RMSE ($MJ\,m^{-2}$)				RRMSE (%)				MBE ($MJ\,m^{-2}$)			
	Min	Max	Mean	sd	Min	Max	Mean	sd	Min	Max	Mean	sd
Ångström–Prescott												
Ground-based	1.3	4.1	2.0	0.7	7.2	36.9	14.8	6.2	−1.7	0.7	−0.2	0.5
MSG-based	1.3	6.0	2.3	1.0	7.1	53.6	17.6	9.2	−1.1	4.5	0.4	1.4
MSG-interpolated	1.3	5.4	2.2	0.9	7.3	49.2	17.2	8.5	−1.3	3.7	0.3	1.2
MCYFS	1.3	5.8	2.3	1.1	7.2	51.7	17.1	8.5	−2.5	4.3	0.3	1.3
Supit–Van Kappel												
Ground-based	1.7	6.0	2.4	0.8	9.3	38.6	19.7	4.8	−1.3	1.7	−0.1	0.4
MSG-based	1.8	6.1	2.6	0.9	10.6	42.9	21.2	5.2	−2.3	3.2	0.2	0.8
MSG-interpolated	1.8	6.4	2.9	1.0	12.0	42.4	22.9	6.1	−3.9	4.2	0.6	1.2
MCYFS	1.8	6.6	2.9	0.9	13.5	41.7	23.1	5.7	−4.2	2.9	−0.9	1.2
Hargreaves												
Ground-based	2.4	6.5	3.4	0.6	13.1	46.5	29.2	6.4	−1.5	1.4	0.0	0.4
MSG-based	2.7	7.0	3.5	0.7	13.6	48.4	29.9	6.3	−3.8	3.4	0.1	0.8
MSG-interpolated	2.7	9.6	3.7	0.9	14.3	52.1	31.3	7.0	−7.7	4.8	0.2	1.4
MCYFS	2.8	11.9	3.8	0.9	13.9	64.9	31.7	7.2	−10.3	3.7	−0.9	1.3

The resulting solar radiation time series were compared against the measured solar radiation for the selected stations. *A priori* we assumed that models employing *ground-based* coefficients would provide more accurate results as compared to models employing the Meteosat Second Generation-based coefficients, because measured solar radiation values were used for calibration.

The model's performance was assessed on the basis of the RMSE and the mean bias error (MBE) as indicators of non-systematic and systematic errors respectively. We also calculated a relative root mean square error (RRMSE) by dividing the RMSE by the average value of all measured values. In addition, for each model we applied a two-sample paired *t*-test to examine whether the average RRMSE for all stations was significantly different between the sets of model coefficients used.

The accuracy of the interpolation of the *Meteosat Second Generation-based* model coefficients was tested by the 10-fold cross-validation method (Burman, 1989). In this method, the stations were randomly allocated to 10 groups. Then, one after another, each group was treated as the testing group, while the remaining stations served as the training set. Validation was thus performed 10 times, and the final validation result was the average of 10 validations.

4. Results

4.1. Evaluation of the models employing Meteosat Second Generation-based calibration

The Meteosat Second Generation-based calibration does not significantly improve the performance of the solar radiation model in comparison to the current Ångström–Prescott model implementation in the MCYFS ($p = 0.38$, $n = 39$ [stations]). For 18 weather stations, the solar radiation estimates derived from the model using Meteosat Second Generation-based calibrated coefficients are less accurate than the estimates of the current MCYFS, whereas estimates for the other 21 weather stations have a greater accuracy (Fig. 2a). The average RRMSE is 17.6% for Meteosat Second Generation-based calibration, as against a slightly lower value (17.1%) for the MCYFS. The MBE varies between −1.1 and 4.5 $MJ\,m^{-2}$ (Table 1). At four stations, the mean bias error is greater than 3 $MJ\,m^{-2}$.

The solar radiation estimates derived from the Meteosat Second Generation-based calibrated Supit–Van Kappel model are significantly more accurate than those derived from the current coefficients implemented in MCYFS (Table 2). The average RRMSE for estimates using the *Meteosat Second Generation-based* calibrated model is 1.9% lower than when implementing the MCYFS ($p < 0.01$, $n = 134$ [stations]). For 93 out of the 134 stations analysed, *Meteosat Second Generation-based* calibration provides more accurate results than the current MCYFS (Fig. 2c).

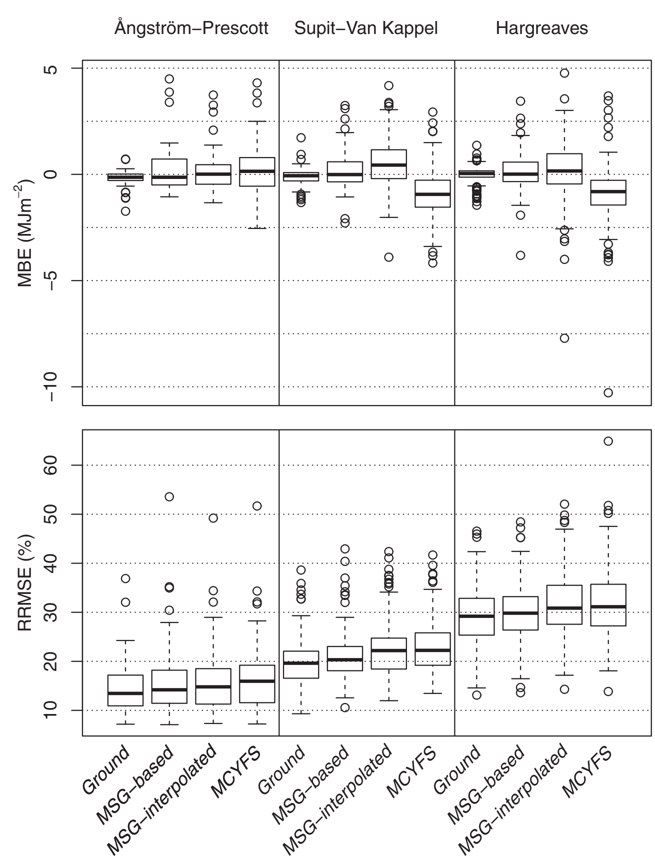

Fig. 1. Box plot of the MBE and RRMSE calculated against measured data at every station for the four methods of the calibration (ground-based, Meteosat Second Generation-based, Meteosat Second Generation-based-interpolated, and current MARS Crop Yield Forecasting System [MCYFS]) for the Ångström–Prescott, Supit–Van Kappel and Hargreaves models. Circles represent values exceeding 1.5 times the interquartile range.

Table 2
Significance of the differences between the performance (RRMSE) of the Ångström–Prescott, Supit–Van Kappel and Hargreaves models depending on the calibration method: ground-based, Meteosat Second Generation-based, Meteosat Second Generation-based-interpolated, and the current MCYFS.

X	*Y*	One-sided *t*-test ($X=Y$ vs $X<Y$) for paired samples								
		Ångström–Prescott			Supit–Van Kappel			Hargreaves		
		t	df	*p*	*t*	df	*p*	*t*	df	*p*
$RRMSE_{ground}$	$RRMSE_{MSG}$	−3.65	38	0.0004	−8.05	133	1.99×10^{-13}	−7.69	221	2.35×10^{-13}
$RRMSE_{MSG}$	$RRMSE_{interp}$	0.90	38	0.81	−6.69	133	2.80×10^{-10}	−7.55	221	5.60×10^{-13}
$RRMSE_{ground}$	$RRMSE_{MCYFS}$	−3.40	38	0.0008	−11.25	133	2.46×10^{-21}	−10.30	221	7.00×10^{-21}
$RRMSE_{MSG}$	$RRMSE_{MCYFS}$	−0.31	38	0.38	−6.18	133	3.70×10^{-9}	−7.62	221	3.69×10^{-13}
$RRMSE_{interp}$	$RRMSE_{MCYFS}$	−0.67	38	0.25	−0.66	133	0.25	−1.99	221	0.02

The temperature-based Hargreaves model calibrated using Meteosat Second Generation data yields more accurate solar radiation estimates than the currently used MCYFS coefficients for the same model. The average RRMSE of the Meteosat Second Generation-based calibrated model is significantly lower ($p<0.01$, $n=222$ [reference stations]) as compared to the MCYFS implementation (29.9% vs 31.7%). The average MBE of the Meteosat Second Generation-based calibrated model is 0 MJ m^{-2}, whereas the MCYFS underestimates the solar radiation values (Fig. 1) with the average MBE of −0.9 MJ m^{-2}. For only 9 of the 222 stations did the current MCYFS coefficients yield more accurate results than the Meteosat Second Generation-based calibration.

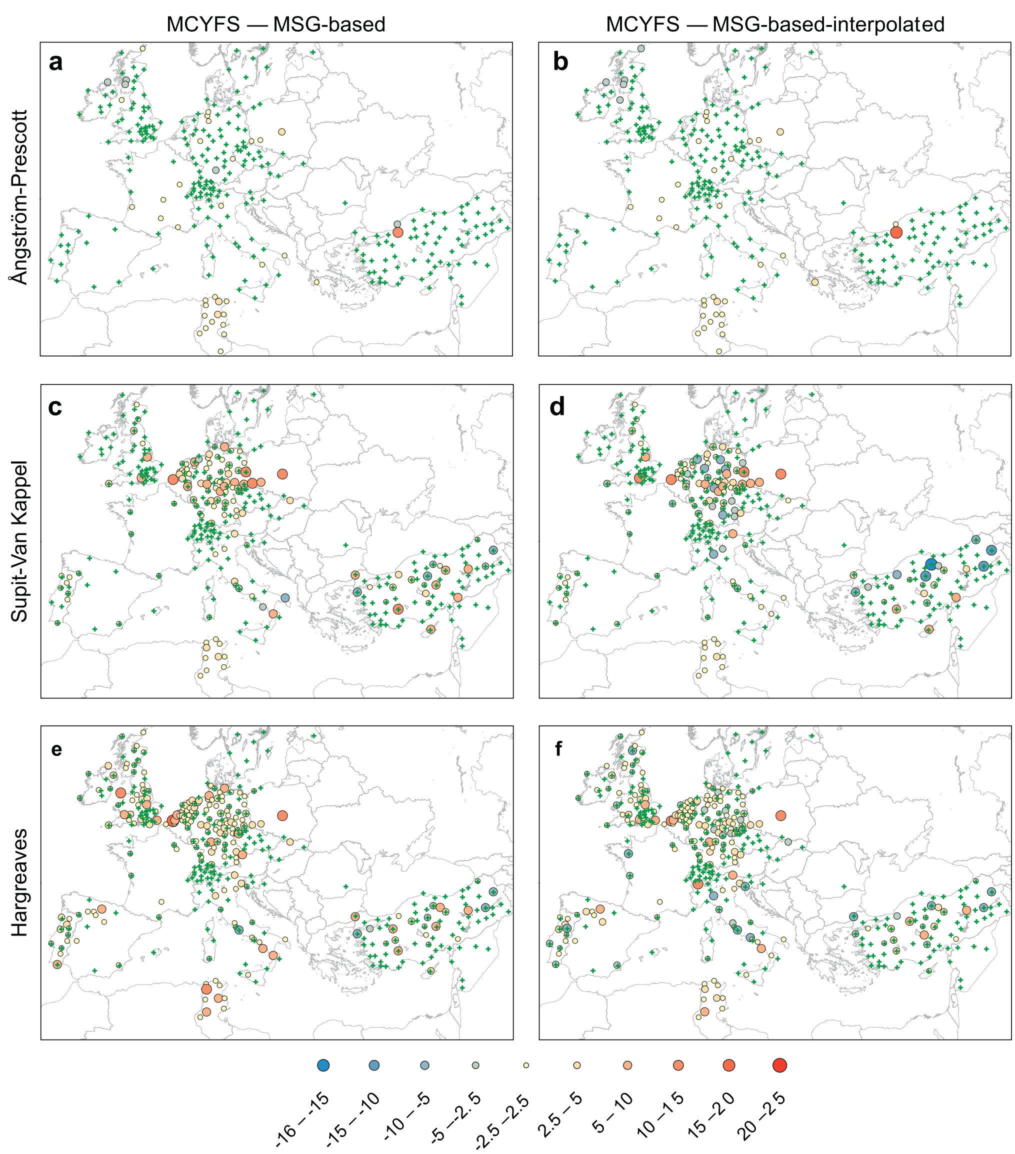

Fig. 2. Differences in RRMSE of the Ångström–Prescott, Supit–Van Kappel and Hargreaves models depending on the coefficients used: current MARS Crop Yield Forecasting System [MCYFS] minus Meteosat Second Generation-based [MSG-based] and current MCYFS minus Meteosat Second Generation-based interpolated. Red dots indicate stations where Meteosat Second Generation-based solar radiation estimates proved more accurate than the current MCYFS estimates. Weather stations used in the MCYFS for interpolating model coefficients are marked by green crosses.

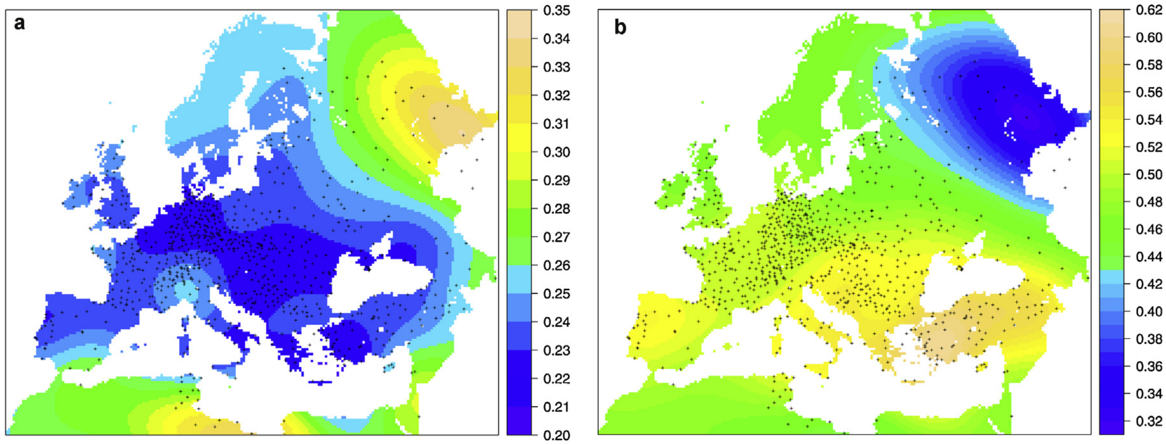

Fig. 3. Interpolation results of Ångström–Prescott model coefficients: predicted A_a (a) and A_b (b). Weather stations used for interpolation are marked by black crosses.

4.2. Interpolation of model coefficients

Fig. 3 presents the results of the spatial interpolation of Ångström–Prescott model coefficients A_a and A_b (Eq. (1)) that were derived from the Meteosat Second Generation-based calibration. The mean absolute interpolation error was 0.014 ± 0.016 for A_a and 0.025 ± 0.03 for A_b.

The interpolated S_a and S_b coefficients of the Supit–Van Kappel model (Fig. 4a and b) are in a similar range to the coefficients reported by Supit and van Kappel (1998). Only the values of the coefficients estimated for the high latitudes of Scandinavia and Russia were not reported by Supit and van Kappel (1998) because this region contained no test sites. The variability of the S_c coefficient (Fig. 4c) is greater than described by (Supit and van Kappel, 1998). The mean absolute cross-validation residuals were: 0.011 ± 0.01 for S_a, 0.039 ± 0.272 for S_b, and 0.218 ± 0.272 for S_c.

Fig. 5 presents the results of the spatial interpolation of the Hargreaves model coefficients H_a and H_b (Eq. (3)) derived from the Meteosat Second Generation data. H_a values are relatively high for areas around the Mediterranean Sea (Fig. 5a), unlike the results reported by van Kappel and Supit (1998). Based on the cross-validation the mean absolute interpolation error H_a was 0.013 ± 0.012 and 0.313 ± 0.396 of H_b.

4.3. Evaluation of the models employing interpolated coefficients

The Ångström–Prescott model employing *Meteosat Second Generation-based* yields *interpolated* coefficients with a similar accuracy to the model coefficients currently used in the MCYFS. The RRMSE resulting from the Meteosat Second Generation-based interpolated calibration is not significantly different from the RRMSE of the MCYFS implemented coefficients, i.e. 17.2% vs 17.1% (Tables 1 and 2).

The Supit–Van Kappel model employing *Meteosat Second Generation-based-interpolated* coefficients has an average RRMSE of 22.9%, which is 1.7% lower than for the model employing locally derived Meteosat Second Generation-based coefficients. The box-plot of the mean bias error (Fig. 1) illustrates that for most stations the model overestimates solar radiation. For the MCYFS, Fig. 1 shows an underestimation of the solar radiation values. The RRMSE of the Meteosat Second Generation-based interpolated calibration (22.9%) is not significantly less ($p = 0.25$) than the RRMSE of the MCYFS implemented coefficients (23.1%). For 74 of the 134 stations tested, solar radiation derived from the MCYFS coefficients shows a greater error than for *Meteosat Second Generation-based-interpolated* coefficients (Fig. 2d).

The average RRMSE of the Hargreaves model employing *Meteosat Second Generation-based interpolated* coefficients is significantly lower ($p = 0.023$) than for the model using MCYFS-implemented coefficients (31.3% vs 31.7%). The model calibrated using Meteosat Second Generation data has a mean bias error of 0.2 MJ m^{-2}, which is less in magnitude than the −0.9 MJ m^{-2} obtained when implementing the MCYFS.

5. Discussion

The performance of the Ångström–Prescott model did not improve following Meteosat Second Generation-based calibration. For the Supit–Van Kappel and Hargreaves models, estimates of daily solar radiation using the Meteosat Second Generation-based calibrated model coefficients were more accurate than estimates obtained from the method currently used in the MCYFS.

At most locations distant from the MCYFS' *reference stations*, both *Meteosat Second Generation-based* and *Meteosat Second Generation-based-interpolated* coefficients performed with a greater accuracy than the MCYFS (Fig. 2a and b). Thus, it can be expected that the average error of the model employing MCYFS' coefficients would be greater for locations far from the weather stations for which the model coefficients were calculated based on ground measurements of solar radiation.

The accuracy of solar radiation estimates from the Supit–Van Kappel model employing Meteosat Second Generation-based coefficients was statistically significantly greater. In other words, for weather stations in the MCYFS database, coefficients derived from Meteosat Second Generation increased the overall accuracy of modelled solar radiation.

The Hargreaves model, which uses only temperature range as input, was the least accurate of all the models analysed. This agrees with Mavromatis (2008) and Trnka et al. (2005), who found that temperature-based models provide less accurate estimates of solar radiation. For temperature-based models, strong advection at some locations causes models to perform poorly and calibration cannot effectively cater for this. The improvement gained by the Meteosat Second Generation-based calibration comparing to the MCYFS implementation was significant. Application of the Meteosat Second Generation-based coefficients would increase the accuracy of the solar radiation estimates in comparison to the method currently used in the MCYFS.

The interpolation of the Meteosat Second Generation-based coefficients did not significantly reduce the performance of Supit–Van Kappel and Hargreaves models compared to MSG-based calibration. A significant improvement in comparison with

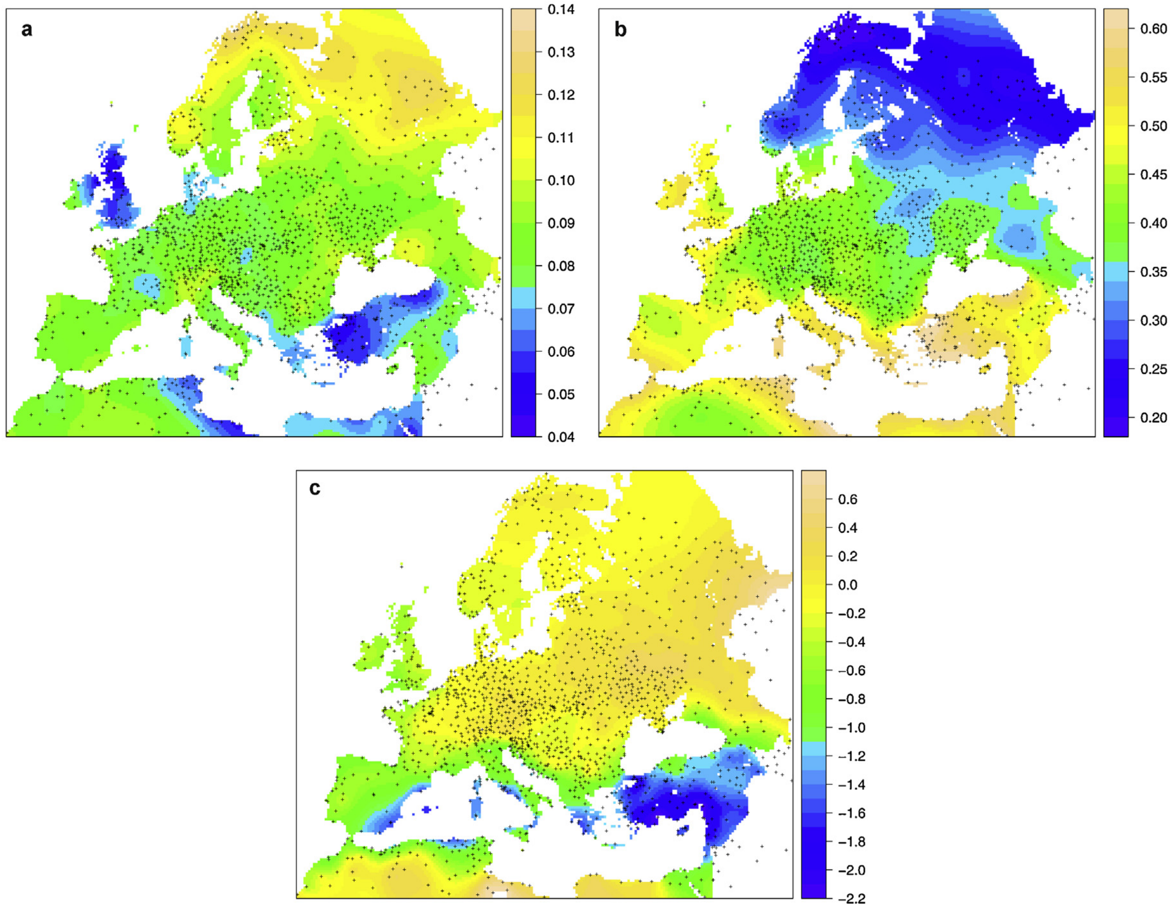

Fig. 4. Interpolation results of Supit–Van Kappel model coefficients: predicted S_a (a), S_b (b) and S_c (c). Weather stations used for interpolation are marked by black crosses.

current implementation in MCYFS was observed for Hargreaves model. This demonstrates that interpolated coefficients can be used for the estimation of solar radiation with good accuracy. Nevertheless, the ordinary kriging used in this study does not account for other possible explanatory variables that may influence the value of the model coefficients, such as altitude. More advanced interpolation techniques such as co-kriging may result in higher interpolation accuracies, and therefore, in a more accurate estimate of solar radiation for locations not used in the interpolation. However, a comparison of interpolation techniques was outside the scope of this paper, but would be valuable for future study.

The quality of the solar radiation measurements used for the evaluation of the models needs to be considered when interpreting results for individual weather stations. Although data included in the MARS meteorological database was manually checked for inconsistencies and outlier values, the high quality and consistency of the time series cannot be guaranteed for each individual station. The creation of a homogenous database of accurate solar radiation measurements is challenging, especially for an entire

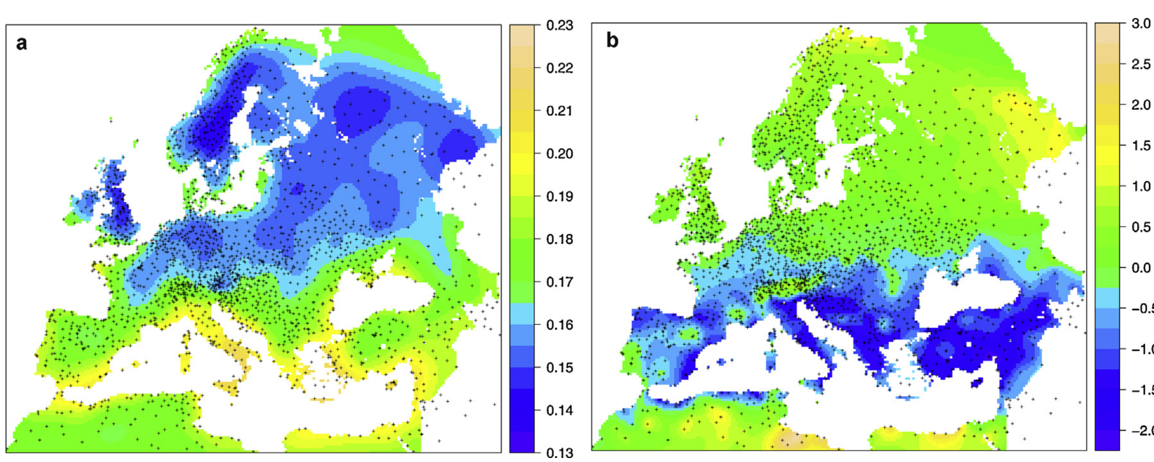

Fig. 5. Interpolation results of Hargreaves model coefficients: predicted H_a (a) and H_b (b). Weather stations used for interpolation are marked by black crosses.

continent. The diversity of the instruments used that have different accuracies and the need of frequent instrument calibration make creation of such a database almost unfeasible. For this reason, databases that provide consistent quality-controlled surface radiation data, such as the Baseline Surface Radiation Network (Ohmura et al., 1998), include a very limited number of locations. For this study we decided instead to use a large number of weather stations that measure solar radiation, even if their accuracies may differ. The better spatial coverage makes our analysis on the performance of solar radiation models more representative for the whole continent. However, we recommend further evaluation of our satellite-based calibration method and obtained coefficients with another independent high-quality solar radiation dataset.

The Meteosat Second Generation-based calibration can improve the quality of the solar radiation estimation in the MCYFS. Although in this paper we do not test if this would result in more accurate crop yield estimates, Roerink et al. (2012) showed that direct use of accurate and spatially-continuous Meteosat Second Generation-based solar radiation estimates represents a significant improvement compared with the measured and modelled solar radiation used in the MCYFS. This implies that more accurate solar radiation estimates translate into more accurate crop yield estimates.

Currently the greatest obstacle to using satellite-based estimates directly in crop models is the limited temporal coverage: crop yield forecasting systems (such as the MCYFS) require long (>10–15 years) time series to perform a regression between historic simulated and reported crop yields. For Europe and Africa, options do exist however to extend the Meteosat Second Generation time series back in time. A recently published solar radiation product from the Meteosat First Generation (Posselt et al., 2010) could be used to extend the Meteosat Second Generation product from 1983 to present. Even if this extension proves to have benefits (including good inter-comparability of both data sets), our newly calibrated solar radiation models could still be used to fill the existing and future gaps in the satellite-based estimates, and to calculate solar radiation for the years before 1983, for which Meteosat-based solar radiation estimations are not available.

6. Conclusions

Our study demonstrates that solar radiation models calibrated using satellite-based Meteosat Second Generation data estimate daily solar radiation with greater accuracy than models employing coefficients derived from ground measurements and then interpolated (i.e. the current MCYFS implementation). The improvement in model performance was statistically significant for the Supit–Van Kappel and Hargreaves models, whereas for the Ångström–Prescott model the improvement was apparent only for stations located in a sparse network of reference stations.

We validated our calibration method for European weather stations. However, the same approach can easily be transferred to Africa, which is contained within the Meteosat Second Generation disc. In addition, our method may prove useful for other continents when using solar radiation data derived from other geostationary satellites (i.e. INDOEX, GOES E, GOES W, and GMS). Application of our method does not require a specialist remote-sensing knowledge. Instead, the calibration of solar radiation models can be performed using published satellite-based solar radiation datasets. Alternatively for Europe, our interpolated model coefficients could be used directly. However, based on the results presented in this study, we highly recommend deriving location-specific coefficients instead of using interpolated ones.

This paper was driven by the need to improve solar radiation estimates, which could subsequently be used in an operational crop growth modelling system (here in particular the MCYFS). Notwithstanding, our calibration method and model coefficient maps for Europe can be of interest of a broader community than crop growth modellers only. This includes anyone that requires solar radiation estimates as input to further studies on a regional, continental, or potentially global scale, for example in the domains of agriculture, forestry, solar energy, or climate studies.

Acknowledgement

The authors would like to thank the two anonymous reviewers for their valuable comments and suggestions which helped to improve the article.

References

Abraha, M.G., Savage, M.J., 2008. Comparison of estimates of daily solar radiation from air temperature range for application in crop simulations. Agric. For. Meteorol. 148, 401–416.
Ångström, A., 1924. Solar and terrestrial radiation. Report to the international commission for solar research on actinometric investigations of solar and atmospheric radiation. Q. J. R. Meteorol. Soc. 50, 121–126.
Baruth, B., Genovese, G., Leo, O. (Eds.), 2007. CGMS version 9.2—user manual and technical documentation. JRC40605. EUR 22936 EN.
Bojanowski, J.S., 2012. sirad: functions for calculating daily solar radiation. R package version 1.0-8/r17. http://R-Forge.R-project.org/projects/sirad/
Boogaard, H.L., Eerens, H., Supit, I., van Diepen, C.A., Piccard, I., Kempeneers, P., 2002. Description of the MARS Crop Yield Forecasting System (MCYFS). Joint Research Centre, Study contract number 19226-2002-02-F1FED ISP NL.
Brisson, A., Borgne, P.L., Marsouin, A., 1999. Development of algorithms for Surface Solar Irradiance retrieval at O&SI SAF Low and Mid Latitudes. Météo-France/SCEM/CMS O & SI SAF Low and Mid Latitudes.
Burman, P., 1989. A comparative study of ordinary cross-validation, v-fold cross-validation and repeated learning-testing methods. Biometrika 76, 503–514.
Cressie, N.A.C., 1991. Statistics for Spatial Data. John Wiley & Sons.
Frouin, R., Lingner, D.W., Gautier, C., Baker, K.S., Smith, R.C., 1989. A simple analytical formula to compute clear sky total and photosynthetically available solar irradiance at the ocean surface. J. Geophys. Res. 94, 9731–9742.
Geiger, B., Meurey, C., Lajas, D., Franchisteguy, L., Carrer, D., Roujean, J.-L., 2008. Near real-time provision of downwelling shortwave radiation estimates derived from satellite observations. Meteorol. Appl. 15, 411–420.
Hargreaves, G.L., Hargreaves, G.H., Riley, J.P., 1985. Irrigation water requirements for Senegal River basin. J. Irrig. Drain. Eng. 111, 265–275.
Ineichen, P., Barroso, C.S., Geiger, B., Hollmann, R., Marsouin, A., Mueller, R., 2009. Satellite Application Facilities irradiance products: hourly time step comparison and validation over Europe. Int. J. Remote Sens. 30, 5549–5571.
Mavromatis, T., 2008. Estimation of solar radiation and its application to crop simulation models in Greece. Clim. Res. 36, 219–230.
Ohmura, A., Dutton, E.G., Forgan, B., Fröhlich, C., Gilgen, H., Hegner, H., Heimo, A., Konig-Langlo, G., McArthur, B., Müller, G., Philipona, R., Pinker, R., Whitlock, C.H., Dehne, K., Wild, M., 1998. Baseline Surface Radiation Network (BSRN/WCRP): new precision radiometry for climate research. Bull. Am. Meteorol. Soc. 79, 2115–2136.
Pebesma, E., 2004. Multivariable geostatistics in S: the gstat package. Comput. Geosci. 30, 683–691.
Posselt, R., Müller, R., Trentmann, J., Stöckli, R., 2010. Surface radiation climatology derived from Meteosat First and Second Generation satellites. European Geophysical Union. Geophys. Res. Abstracts, Vienna.
Prescott, J.A., 1940. Evaporation from a water surface in relation to solar radiation. Trans. R. Soc. South Aust. 64, 114–125.
R Development Core Team, 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN: 3-900051-07-0 http://www.R-project.org/
Roerink, G.J., Bojanowski, J.S., de Wit, A.J.W., Eerens, H., Supit, I., Leo, O., Boogaard, H.L., 2012. Evaluation of MSG-derived global radiation estimates for application in a regional crop model. Agric. For. Meteorol. 160, 36–47.
Stein, M.L., 1999. Interpolation of Spatial Data: Some Theory for Kriging (Springer Series in Statistics), 1st ed. Springer, New York.
Supit, I., van Kappel, R.R., 1998. A simple method to estimate global radiation. Sol. Energy 63, 147–160.
Trnka, M., Žalud, Z., Eitzinger, J., Dubrovský, M., 2005. Global solar radiation in Central European lowlands estimated by various empirical formulae. Agric. For. Meteorol. 131, 54–76.

van Diepen, C.A., Wolf, J., van Keulen, H., 1989. WOFOST: a simulation model of crop production. Soil Use Manage. 5, 16–24.
van Kappel, R.R., Supit, I., 1998. Geographical extension of the EU global radiation study (EUR 15745) so as to include also Eastern Europe and Northern Africa. EUR 16371 EN, JRC, Ispra, Italy.
Winslow, J.C., Hunt, E.R., Piper, S.C., 2001. A globally applicable model of daily solar irradiance estimated from air temperature and precipitation data. Ecol. Model. 143, 227–243.
Wörner, H., 1967. Zur Frage der Automatisierbarkeit der Bewölkungsangaben durch Verwendung von Strahlungsgrößen. Abh. Met. Dienst. DDR 11.