Effective Accuracy of Satellite-Derived Hourly Irradiances

A. Zelenka\textsuperscript{1}, R. Perez\textsuperscript{2}, R. Seals\textsuperscript{2}, and D. Renne\textsuperscript{3}

With 6 Figures

Received August 15, 1997
Revised March 4, 1998

Summary

Estimates of hourly global irradiance based on geostationary satellite data with a resolution of several (2 to 10) kilometres reproduce ground-measured values with a Root Mean Square Error (RMSE) of typically 20\% to 25\%. The different components of this RMSE have been enumerated by several authors but, due to the lack of adequate measurements, their respective importance is not well settled. In the present study we attempt to quantify these components from a practical point of view, that is from the point of view of users having to rely on time/site specific irradiance data. We conclude that the intrinsic, or "effective" RMSE is more along the line of 12\%. This effective RMSE is the measure of the methodological imprecision (satellite-to-irradiance conversion models). The remaining part of the overall RMSE is the amount by which spatially averaged satellite-derived estimates are, by their very nature, bound to differ from ground-measured local insolation.

1. Introduction

Hourly global irradiance estimates derived from satellite visible-channel data are reported to have an RMSE of typically 20 to 25\% (80–100 \text{Wm}^{-2} \text{hr}^{-1}) when compared with ground-based, single-site or regional network measurements (Schmetz, 1989; Hay, 1993; Noia et al., 1993; Bayer et al., 1996).

This overall RMSE originates from three basic groups: Group (1) encompasses both the simplifying assumptions in the models converting the satellite radiometric response into ground irradiance (essentially for retrieving irradiances from radiances and transforming narrowband, spectrally-filtered into broadband, unfiltered instrumental responses) and the satellite measurement errors. Group (2) consists of the errors inherent to pyranometric measurements, while group (3) comprises the difference between instantaneous pixels extended in space and single point measurements integrated in time. Geoverification, i.e. the correct identification of the pixels containing the measuring sites, as well as cloud shadow effects at low solar elevations, also belong to this third group.

The first group constitutes the satellite's (modelling) intrinsic error, while the second and third group constitute intercomparison errors.

Considering the hourly RMS difference between neighbouring ground measurements as a function of distance (Fig. 1), we note that the 20–25\% level is reached within only 20–30 km (Perez et al., 1997). This means that for any application requiring time/site specific data, the user should rely on the satellite rather than on a "neighbouring" ground station if the latter operates further away than 20–30 km from the site.

Another conspicuous feature of the RMS difference versus distance diagram (the raw variogram in terms of geostatistics) is the sharp rise near the origin: within the first ten kilo-
metres, the RMSE reaches 17%. This region is shaped by the micro-variability of the irradiation field and by measurement uncertainties, which are both components of the overall satellite estimation RMSE of 20–25%. Measurement errors for properly maintained solar radiation networks are generally accepted to be 3 to 5%. In order to quantify the impact of the space/time sources of error mentioned above (error group 3) and, thus, to isolate the intrinsic error of satellite prediction (group 1), knowledge of the micro-variability of the irradiation field would be required. The meso-scale variability is well documented albeit mostly for daily sums or longer integration times. Dugas and Heuer (1985) and Hay and Hanson (1985) summarize the knowledge in this field and the latter authors already note that “for the short(er) time intervals solar radiation measurements provide little more than point estimates”. Reliable measurements at the microscale are not commonly available at present, but even if there were some, it could be that the shape of the variogram near the origin would still remain elusive due to the discontinuous and fractal nature of cloud patterns on an hourly time scale (Perez et al., 1993).

Therefore, we develop an alternative strategy to set a reliable figure on the “effective” RMSE. We base our considerations on one year of operation with a simple model based on GOES-8 measurements applied to the Northeastern US. This model is described in Section 3, after presentation of the ground and satellite data in Section 2. Section 4 describes the strategy and reports the evidence that can be gathered, among others, about the influence of the micro-variability on the satellite estimates. This allows us to conclude (Section 5) that, from the users’ point of view, the estimations are definitely more accurate than suggested by the 20–25% RMSE.

2. Experimental Data

2.1 Surface Network Irradiation Measurements

The northeastern U.S. network consists of twelve stations (Fig. 2). Stations include the State University of New York Albany research site, with WMO Class 1 instrumentation, and eleven rotating shadowband radiometers (RSR). The eleven RSRS had been deployed to serve the research needs of the electric utility industry in southern New York and Massachusetts. Thus, their deployment is not ideal from a geographical standpoint. Data from the network were processed and analysed by the authors (Perez et al., 1997). This process comprised automatic and visual data quality control as well as calibration.
including, when applicable, a posteriori calibration drift correction.

The Swiss automatic meteorological network ANETZ consists of 72 stations, 67 of which are equipped with Kipp & Zonen CM6 pyranometers (Fig. 3). The network operates in an unattended fashion. Each site is calibrated once a year. Hourly sums are scrutinized, but the 10-minute sums (the unity integration time) are raw data (Zelenka, 1984). Hence, in this study, implausible 10-minute values have been rejected. The RMS difference versus distance investigation is limited to 31 sites with altitude below 800 m. Site pairs separated by the main crests of the Alps were discarded. The data displayed in Figs. 1 and 6 are 6-year averages (1985–1990) for the month of July which shows a low variability across the country. This choice helps keep undesired “noise” as low as possible for the benefit of the study.

Data from the NE US network are used both for satellite model ground truth and for analysing geostatistical features of regional solar radiation fields. Data from the Swiss network are used only for the latter purpose.

2.2 Satellite Data

The satellite data consist of intermediate resolution images from the visible channel of NOAA's geostationary weather satellite GOES-8. Ground resolution of these images in the study region is of the order of 10 km latitude by 13 km longitude.

Navigated images covering the North American continent and the Atlantic Ocean are distributed on an hourly basis through the Internet Data Distribution System (1995) to participating US Universities. A portion of these images, covering the mid-Atlantic region of the US, has been locally archived since May 1995. Based on our processing of thousands of images, typical navigation errors are of the order of 2 to 3 km.

3. The Operational Satellite-to-Irradiance Conversion Model

3.1 Formulation

Retrieval techniques of surface shortwave irradiance from satellite observations (hereafter referred to as “satellite models”) rely on the almost linear coupling that exists between irradiance at the top of the atmosphere and at the surface. Straightforward exploitation of this coupling leads to (Eq. (9) of Schmetz (1989))

$$
\tau_{\text{atm}}(n) = \tau_{\text{atm}}(0)(1 - n)
$$

where $\tau_{\text{atm}}(n)$ and $\tau_{\text{atm}}(0)$ are the atmospheric transmittances for the cloudy and clear sky, respectively. The cloudiness is expressed with the cloud index $n$ (Cano et al., 1986)

$$
\alpha_{\text{toc}} = \alpha_{\text{toc min}} / (\alpha_{\text{toc max}} - \alpha_{\text{toc min}})
$$

where $\alpha_{\text{toc}}$ is the instantaneous planetary albedo, while $\alpha_{\text{toc min}}$ corresponds to a clear, clean and dry sky, and $\alpha_{\text{toc max}}$ corresponds to a heavily

Fig. 3. Spatial distribution of measuring sites below 1000 m altitude in the Swiss network. Sites extend from 6.13°E (Geneva, GVE) to 9.53°E (Chur, CHU) and 45.85°N (Stabio, SBO) to 47.68°N (Schaaffhausen, SHA). The smallest distances are located in the Zurich area (5.7 km between Kloten, KLO and Reckenholz, REH) and in the South (7.1 km between Locarno, LOC and Magadino, MAG).
overcast sky. These planetary albedos are radiant flux densities (normalized irradiances), whereas the satellite observes spectrally filtered irradiances. Retrieval of irradiances from irradiances requires correction for spectral conversion and anisotropy of backscattering processes as functions of observing direction and illumination angle. This "viewing geometry" must also be standardized to allow comparisons between widely spaced scenes observed at different times. Thus, the satellite-observed irradiances need careful correction and normalization, as explained below, before their use in the calculation of the cloud index $n$.

3.2 Clear Sky Transmissivity

We select the model of Kasten et al. (1984), which has been derived from 10 years of observations by the radiometric network of the German Weather Service:

$$\tau_{\text{sun}}(0) = 0.84\exp(-0.0277 T_L m)$$  \hspace{1cm} \text{(3)}

An altitude-dependent version of this model also exists (Kasten, 1990). Besides the relative air mass $m$, equation (3)'s only variable is the Linke turbidity factor $T_L$. For the temperature mid-latitudes, monthly averages of $T_L$ listed by Kasten et al. (1984, Table 1), or interpolation between these averages according to the day in the year, lead to good results (Davies et al., 1988). Alternatively, more accurate regional values may be used if available (as in the present study, where long term measurements allow determination of seasonal values). If these values are not available, then the WMO (1981) formulae giving $T_L$ as a function of climatological parameters prove effective. Note that with this $\tau_{\text{sun}}(0)$, the resulting satellite model may be considered as "physical" rather than "empirical" since it does not rely on a correlation between satellite response and specific ground measurements.

3.3 Cloud Index

The relation between the radiance observed by the satellite, $L_{\text{sat}}$, and the planetary albedo is discussed by, e.g., Koepeke and Kriel (1987). A discussion more closely related to the cloud index is given by Moussu et al. (1989). Here, we follow the practical approach adopted by Beyer et al. (1996) which adheres to the first principles underlying Eqs. (1) and (2). Since the cloud index $n$ is a quotient of planetary albedos $\alpha_{\text{toa}}$ all the proportionality constants between $L_{\text{sat}}$ and $\alpha_{\text{toa}}$ cancel. In particular, the calibration factor converting the satellite counts $C_{\text{sat}}$ (or $\sqrt{C_{\text{sat}}}$ in the case of GOES-8) into $L_{\text{sat}}$ does not need to be known. Thus, introducing a Lambertian normalized satellite count

$$CC = C_{\text{sat}}/(r^2 \cos \theta_z)$$  \hspace{1cm} \text{(4)}

leaves us with

$$n = (CC - g_{\text{min}} CC_{\text{min}})/(g_{\text{max}} CC_{\text{max}} - g_{\text{min}} CC_{\text{min}})$$  \hspace{1cm} \text{(5)}

where $\theta_z$ is the solar zenith angle, $r^2$ is the Sun-Earth distance modulating factor, and $g, g_{\text{max}}$ and $g_{\text{min}}$ are functions introduced to account for non-Lambertian effects. These effects arise from readily-observed scattering mechanisms. Comparing ground measurements, when available, with irradiance components modelled (Perez et al., 1990) from satellite-derived global irradiance has also proven beneficial to identify the cause of systematic non-Lambertian optical effects. In fact, only two such effects prove to leave a conspicuous signature. These effects are the clear atmosphere air mass effect and the back-scattering effect.

3.4 Clear Atmosphere Air Mass Effect

Figure 4 displays one year of normalized counts (Eq. 4) for the pixel covering the ASRC site at Albany, N.Y., as a function of the solar zenith

<table>
<thead>
<tr>
<th>Month</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
<th>XI</th>
<th>XII</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kasten $T_L$</td>
<td>3.8</td>
<td>4.2</td>
<td>4.8</td>
<td>5.2</td>
<td>5.4</td>
<td>6.4</td>
<td>6.3</td>
<td>6.1</td>
<td>5.5</td>
<td>4.3</td>
<td>3.7</td>
<td>3.6</td>
</tr>
<tr>
<td>NE US $T_L$</td>
<td>1.9</td>
<td>2.1</td>
<td>2.5</td>
<td>3.2</td>
<td>3.9</td>
<td>5.1</td>
<td>5.9</td>
<td>4.4</td>
<td>3.0</td>
<td>2.4</td>
<td>2.1</td>
<td>2.0</td>
</tr>
</tbody>
</table>
angle $\theta_z$. Not unexpectedly, $CC_{\text{max}}$ is very stable except for the few outliers corresponding to skewed cloud geometries. Thus, $g_{\text{max}}$ is equal to one in Eq. (5). The count $CC_{\text{min}}$ also behaves stably up to about $\theta_z = 70^\circ$ but then it increases with decreasing solar elevation. This behaviour is reminiscent of Lacis and Hansen (1974) $g_{\text{max}}$ for a Rayleigh atmosphere above a black earth, namely $g_{\text{max}}(\theta_z) = 0.28/(1+6.43\cos\theta_z)$.

The correction $g_{\text{min}}$, necessary to restore the full dynamic range of $n$, therefore follows from a simplified calculation of the planetary albedo for slanting incidence, where the atmosphere continues to scatter while the ground progressively dims. From a simple box model and a direct application of Kasten's clear sky formula (Eq. 3), assuming a 15%–85% diffuse-direct split at normal solar incidence, this correction may be expressed as:

$$g'_{\text{min}} = 0.85 + 0.15\left(\cos\theta_z + \sin\theta_z\exp[0.04(1 - m)]\right)/\left[\cos\theta_z(\cos\theta_z + \sin\theta_z)\right]$$

(6)

3.5 Back-Scattering Effect

Figure 5 displays the same counts as a function of the Sun-to-satellite angle $\psi$. Again, there is no evidence for any dependency of $CC_{\text{max}}$ on this angle, while for $CC_{\text{min}}$ the Rayleigh and aerosol backscattering anisotropy become evident for $\psi < 25^\circ$. As noted by Beyer et al. (1996), a possible correction will have the shape of the phase function of Rayleigh scatter, i.e., $(1+\cos^2\psi)$. However, as we have no guaranty that our $CC_{\text{min}}$ record really encompasses dry, clean atmospheres, and because other potential directional effects may intermingle, we prefer to fit the required correction to the lower envelope of the data set. For $\psi < 50^\circ$, the result is:

$$g''_{\text{min}} = 1 + 0.7[(50^\circ - \psi)/50^\circ]^2$$

(7)

Note that this effect is physically independent from the airmass effect. In addition, given the position of the Northeastern US pixels relative to the satellite position, the signature of this effect could also be observed independently (i.e., at different times of day). Hence, in the model the two corrections can simply be multiplied as described below.

3.6 Operational Logistics

As evidenced above, the crucial step for continuous operation is to establish and monitor the full dynamic range of the cloud index $n$. While care
for $CC_{\text{max}}$ stability is ensured by periodic inspections, $CC_{\text{min}}$ is calculated for every pixel of every satellite scene as $CC_{\text{min}} = CC/g_{\text{min}}$ with $g_{\text{min}} = g'_{\text{min}}g''_{\text{min}}$. The necessary adaptation to changing vegetation and soil properties is ensured by building the average of the ten lowest values from an evolving list (trailing window) of 5 days in winter and 18 days otherwise. The size of the trailing windows is not critical and these choices were optimized for our entire data set. The shorter winter window reflects the need to frequently reset the model’s minimum count because each snowfall noticeably increases the ground albedo. Scenes with snow covered ground have not received special attention here, although deliberate overriding of the trailing window upon each snow event occurrence is possible for more elaborate operations.

4. Performance Assessment of Effective RMSE

4.1 Conventional RMSE

The model is validated against the 12 measuring sites in the NE US described in Section 2. The 12 corresponding (closest) pixels are pre-determined assuming an ideal image orientation. No attempt is made a posteriori for correcting small misalignments of the actual images. Gauged in this way against 43,000 ground-truth station hours distributed from May to December 1995 for the whole network and, additionally, from January to December 1996 for the Albany site, the average satellite RMS error for global irradiance is 81 Wm$^{-2}$, or 23% of the average global irradiance, while the average mean bias error is 4 Wm$^{-2}$, or 1.2% of the measured average. These figures are very similar to those reported in various review papers (Section 1). The question is, how much of this 23% is due to the satellite-to-irradiance conversion model itself and how much is due to the shortcomings of the comparison method as well as to the other sources enumerated in the introduction?

4.2 Conventional RMSE with Optimum Pixel Approach

Möser (1983) and Möser and Raschke (1984) have investigated the number of neighboring pixels required for averaging around the closest pixel in order to minimize the validation RMSE. This is a probabilistic approach (Schmetz, 1989) for making the satellite’s estimates (discrete in
time, extended in space) comparable with ground measurements (discrete in space, extended in time). Rather than repeating this procedure, we follow an alternative path, more along the line of our approach.

We consider a block of 9 pixels centred at the nearest pixel, and determine which of the nine pixels exhibits lowest RMSE. For the majority of the sites and cases this happens to be the West to Northwest pixel. The wind blows predominantly from this direction, and the satellite scan is completed in the first quarter of the pyranometer integration hour. So it is plausible that, at the instantaneous scale, this West to Northwest pixel is frequently more representative of the conditions that will prevail during the hour at the closest (central) pixel than this closest pixel itself. Note that Mose (1982, 1983) already conjectured that use of elliptical averaging domains oriented in the predominant wind direction would improve validation results.

Switching from the closest to these “optimal” pixels (which, otherwise, remain fixed as before) brings the overall RMSE down to 19%, corresponding to 70 Wm\(^{-2}\) (see Table 2 for a four sites sample). Thus, the RMSE attributable to the heterogeneity of the data amounts to about 13%.

### 4.3 Effective RMSE

Irradiation fields are not as homogeneous in space as intuitively perceived. This is illustrated in Fig. 6, which extends the representation of RMS differences as a function of distance (Fig. 1) to various integration times, namely 10-minutes (only for the Swiss network), hourly, daily and monthly sums. The strong variability of the 10-minute and hourly sums, remarkably consistent in the two climatically distinct networks, over even short distances, explains why an estimate can be registered as inaccurate, even though it may be correct for several points in the pixel. The RMS difference at sub-pixel distances may be considered as the intrinsic lower limit of validation accuracy. This difference, termed “nugget effect” in geostatistics, has its origin in both measurement errors by the surface instruments and the micro-scale spatial resolution limitations of the satellite sensors. For hourly sums (Fig. 6) the nugget effect is of the order of 14–15%. Considering the 19% RMSE above leads us to an effective, or intrinsic RMSE, that is, the one traceable to the satellite-to-irradiance conversion model, of 12–13%. Two other possible indirect measures of this intrinsic error may be obtained from (1) a self similarity/ideal pixel analysis and (2) a locally homogeneous conditions analysis.

- **Self similarity/ideal pixel** approach: We return to the 9 pixel blocks, but centre them now on the

### Table 2. % RMSEs at 4 selected Sites and Different Pixel Combinations

<table>
<thead>
<tr>
<th>Site</th>
<th>closest pixel</th>
<th>optimal pixel</th>
<th>homogenous</th>
<th>ideal pixel/best-of-9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albany</td>
<td>23.27</td>
<td>20.65</td>
<td>17.77</td>
<td>11.25</td>
</tr>
<tr>
<td>New Paltz</td>
<td>21.59</td>
<td>18.88</td>
<td>12.55</td>
<td>11.25</td>
</tr>
<tr>
<td>Waltham</td>
<td>21.03</td>
<td>19.12</td>
<td>13.89</td>
<td>11.95</td>
</tr>
<tr>
<td>Mc Arthur</td>
<td>20.22</td>
<td>17.14</td>
<td>13.57</td>
<td>09.92</td>
</tr>
</tbody>
</table>

Fig. 6. Relative RMS difference between network sites as a function of distance for various integration times. As in Fig. 1, the points for the Swiss network are binned because of the very large number of possible station pairs. Points for the Northeast US network reflect each pair of stations. The consistency between the two climatically-distinct networks is conspicuous.
optimal pixel. Then considering that the fractal (hence self similar) structure of cloud patterns allows us to draw conclusions at the micro-scale (sub-pixel) from observations at the meso-scale (9 pixel block), we select, within each block, the estimated value closest to the observation. The resulting RMSE then drops, for almost all locations (Table 2), to 11%. Somewhere near the considered pixel (hence within the considered pixel when accounting for fractal effects, remaining navigation imprecision and time/space differences) the RMSE is roughly equal to the conventional RMSE minus the nugget effect.

- **Locally homogeneous conditions** approach: Homogeneous conditions are selected when the values within the 9 pixel block fluctuate within a 10% range, that is when no strong cloudiness gradient or scattered cloud conditions are present in the immediate vicinity of the ground station. The effect of cloud field structures and navigation imprecision should be minimized during these conditions. We restrict our data samples to those 4 sites (Albany, New Paltz, Dutchess and Waltham) where the ground albedo is expected to be fairly homogeneous across the whole block. This four station sample exhibits about the same closest pixel RMSE as the entire data set (22%). This RMSE drops down to 14% for locally homogeneous conditions (Table 2).

Hence both indirect approaches confirm our estimate of hourly intrinsic accuracy (of the order of 12–13%) with a nugget effect of similar magnitude.

5. Conclusions

The relative RMSE of hourly irradiance estimates with a simple satellite-based model amounts to 23% after verification against 8 site-years of data in the Northeastern US. This value compares favourably with the value of 20–25% reported world-wide in current literature. Validation results for several conditions allows us to determine the contribution of the various components of this RMSE. A substantial portion of this error is attributed to measurement errors by the surface instruments and, more importantly, to the genuine mico-variability of the irradiation field. The satellite pixel-to-irradiance conversion error is only 12–13%, which is considerably less than conventionally assumed. It is remarkable that this conclusion is reached with a very elementary satellite-to-irradiance conversion model, requiring only pixel brightness as an input. The use of more sophisticated algorithms (e.g., Charlock and Alberta, 1996; Chou and Zhao, 1997) could only strengthen this conclusion through a further reduction of this intrinsic error.

The practical meaning of this effective RMSE should not be underestimated: the precision of satellite-derived irradiances is likely to be of the order of 12%, while the RMS difference between irradiances 5 km apart already amounts of 15%. So if the modelled value differs, e.g., by 20% from the observation at one verification site, it is very likely to be registered as correct only a few kilometres away. With this in mind, we strongly recommend the use of satellite-based irradiance estimates, even near a measuring site, when site/time specific information is needed. Assessment of the effective RMSE has also implications in the process of merging satellite-derived estimates with ground data (e.g., for mapping purposes, D’Agostino and Zelenka, 1992), where an important issue is how much the combined values are allowed to differ from observations at the network sites.

Acknowledgements

This work was supported by contract XAH-515-22-201 from the National Renewable Energy Laboratory, operated for the U.S. Department of Energy by the Midwest Research Institute. Discussions with Thomas Konzelmann and his valuable comments were greatly appreciated.

References


Internet Data Distribution System, 1995: UCAR, Boulder, CO.


Kasten, F., 1990: Personal communication.


Müser, W., 1982: Personal communication.


Authors’ addresses: Dr. A. Zelenka (zelenka@zma.ch), Swiss Meteorological Institute, 8044 Zurich, Switzerland; Dr. R. Perez (perez@ascr.cestm.albany.edu) and R. Seals, The University at Albany, ASRC-CESTM, 251 Fuller Rd., Albany, NY 12203, U.S.A.; Dr. D. Renne (david_renee@nrel.gov), National Renewable Energy Laboratory, 1617 Cole Bd. Golden, CO 80401-3393, U.S.A.