An auto-calibration procedure for empirical solar radiation models*

J. S. Bojanowski1,2, M. Donatelli1,3, A. K. Skidmore4, and A. Vrieling4

1 European Commission, Joint Research Centre, Via E. Fermi 2749, I-21027 Ispra (VA), Italy
2 Federal Office of Meteorology and Climatology MeteoSwiss, Climate Services, Krähhüblstrasse 58, 8044 Zürich, Switzerland
3 Agriculture Research Council, Research Centre for Industrial Crops, I-40128 Bologna, Italy
4 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

Abstract

Solar radiation data are an important input for estimating evapotranspiration and modelling crop growth. Direct measurement of solar radiation is now carried out in most European countries, but the network of measuring stations is too sparse for reliable interpolation of measured values. Instead of direct measurements, solar radiation may be estimated from empirical solar radiation models that employ more commonly measured variables or direct outputs of general and regional circulation models (such as air temperature). Coefficients for these models are site-dependent. This usually implies that they are estimated for stations with direct radiation measurements, but need to be interpolated for other locations. In this paper, we introduce a procedure to auto-calibrate empirical solar radiation models that are based on daily air temperature range, i.e. Bristow and Campbell (1984), and Hargreaves (1985). Meteosat Second Generation data were used to create two static look-up tables of mean cloud cover and clear-sky transmissivity as input for the auto-calibration procedure. We demonstrate that daily solar radiation can be accurately estimated from daily air temperature range measurements without site-specific empirical coefficients that require stations that measure solar radiation. The average relative root mean square error for our auto-calibrated models was comparable to ground-measurement-based calibration; only 1% higher for the Bristow and Campbell model ($p < 0.05, n = 126$), and 2% higher for the Hargreaves model ($p < 0.05, n = 126$). The mean bias error, relative mean bias error and the slope of linear regression were not statistically different in comparison to ground-measurement-based calibration for the Bristow and Campbell model. When our new solar radiation retrieval algorithm is used to estimate evapotranspiration, we found similar accuracies when using solar radiation input from ground- and auto-calibration. We conclude that our auto-calibration procedure results in accurate solar radiation retrievals, and requires only daily air temperature time series as input. The same procedure could easily be applied to other empirical solar radiation models.

---

†MeteoSwiss, Zürich, Switzerland, T: +41 44 256 93 56, F: +41 44 256 92 55, jedrzej.bojanowski@meteoswiss.ch
1 Introduction

Incoming global surface solar radiation is the main input for estimating evapotranspiration and the accumulation of plant biomass when simulating crop growth. High-quality solar radiation measurements are becoming increasingly available, but the network of measuring stations is too sparse for reliable spatial interpolation of measured values. In addition, prediction of crop yields requires solar radiation estimates at similar spatial and temporal resolutions as for other weather variables derived from global and regional circulation models. This cannot be achieved merely with ground solar radiation measurements.

Different approaches have been developed to estimate solar radiation, which can be grouped into: (1) physically-based radiative transfer models, (2) empirical models, based on the statistical relationship between measured meteorological variables and incoming solar radiation, and, more recently, (3) Bayesian neural network methods (Tymvios et al., 2005; Yacef et al., 2012). Physically-based radiative transfer models generally require a large number of input variables. Although empirical models are less data-demanding with respect to input variables, the accuracy of these models largely depends on reference solar radiation data, required to calibrate model coefficients. Accurate calibration can only be achieved for locations where solar radiation is measured. For locations without solar radiation measurements, model coefficients are interpolated (e.g. Bechini et al., 2000; Fodor and Mika, 2011; Van Kappel and Supit, 1998; Miller et al., 2008), which can result in larger errors (Bojanowski et al., 2013).

Most empirical models utilise the daily range between minimum and maximum air temperature. Clear days show a greater air temperature range: daytime air temperatures are high because clouds do not absorb incoming solar radiation; night-time air temperatures on the other hand are low because infrared radiation is emitted from the earth’s surface to the atmosphere and not radiated back by clouds. This relationship is however weaker during conditions of advection, which reduces the performance of air temperature-based empirical models in some regions for specific periods during the year. Despite this limitation, these empirical models have proved effective to accurately estimate solar radiation at several locations (e.g. Abraha and Savage, 2008; Bristow and Campbell, 1984; Grant et al., 2004; Hargreaves et al., 1985; Trnka et al., 2005).

Two common empirical solar radiation models based on daily air temperature range are the models proposed by Bristow and Campbell (1984) and Hargreaves et al. (1985). The Bristow and Campbell model exploits a saturation-type, exponential relationship between daily total solar radiation and daily air temperature range. In contrast, the Hargreaves model uses a linear relationship between solar radiation and the square root of daily air temperature range. These models are typically calibrated based on measured solar radiation, resulting in site-specific coefficients. Locations measuring solar radiation are relatively sparse over Europe, so approaches for calibrating the models for weather stations without solar radiation measurements would be useful.

The auto-calibration procedure, which we present in this paper, can be used for the calibration of air temperature-based solar radiation models without solar radiation measurements. The term ‘auto-calibration’ is used here to indicate a calibration without reference solar ra-
diation data. Our procedure is fundamentally based on the assumption formulated by Allen (1997) that on clear-sky days the model should approximate but not over-predict potential solar radiation. Potential radiation can easily be calculated for any location, as it is a function of location, day of the year, and atmospheric composition. Thus, the auto-calibration algorithm firstly identifies cloud-free days, and secondly optimises the model’s empirical coefficients to meet Allen’s assumption. This is done for a specific location using only a time series of daily air temperature ranges thereby allowing solar radiation to be estimated without prior calibration with measured solar radiation data.

The specific objectives of this study are: (1) to introduce a procedure to auto-calibrate empirical solar radiation models, (2) to evaluate the performance of the auto-calibration procedure based on the Bristow and Campbell, and Hargreaves air temperature-based models for European weather stations, and (3) to analyse how retrievals of solar radiation through auto-calibrated models affect the estimation of evapotranspiration. We hypothesise that solar radiation retrieved from auto-calibrated models has an accuracy similar to retrievals from models calibrated with measured solar radiation data, and consequently evapotranspiration estimates are also similar.

2 Data

2.1 Data from weather stations

Meteorological data were obtained from 126 weather stations. These data were extracted from the Joint Research Centre’s Monitoring Agricultural Resources Unit (JRC-MARS) database, which is the input for the MARS Crop Yield Forecasting System (Baruth et al., 2007; Boogaard et al., 2002). The stations range from latitude 34°N to 58°N, from longitude 9°W to 48°E and in altitude from −5 m to 1677 m. They are located in ten countries: France, Germany, Italy, the Netherlands, Poland, Portugal, Spain, Tunisia, Turkey, and the United Kingdom. Each station reported daily values of maximum (\(T_x\)) and minimum (\(T_n\)) air temperatures measured 2 metres above the ground, and daily solar radiation measurements (\(I_s\)) for at least 60% of the days during 2005–2010. All but six stations provided daily wind speed and water vapour pressure means, thus allowing for the estimation of evapotranspiration.

2.2 Meteosat Second Generation data

We used Meteosat Second Generation data to create two look-up tables to be used within the auto-calibration procedure: (1) a map of mean annual cloud fractional cover, and (2) a map of clear-sky atmospheric transmissivity. The data used to estimate cloud fractional cover are and for the clear-sky atmospheric transmissivity retrieval are described in the following subsections. It should be emphasized that current satellite data are not required for using the auto-calibration procedure, since, once calculated, two satellite-derived maps (look-up tables) are stored within the auto-calibration algorithm. Alternative methods to those used here exist to derive both of these maps. Clear-sky transmissivity may alternatively be estimated based on atmospheric composition (e.g. Fortin et al., 2008), while a map of mean cloud fractional cover can be derived from interpolation of values measured at weather stations. However, the MSG-derived estimates have the advantage of being spatially continuous and thus allow deriving the mean cloud fractional cover and clear-sky transmissivity maps without interpolating ground measurements.
2.2.1 Cloud fractional cover

To obtain a map of mean annual cloud fractional cover we used five years (2007–2011) of the monthly cloud fractional cover product provided by EUMETSAT’s Satellite Application Facility on Climate Monitoring (CM-SAF). The product is derived from 15-minute pixel-based cloud detection from visible and near-infra-red Meteosat Second Generation SEVIRI data (Derrien and LeGléau, 2005), developed by the Satellite Application Facility for the project ‘Support to Nowcasting and Very Short Range Forecasting’ (Geiger et al., 2008). The monthly cloud fractional cover is delivered by the CM-SAF at 15 km × 15 km resolution. For each pixel, we calculate the average of all monthly values to obtain the mean annual cloud fractional cover (in percent).

2.2.2 Surface solar radiation

To estimate clear-sky atmospheric transmissivity, we used six years (2005–2010) of the down-welling surface short-wave radiation flux (DSSF) daily product derived from the Meteosat Second Generation satellite data, for which the pixel size is 3 km at the equator and approximately 5 km in Central Europe (Geiger et al., 2008). The solar radiation product is generated and made freely available by the Land Surface Analysis Satellite Applications Facility (LSA-SAF) and further processed by the Flemish Institute for Technological Research (VITO) on behalf of the Joint Research Centre’s Monitoring Agriculture Resources (MARS) Unit. The solar radiation obtained for wavelengths between 0.3 μm and 4.0 μm is considered in this product. In the retrieval scheme used, the down-welling surface short-wave radiation flux is approximated by multiplication of the top-of-atmosphere solar radiation and the effective transmittance of the atmosphere or cloud atmosphere system (Geiger et al., 2008). The effective transmittance is calculated using two approaches depending on whether a given pixel is classified as clear or cloudy. This classification is done based on the 15-minute cloud mask, which monthly derivative is described in Section 2.2.1.

The approach used for the cloud-free pixels does not employ the Meteosat Second Generation satellite signal. The effective transmittance is estimated as a function of atmospheric composition following Frouin et al. (1989), employing the modelled water vapour column density taken from the European Centre for Medium-Range Weather Forecasts (ECMWF) and the amount of ozone retrieved from the Total Ozone Mapping Spectrometer (TOMS). The methodology applied for the cloudy pixels follows Brisson et al. (1999) and employs 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 radiation transfer in the cloud-atmosphere-surface system, assuming that the whole pixel is covered by a homogeneous cloud layer (Ineichen et al., 2009).

3 Methods

3.1 Solar radiation models based on daily air-temperature range

The two empirical air-temperature-based solar radiation models used in this study were the Bristow and Campbell model, and the Hargreaves model. Because most weather stations measure daily air-temperature range, these models can be easily applied for many stations. Bristow and Campbell model estimates solar radiation at the earth’s surface, $I_s$ (MJ m$^{-2}$ d$^{-1}$), as:

$$I_s = \tau_i I_x$$

(1)
where $I_x$ (MJ m$^{-2}$ d$^{-1}$) is daily extra-terrestrial solar radiation and $\tau_i(\%)$ is the daily atmospheric transmissivity for day $i$. The main difference compared with existing empirical solar radiation models is the way in which $\tau_i$ is estimated.

The daily extra-terrestrial solar radiation ($I_x$) is calculated in the same way by both solar radiation models. The method uses solar geometry (e.g. Campbell and Norman, 1998; Sellers, 1965; Spitters et al., 1986) and the solar constant ($S_0 = 4.921$ MJ m$^{-2}$ h$^{-1}$), which is corrected to reflect the actual Earth–Sun distance for a specific day of the year. To account for solar geometry changes during the day, a day is sub-divided into hourly units. Hourly extra-terrestrial radiation ($I_{x,hr}$) is calculated as:

$$I_{x,hr} = S_0 d \cos \theta_{hr}$$  \hspace{1cm} (2)

where $\theta_{hr}$ is the solar incidence angle relative to the normal to the land surface (Allen et al., 2006). The term $d$ expresses the correction factor for Earth–Sun distance which is calculated for a day of the year ($i$) as:

$$d = 1 + 0.0334 \cos(0.1721i - 0.0552)$$  \hspace{1cm} (3)

Daily extra-terrestrial solar radiation, $I_x$, is an integration of hourly values from sunrise to sunset.

Bristow and Campbell (1984) proposed an exponential relationship between solar radiation and daily air temperature range. In their approach $\tau_i$ is defined as:

$$\tau_i = \tau \left[1 - \exp \left(\frac{-B_b \Delta T_i^{B_c}}{\Delta T_m}\right)\right]$$  \hspace{1cm} (4)

where $\tau$ is the clear-sky transmissivity, $B_b$ and $B_c$ are the model coefficients and $\Delta T_m$ is the average monthly air temperature range. Clear-sky transmissivity is described in the next Section 3.2.

Hargreaves et al. (1985) proposed the following simple linear equation between $I_s$ and daily air temperature range ($\Delta T_i$ ($^\circ$C)):

$$I_s = I_x H_a \sqrt{\Delta T_i} + H_b$$  \hspace{1cm} (5)

where $H_a$ and $H_b$ are site-specific empirical coefficients. The air temperature range for day $i$ used in the Hargreaves model (Eq. (5)) is calculated as the difference between maximum day-time ($T_x$ ($^\circ$C)) and minimum night-time ($T_n$ ($^\circ$C)) air temperature:

$$\Delta T_i = T_x(i) - T_n(i)$$  \hspace{1cm} (6)

Relating effective transmissivity to a simple difference between maximum and minimum air temperature may be adequate for sites that are not significantly affected by advection (Bristow and Campbell, 1984). The advection of cold or warm air masses can result in $\Delta T_i$ not reflecting the impact of cloud cover on incoming radiation during the day and outgoing radiation during the night. Therefore, the occurrence of advection can reduce the accuracy of air-temperature-based model estimates (Winslow et al., 2001).

To account for this limitation, Bristow and Campbell (1984) proposed the following adapted equation for $\Delta T_i$:

$$\Delta T_i = T_x(i) - \frac{T_n(i) + T_n(i+1)}{2}$$  \hspace{1cm} (7)

The mean of the two consecutive daily minimum air temperatures for days $i$ and $i+1$ is used to reduce the effect of large-scale hot or cold air masses arriving at the measurement location.
(Bristow and Campbell, 1984). Without this adaptation, incoming warm air masses can result in overestimation and cold air masses in underestimation of solar radiation.

The Hargreaves model requires calibration of both empirical coefficients $H_a$ and $H_b$ (Eq. (5)), whereas the Bristow and Campbell model requires calibration only of the $B_b$ coefficient (Eq. (4)). Several authors have demonstrated that the Bristow and Campbell model has the same performance using a constant $B_c$ coefficient, which can be 1.5 (Coops et al., 2000; Thornton and Running, 1999), 2.0 (Abraha and Savage, 2008), 2.02 (Grant et al., 2004) or 2.4 (Bandyopadhyay et al., 2008). In this study we used a value of 2.0 for $B_c$.

3.2 Clear-sky transmissivity

The clear-sky transmissivity ($\tau$) is the ratio between solar radiation measured at ground level during cloud-free days and extra-terrestrial radiation ($I_x$), and is a function of atmospheric turbidity. The most accurate approach to calculate $\tau$ makes direct use of measured solar radiation, as in Allen (1996) and Allen et al. (2005). Although we applied the same approach, instead of using solar radiation measurements from stations, we used the solar radiation estimates from Meteosat Second Generation to calculate site-specific clear-sky transmissivity values that are stable over time. The main advantage is the availability of these data for large regions, rather than only a limited number of stations. We first calculated, for every day during 2005–2010, the atmospheric transmissivity as the ratio between (1) the down-welling surface short-wave radiation flux (DSSF) product derived from the Meteosat Second Generation satellite data (Section 2.2.2) and (2) extra-terrestrial solar radiation (Eq. 2). Then, for each pixel and year we retained three per cent of the days with the highest atmospheric transmissivity, assuming those days were cloudless. The median of these three-per-cent values was taken as the value for the clear-sky transmissivity of that pixel. We then created a map of clear-sky transmissivity containing a single transmissivity value per pixel. The map is used as a static element in the auto-calibration procedure and is not updated during auto-calibration.

3.3 Auto-calibration procedure

In this study we present a procedure for auto-calibration of the $B_b$ coefficient of the Bristow and Campbell model (Eq. 4) and the $H_a$ and $H_b$ coefficients of the Hargreaves model (Eq. 5). We used this procedure to derive model coefficients for each individual weather station of the 126 used in this study based on air temperatures from the period 2005–2007. In this section we describe the main steps for implementing the new procedure following the flowchart shown in Fig. 1.

Our procedure assumes that during cloudless conditions, solar radiation cannot exceed the daily potential solar radiation ($I_{pot}$), which is defined as:

$$I_{pot} = \tau I_x$$

(8)

Following this equation (8) we calculated the daily potential solar radiation for each Meteosat pixel from our clear-sky transmissivity estimates (Section 3.2) and the daily extra-terrestrial solar radiation (Section 3.1). Subsequently, we estimated daily solar radiation ($I_s$) from the solar radiation models using the following default values for the model coefficients: $B_b = 0.12$ for the Bristow and Campbell model and $H_a = 0.16$ and $H_b = 0.10$ for the Hargreaves model. The default Hargreaves model coefficients were taken as the median values for Europe based on the maps published in Bojanowski et al. (2013). We repeated the procedure applied in that study to derive the median value of the Bristow and Campbell $B_b$ coefficient. Then we calculated $\Delta I$ – the absolute difference between potential ($I_{pot}$) and estimated ($I_s$) daily solar
Clear-sky transmissivity ($\tau$)

Look-up table

$I_{pot} = \tau I_x$

Hargreaves model:

$I_s = 0.16I_x\sqrt{\Delta T_i} + 0.1$

Bristow & Campbell model:

$I_s = I_x\tau \left[ 1 - \exp \left( -0.12\tau^2 \right) \right]$

Input data

Extra-terrestrial radiation ($I_x$)

Input data

Air temperature range:

Hargreaves model:

$\Delta T_i = T_{i(i)} - T_{n(i)}$

Bristow & Campbell model:

$\Delta T_i = T_{x(i)} - \frac{T_{n(i)} + T_{n(i+1)}}{2}$

Air temperature range:

Input data

Distance to the coast ($dc$)

Input data

$\Delta I = I_{pot} - I_s$

Select $P$ percent of days with the lowest $\Delta I$

$P = \begin{cases} 
30, & C < 25 \\
-68\ln C + 0.92C + 225, & 25 \leq C \leq 75 \\
1, & C > 70 
\end{cases}$

Mean annual cloud fractional cover ($C$)

Look-up table

$\epsilon = \begin{cases} 
0.1, & \text{if } dc < 15 \text{ km} \\
0.5, & \text{if } dc \geq 15 \text{ km} 
\end{cases}$

Least Square Regression based on selected days:

Hargreaves model:

$I_{pot} = I_xH_a\sqrt{\Delta T_i} + H_b$

Bristow & Campbell model:

$I_{pot} - \epsilon = I_{pot} \left[ 1 - \exp \left( -B_b\Delta T^2 \right) \right]$

Output

Hargreaves: $H_a$ and $H_b$

Bristow & Campbell: $B_b$

Figure 1: Flowchart of the auto-calibration method for the daily air-temperature-based solar radiation models applied to the 126 weather stations over Europe.
radiation. The days with differences closest to 0 were then selected as potentially cloud-free. However, the average number of clear-sky days per year varies among locations. Thus, we used the map of the annual mean cloud fractional cover (Section 2.2.1) as a proxy for the number of days to be selected as cloud-free for a given location (defined by $P(\%)$). An empirically-derived logarithmic function was chosen to link $P$ to the annual mean cloud fractional cover ($C(\%)$). The Hargreaves and Bristow and Campbell models (tested for 2005–2007) gave the lowest relative root mean square error (RRMSE) when the function was formulated as follows:

$$
P = \begin{cases} 
30, & C < 25 \\
-68 \ln C + 0.92C + 225, & 25 \leq C \leq 75 \\
1, & C > 70
\end{cases}
$$

(9)

For locations with a high number of clear-sky days, where $C$ is below 25%, $P$ is set to 30%. This limits the number of days used to derive the model coefficients thus avoiding overestimation of solar radiation, which can be caused by erroneous selection of a cloudy day as cloud-free. Similarly, for locations with a low number of clear-sky days, where $C$ is above 70%, $P$ is set to 1%. This ensures that the number of days used to derive the model coefficients is not too low. We assume that at every location at least 4 days ($\approx 1\%$) per year are cloudless; this should reduce the possibility that a selected day is in fact a cloudy day.

Next, we perform a non-linear least squares regression implemented in the R-package sirad (Bojanowski, 2013; R-project, 2013) to derive coefficients for the solar radiation models. This regression method determines the model coefficients providing 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. We assume that the reference solar radiation data (dependent variable) are equal to the $I_{pot}$ for selected clear-sky days, while the air temperature range is the independent variable. This relationship was directly used as the basis of the least square regression for the Hargreaves model. The Bristow and Campbell model, however, needs an additional constraint. By definition, during clear-sky days, $\tau_i$ is equal to the clear-sky transmissivity ($\tau$). Using the equation of the Bristow and Campbell model (Eq. 4), atmospheric transmittance cannot equal the clear-sky transmissivity. To account for this when auto-calibrating the Bristow and Campbell model, the dependent variable for the least square regression is equal to $I_{pot}$ reduced by a constant value ($\epsilon$). We have tested the model for every weather station for a period 2005–2007 using different values for $\epsilon$. The best performance of the model (lowest RRMSE) was reached for $\epsilon$ equals 0.1 MJ m$^{-2}$ for locations less than 15 km from the coast and 0.5 MJ m$^{-2}$ for locations more distant from the coast.

### 3.4 Evaluation of the auto-calibrated models

We applied the auto-calibration procedure to each individual station using the daily air temperature data for the period 2005–2007. This resulted in a set of station-specific model coefficients for Bristow and Campbell, and Hargreaves solar radiation models. In addition, we obtained a reference dataset of model coefficients through normal ‘ground-based’ calibration using station air temperature and solar radiation measurements for the period 2005–2007. For the ground-based calibration, we used a least square regression in which the measured solar radiation data were used as the dependent variable.

The second half of the series (2008–2010) was reserved for evaluating the two solar radiation models using both the ground-calibrated and auto-calibrated coefficients. Thus, for each station the model coefficients were applied to the air temperature time series from 2008–2010, resulting in solar radiation time series per station, per model and per calibration method. These time series were compared against the measured solar radiation for the selected stations.
We assessed model performance with six estimators: the mean bias error (MBE), the relative MBE (RMBE), the root mean square error (RMSE), the relative RMSE (RRMSE), the slope of the linear regression and the modelling efficiency (EF). Negative EF value indicates that the average value of the measurements gives a better estimate than the simulated values. A perfect fit between observed and simulated data would be indicated by: MBE = 0 MJ m$^{-2}$, RMBE = 0%, RMSE = 0 MJ m$^{-2}$, RRMSE = 0%, slope = 1 and EF = 1.

3.5 Simulation of evapotranspiration

We used measured and modelled solar radiation together with measured maximum and minimum air temperatures, water vapour pressure and wind speed to simulate grass reference evapotranspiration ($ET_0$). Daily $ET_0$ was calculated based on the Penman-Monteith equation implemented in the R-package sirad (Bojanowski, 2013; R-project, 2013). The implementation follows the procedure presented by Allen et al. (1998).

For each station, we simulated short-grass reference evapotranspiration five times. For each simulation, we only changed the daily solar radiation ($I_s$), i.e. measured (as the reference) or modelled with the Bristow and Campbell model and the Hargreaves model using the ground-calibration or auto-calibration procedure. We compared the $ET_0$ simulated using modelled solar radiation with the $ET_0$ simulated using measured solar radiation for the years 2008–2010, which were not used for calibration of the solar radiation models. We used the same performance statistics as listed in Section 3.4 to assess the performance of the differently calibrated solar radiation models regarding their ability to simulate short-grass reference $ET_0$.

4 Results

The auto-calibrated models performed with a lower but very similar accuracy compared with the models calibrated using the ground measurements. The mean RRMSE for the auto-calibrated Bristow and Campbell model was 29.04±6.74%, while for ground-based calibrated 29.82±6.80%. The same statistics for Hargreaves model were 28.50 ± 5.99% and 30.75 ± 7.36% respectively. Even if the auto-calibrated models gave on average slightly higher errors, their great advantage is that the model coefficients can be retrieved for every location. This is not possible for ground-based calibration that requires solar radiation measurements as input.

Despite the overall similar accuracy, the use of auto-calibrated model coefficients significantly ($p < 0.05$) decreased the performance of the Bristow and Campbell solar radiation model in comparison to ground-based calibration for three of the six statistics (MBE, RMBE and the slope of linear regression), while the corresponding decrease for the Hargreaves model was significant for all statistics (Table 4). The distribution of the performance statistics among weather stations confirms that our auto-calibration procedure performs more accurately for the Bristow and Campbell model (Fig. 2) than for the Hargreaves model (Fig. 3).

The spatial distribution of the RRMSE shows almost identical behaviour of the two calibration methods for both the Bristow and Campbell (Fig. 4b and d) and Hargreaves models (Fig. 5b and d). Weather stations where MBE visibly increases using auto-calibration are located in the southern countries for the Bristow and Campbell model (Fig. 4a and c). For the Hargreaves model (Fig. 5a and c), the MBE following auto-calibration is higher for most of the weather stations.

Most stations with an outlying MBE value above 2 MJ m$^{-2}$ (Fig. 4c) are those with the greatest difference between clear-sky transmissivity derived from ground measurements and that obtained from Meteosat Second Generation data (Fig. 6). This difference in clear-sky...
Table 1: Performance statistics for the two solar radiation models (Bristow and Campbell, and Hargreaves) for both calibration methods used (ground-based and auto-calibration). Numbers highlighted with **bold font** indicate that the difference between means is statistically significant ($p < 0.05$) for the two-sided *t*-test for paired samples.

<table>
<thead>
<tr>
<th>Model</th>
<th>Calibration Method</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bristow-Campbell</td>
<td>Ground-calibration</td>
<td>0.16</td>
<td>0.52</td>
<td>-0.07</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>Auto-calibration</td>
<td>-1.69</td>
<td>4.30</td>
<td>-1.18</td>
<td>7.93</td>
</tr>
<tr>
<td>Hargreaves</td>
<td>Ground-calibration</td>
<td>3.34</td>
<td>0.57</td>
<td>3.44</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>Auto-calibration</td>
<td>29.04</td>
<td>6.74</td>
<td>29.82</td>
<td>6.80</td>
</tr>
<tr>
<td></td>
<td>Ground-calibration</td>
<td>29.04</td>
<td>6.74</td>
<td>29.82</td>
<td>6.80</td>
</tr>
<tr>
<td></td>
<td>Auto-calibration</td>
<td>28.50</td>
<td>5.99</td>
<td>30.75</td>
<td>7.36</td>
</tr>
<tr>
<td>Slope</td>
<td>Ground-calibration</td>
<td>0.99</td>
<td>0.06</td>
<td>0.99</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>Auto-calibration</td>
<td>0.84</td>
<td>0.09</td>
<td>0.86</td>
<td>0.06</td>
</tr>
<tr>
<td>EF</td>
<td>Ground-calibration</td>
<td>0.82</td>
<td>0.08</td>
<td>0.81</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Auto-calibration</td>
<td>0.83</td>
<td>0.06</td>
<td>0.80</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Figure 2: Box plots of the performance estimators of the Bristow and Campbell model for ground- and auto-calibration evaluated against measured solar radiation data. Boxes denote 1st and 3rd quartiles (with thick horizontal median line). Whiskers indicate largest and lowest values within 1.5 times the interquartile range, while circles represent values beyond this.
Figure 3: Box plots of the performance estimators of the Hargreaves model for ground- and auto-calibration evaluated against measured solar radiation data. Boxes denote 1st and 3rd quartiles (with thick horizontal median line). Whiskers indicate largest and lowest values within 1.5 times the interquartile range, while circles represent values beyond this.
Figure 4: Relative root mean square error (RRMSE) and mean bias error (MBE) for the Bristow and Campbell model based on evaluation of the model against measured data: ground-calibration (top row) and auto-calibration (bottom row).
Figure 5: Relative root mean square error (RRMSE) and mean bias error (MBE) for solar radiation obtained with the Hargreaves model using ground- and auto-calibrated coefficients, as evaluated against measured data: ground-calibration (top row) and auto-calibration (bottom row).
transmissivity resulted in the maximum potential solar radiation of the clear-sky days (Eq. 8) used for model calibration being different from the equivalent maximum values of the measured solar radiation.

Fig. 7 and 8 show scatter plots of solar radiation, where each point represents a single solar radiation value for each weather station and day. Fig. 7a and 8a reflect the exponential nature of the Bristow and Campbell model and the linear nature of the Hargreaves model. For the Hargreaves model, points representing one location are situated along a straight line, as data at both axes are calculated with linear equations (Fig. 8a). For the Bristow and Campbell model, points representing one location are situated along a non-linear line, as data at both axes are calculated with non-linear (exponential) equations (Fig. 7a). Scatter plots comparing measured against estimated solar radiation with auto-calibrated or ground-calibrated models show similar performance by the Bristow and Campbell (Fig. 7b and c) and Hargreaves models (Fig. 8b and c).

4.1 Evapotranspiration simulation

Estimates of $ET_0$ were significantly more accurate, for three of the six statistics (MBE, RMBE and the slope of linear regression), using solar radiation retrieved from the ground-calibrated Bristow and Campbell model as compared to the auto-calibrated retrievals. The average RMSE, RRMSE and EF presented in Table 4.1 demonstrate a statistically significant difference ($p < 0.05$). The MBE does not show any difference for the use of solar radiation derived from both calibrations.

The accuracy of the $ET_0$ estimation using solar radiation derived from the Hargreaves model differs depending on the model calibration method. For all the statistics (Table 4.1), the perfor-
Figure 7: Daily solar radiation estimates from the Bristow and Campbell model: (a) auto-calibration vs ground-based calibration, (b) auto-calibration vs measured solar radiation, and (c) ground-based calibration vs measured solar radiation. Lines represent: 1:1 (solid) and linear regression (dashed).
Figure 8: Daily solar radiation estimates from the Hargreaves model: (a) auto-calibration vs ground-based calibration, (b) auto-calibration vs measured solar radiation, and (c) ground-based calibration vs measured solar radiation. Lines represent: 1:1 (solid) and linear regression (dashed).
performance changed significantly \((p < 0.05)\) when using the auto-calibrated model to estimate solar radiation. All the statistics except slope indicate that the performance of the Penman-Monteith model decreased when using solar radiation derived from the auto-calibrated Hargreaves model. In contrast, the slope indicates an improvement in the mean value from 0.96 to 0.99.

No clear spatial pattern is apparent for weather stations with the greatest difference in the RRMSE between the two calibrations of the Bristow and Campbell model. For the Hargreaves model, stations with the greatest reduction in accuracy between both calibrations are located at latitudes north of 50\(^\circ\)N.

## 5 Discussion

The auto-calibration procedure presented here can calculate the coefficients of the solar radiation models for every location where daily air temperature range is measured, and it does not require solar radiation measurements that are only obtained at a relatively limited number of stations. The accuracy when applying these coefficients is lower than the accuracy obtained using ground calibration, i.e. through a time series of measured solar radiation. For ground calibration, spatial interpolation of the calibrated model coefficients is required in order to apply the models elsewhere. In comparison, our auto-calibration approach extracts site-specific model coefficients for any station that measures the daily air temperature range. In Europe alone, at least 8000 weather stations (JRC-MARS database) report daily air temperature range since 2005. The only additional input required for calibration is a single estimate of clear-sky transmissivity and cloud fractional cover. We showed that these can be simply constructed from Meteosat time series, although they may also be derived from other sources.

The accuracy of the auto-calibrated Hargreaves model \((RRMSE = 30.75 \pm 7.36\%\); \(MBE = 0.70 \pm 1.03 \, \text{MJ m}^{-2}\)) can be compared to previous studies that applied this model in Europe without prior ground-based calibration. Bojanowski et al. (2013) used the Hargreaves model with coefficients calculated based on ground measurements and then interpolated (as proposed by Supit and Van Kappel, 1998; Van Kappel and Supit, 1998). The Hargreaves model employing these coefficients achieved a RRMSE of 31.7 \(\pm 7.2\%\) and a \(MBE\) of \(-0.9 \pm 1.3 \, \text{MJ m}^{-2}\). In the same study, the Hargreaves model was calibrated for each weather station using Meteosat Second Generation data instead of ground measurements. In this way, an average RRMSE of 29.9 \(\pm 6.3\%\) and \(MBE\) of 0.1 \(\pm 0.8 \, \text{MJ m}^{-2}\) were achieved for solar radiation estimates. Bojanowski et al. (2013) provide only station-specific model coefficients for a limited number
of stations. For other stations, a new calibration with Meteosat Second Generation data would be needed. This is not a crucial limitation because these data are freely available. The R-package sirad (Bojanowski, 2013; R-project, 2013) provides functions for the calibration of solar radiation models based on Meteosat Second Generation data. Since downloading the satellite data and deriving the coefficients still need time and expertise, Bojanowski et al. (2013) provide a map of Hargreaves model coefficients interpolated from the locally derived coefficients based on Meteosat Second Generation data. Employing these maps, the Hargreaves model achieved a RRMSE of 31.3 ± 7.0% and a MBE of 0.2 ± 1.4 MJ m$^{-2}$, hence a slightly higher RRMSE and lower MBE than when using the auto-calibrated coefficients.

Comparison of the performance of the calibration methods discussed above shows, in agreement with Bojanowski et al. (2013), that the use of locally-derived coefficients with Meteosat Second Generation data is to be recommended. However, the advantage of the auto-calibration procedure presented here is its simplicity and ability to estimate (using the proposed software implementation) solar radiation from air-temperature data without additional data other than location, mean annual cloud cover fraction, and clear-sky transmissivity. For Europe, the last two are already incorporated in the software implementation (Bojanowski, 2013; R-project, 2013). This can be an advantage when applying the auto-calibration procedure to air temperature time series generated by general or regional circulation models, ensuring consistency between air temperature and solar radiation data. In that case, the resulting solar radiation estimates can be used for example for crop yield prediction under climate change scenarios.

The auto-calibration procedure could potentially be adapted to other solar radiation models, not only air-temperature-based models. Empirical models that estimate solar radiation using sunshine duration (e.g. Ångström, 1924; Prescott, 1940), cloud coverage and air temperature (e.g. Supit and Van Kappel, 1998) or air temperature and precipitation (e.g. Wu et al., 2007) could be calibrated using the same procedure. This would require a network of weather stations with time series (> 5 years) data for these parameters. Such stations exist in many European countries, as well as worldwide. Auto-calibration of other models of this kind would be relevant, because their reported accuracy is frequently greater (Mavromatis, 2008; Trnka et al., 2005; Weiss and Hays, 2004), even though fewer stations measure parameters such as sunshine duration and cloud coverage. As for air temperature, these auto-calibrated models could also be used to link solar radiation to general circulation model outputs.

We tested the auto-calibration procedure only for European weather stations. Further study for other continents could reveal whether the presented procedure can be applied with a similar accuracy in other climatological zones. The required estimates of the clear-sky transmissivity as well as the cloud fractional cover could be derived from other geostationary satellites (Thies and Bendix, 2011) such as: the American Geostationary Operational Environmental Satellites (GOES), the Indian National Satellite (INSAT), the Chinese Fèngyīn and the Japanese Geostationary Meteorological Satellite (HIMAWARI).

6 Conclusions

We demonstrated that daily solar radiation can be estimated from daily air temperature range measurements without previously derived site-specific empirical coefficients. As hypothesised, the accuracy of the auto-calibrated Bristow and Campbell and Hargreaves models is slightly reduced as compared to the use of conventional ground-based calibration. However, for the Bristow and Campbell model the decrease in model performance is not statistically significant for three of the six performance statistics tested here. The Bristow and Campbell model employing our proposed auto-calibration procedure provides more accurate solar radiation estimates than the Hargreaves model when site-specific empirical coefficients derived from ground measure-
ments are not available. The accurate simulation of the daily $ET_0$ based on the solar radiation derived from the auto-calibrated models presented here demonstrates its applicability for crop growth modelling.

We conclude that our auto-calibration procedure results in accurate solar radiation retrievals, and requires only the availability of air temperature time series. These are available for many weather stations, or alternatively are estimated by general circulation models. The same procedure could easily be applied to other empirical solar radiation models that include other parameters besides air temperature.

References


Van Kappel, R. R. and 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.


