Interpolation of daily solar radiation for New Zealand using a satellite data-derived cloud cover surface

Andrew Tait\(^1\) and Ben Liley\(^2\)


Abstract

Daily values of global solar radiation are important inputs for pasture growth models which are often run for rural locations (e.g. where detailed pasture growth data have been collected). However, it is possible that the nearest climate station with daily solar radiation data is located tens of kilometres away from these pasture measurement sites. This paper describes a method of interpolating daily global solar radiation data recorded or estimated at climate stations throughout New Zealand onto a regular 0.05° latitude / longitude (approximately 4 km) grid using a thin plate smoothing spline model. An analysis of the mean annual solar radiation prediction standard error from several spline model runs shows that the lowest prediction error is obtained using two positional variables and a satellite data-derived cloud cover surface. For the daily interpolations, maximum use is made of observational data by combining daily solar radiation and sunshine measurements collected over the period 1972 – 2006. An analysis of the daily solar radiation interpolation error at 20 validation sites shows that the average root mean square error varies between about 4.1 MJ/m\(^2\)/day in the summer months to about 1.5 MJ/m\(^2\)/day in winter (average relative error of between 20 and 30\%). Time series data derived from the daily gridded solar radiation data are currently being used in a pasture growth model for New Zealand.

Keywords
Solar radiation; interpolation; pasture growth model; New Zealand
1. Introduction

Daily measurements of global solar radiation are currently made at around 100 climate stations in New Zealand which are carefully sited to be representative of the general terrain of the surrounding area. Even so, there are many rural locations where the nearest climate station with solar radiation measurements is up to 50 km away. For the purpose of modelling pasture growth at these sites, using solar radiation data from the nearest site may be a significant source of model error. Therefore, there is a need to estimate daily global solar radiation accurately for such locations which could be anywhere in the country.

Numerous methods have been developed to estimate daily solar radiation from other meteorological variables such as air temperature, humidity, and precipitation (e.g., Bristow and Campbell, 1984; Hunt et al., 1998; Goodin et al., 1999; Thornton and Running, 1999; Mahmood and Hubbard, 2002). However these methods still only yield estimates at locations where the required meteorological variables are measured and result in relative root mean square errors (RMSEs) of between around 20 to 60% (depending upon season and model). This paper describes a method of interpolating daily global solar radiation data recorded or estimated at climate stations throughout New Zealand onto a regular 0.05° latitude / longitude (approximately 4 km) grid. Thus, the estimates are now not restricted to locations with measured data. Further, as will be demonstrated, the relative RMSEs from the interpolations are comparable to the best of the prediction models using other meteorological data inputs.

There are several spatial interpolation models which have been used to interpolate climate data from surface observations sites spanning large areas and encompassing complex terrain. The most common models are inverse distance weighting, Gaussian weighting, trend surface analysis (including linear and polynomial regression), kriging (including cokriging), and splines (Borga and Vizzaccaro, 1997; Daly et al., 2002; Hutchinson and Bischof, 1983; Hutchinson and Gessler, 1994; Laslett et al., 1987, Matheron, 1981; Phillips et al., 1992; Saveliev et al., 1998; Seaman and Hutchinson, 1985; and
Thornton et al., 1997). In this study, we have chosen a thin plate smoothing spline model, which has been shown to perform well for the interpolation of mean monthly solar radiation data (Hutchinson et al., 1984) and New Zealand climate data in general (Zheng and Basher, 1995; Sansom and Thompson, 2003; Tait and Turner, 2005; Tait et al., 2006; and Tait and Woods, 2007).

A thin plate smoothing spline model works by fitting a surface to the data with some error allowed at each data point, so the surface can be smoother than if the data were fitted exactly. Each station is omitted in turn from the estimation of the fitted surface and the mean error is found. This is repeated for a range of values of a smoothing parameter then the value that minimises the mean error is taken to give the optimum smoothing. This process is called minimizing the generalised cross validation (GCV). It can be automated once the order of the derivative, which controls the surface roughness, has been chosen. The software used was ANUSPLIN version 4.2 (Hutchinson, 2008).

This paper is structured in the following way: section two presents the results from an analysis of the mean annual solar radiation prediction error from seven spline model runs using different interpolation surfaces; section three reviews the daily solar radiation and sunshine data resource for New Zealand and describes how sunshine data is used to supplement solar radiation data for use in the daily interpolations; section four compares two maps of mean daily solar radiation – one derived from interpolated mean solar radiation values and the other from averaging all the daily solar radiation interpolations; section five shows the results of an analysis of the daily solar radiation interpolation error at 20 validation sites; discussion and conclusions follow in section six.

2. Analysis of the spline model prediction standard error

Values of most meteorological variables, such as air temperature and rainfall, vary in a consistent manner with changes in orography. Thus it makes sense

---

1 The spline computation runs quickly. It takes approximately 3 seconds to interpolate one day’s solar radiation data from about 100 climate stations onto a 0.05° latitude / longitude grid covering all of New Zealand (11,491 grid points) on a UNIX machine with 2 Gigabytes of RAM and a 667 MHz processor.
to interpolate these data using a spline model with two position variables and an orographic variable such as elevation. The broad spatial pattern is determined by the two position variables, while the inclusion of elevation modifies the broad pattern to give more precise representations of the higher resolution variability. Hutchinson (1991) used a trivariate thin plate smoothing spline (latitude, longitude and elevation) to interpolate several meteorological variables across Australia. A trivariate spline, which allows the relationship between the climate variable and three dependent variables (e.g. latitude, longitude and elevation) to vary spatially, was deemed more appropriate for continent-wide interpolations. This is compared with a trivariate partial spline (two positional variables plus a single linear dependence upon elevation), which is more suited to small-scale applications such as mapping the spatial variation of temperature within a single valley.

For the interpolation of mean monthly solar radiation data across Australia, Hutchinson et al. (1984) showed that a trivariate interpolation using a third variable indicative of cloudiness (transformed monthly rainfall surfaces were used in this case) performed better than a simple bivariate interpolation (using position variables only). The mountainous terrain of New Zealand (Figure 1) combined with its maritime location results in a complicated spatial distribution of cloudiness and hence complicated insolation patterns (de Lisle, 1966). To best capture this spatial variability, six trivariate and one bivariate spline model runs were trialled using mean annual solar radiation data from about 100 sites. The surfaces included in the trivariate models (along with two positional variables) were elevation, rainfall (transformed and un-transformed), raindays (a rainday is counted when the daily rainfall total is greater than or equal to 1 mm), relative humidity, and satellite data derived cloud cover. The bivariate model only included the two positional variables.

The rainfall, raindays and relative humidity surfaces were themselves derived from thin-plate smoothing spline interpolations of mean annual values calculated at 1950, 1621 and 515 sites throughout New Zealand, respectively. For rainfall and raindays, an expert-guided mean annual rainfall surface was used as the third variable in the trivariate spline model (Tait et al., 2006) while
Figure 1: A map of New Zealand showing areas above 500 and 1000m, and the location of the 20 validation sites.

Elevation was used for the interpolation of relative humidity. The cloud cover surface was derived from Advanced Very High Resolution Radiometer (AVHRR) PM data collected over the New Zealand region for the period 1996...
using a Bayesian cloud classification algorithm called “SRTex” (Uddstrom and Gray, 1996). The surface represents the average frequency (as a percentage) of cloud-free skies at the times of satellite overpasses.

The ANUSPLIN program used for this analysis produces as part of the suite of possible outputs an estimate of the model prediction standard error. This error statistic is a combination of the variance of the data values and the model standard error estimate of the interpolated values. Confidence intervals of the calculated spline values are estimated by multiplying the prediction standard error by 1.96, the 95 percent two-sided confidence interval of the standard normal distribution. The minimum, mean, and maximum model prediction standard errors for each of the surfaces are shown in Table 1. It can be seen that the satellite data-derived cloud cover surface yields the lowest model prediction error amongst all the candidates analysed here.

Table 1: Spline model prediction standard errors for the interpolation of mean annual solar radiation for New Zealand (X and Y are easting and northing positional variables in metres, New Zealand Map Grid projection). Minimum errors are shown in bold and maximum errors are italicized.

<table>
<thead>
<tr>
<th>Spline model</th>
<th>Variables</th>
<th>Minimum Error (MJ/m²/day)</th>
<th>Mean Error (MJ/m²/day)</th>
<th>Maximum Error (MJ/m²/day)</th>
<th>Mean Relative Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trivariate</td>
<td>X, Y,</td>
<td>0.38</td>
<td>0.46</td>
<td>0.98</td>
<td>3.38%</td>
</tr>
<tr>
<td></td>
<td>Elevation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trivariate</td>
<td>X, Y,</td>
<td>0.39</td>
<td>0.46</td>
<td>1.20</td>
<td>3.54%</td>
</tr>
<tr>
<td></td>
<td>Rainfall</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trivariate</td>
<td>X, Y,</td>
<td>0.39</td>
<td>0.44</td>
<td>0.64</td>
<td>3.31%</td>
</tr>
<tr>
<td></td>
<td>Rainfall*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trivariate</td>
<td>X, Y,</td>
<td>0.61</td>
<td>0.72</td>
<td>1.10</td>
<td>5.50%</td>
</tr>
<tr>
<td></td>
<td>Raindays</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trivariate</td>
<td>X, Y,</td>
<td>0.39</td>
<td>0.44</td>
<td>0.67</td>
<td>3.28%</td>
</tr>
<tr>
<td></td>
<td>Relative</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Humidity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trivariate</td>
<td>X, Y,</td>
<td>0.38</td>
<td>0.42</td>
<td>0.61</td>
<td>3.13%</td>
</tr>
<tr>
<td></td>
<td>Cloud</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cover</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bivariate</td>
<td>X, Y</td>
<td>0.77</td>
<td>1.00</td>
<td>2.20</td>
<td>7.46%</td>
</tr>
</tbody>
</table>

* Rainfall transformed prior to interpolation using tanh(rain/1300) [after Hutchinson et al., 1984], where 1300 is the mean annual rainfall (mm) at the input data sites.

In the following sections, the average frequency of cloud-free skies surface is used for the interpolation of daily solar radiation data. A criticism of this
approach is that the average cloud cover pattern does not capture the day-to-day variability in cloudiness. A much improved method would be to use actual daily cloud cover surfaces for the interpolation of the corresponding measured surface solar radiation data. Regrettably, this was not possible from the SRTex data for two reasons. First, for any one location, the AVHRR instrument on the NOAA 14 satellite made at most two usable measurements per day, 100 minutes apart, as it swept the Earth in sun-synchronous orbit. Thus the satellite data lack the time-integration of the daily climate station-based observations of solar radiation and sunshine hours. Secondly, many satellite observations are unusable, for various reasons, giving an even sparser data set. Thus, all the SRTex data were combined to determine averages over longer times. In the present analysis we have used them over the entire period, but in future work we will look at alternatives, including disaggregating the cloud cover data by synoptic weather type (e.g., using the 12 generalised synoptic types of Kidson, 2000), which should improve the interpolation accuracy.

Satellite data with much greater temporal resolution, such as GMS (Geostationary Meteorological Satellite) hourly data, would avoid the above limitation, but bring other problems. The viewing angle for GMS is poor for New Zealand, and the resolution is much lower. The SRTex algorithm gives not just the presence or absence of cloud, as used here, but classification of cloud type and optical depth. These parameters, in some combination, may also improve the interpolation should the methodology be revised in the future.

3. Daily solar radiation and sunshine hours data
The majority of New Zealand’s climate data are stored in the National Climate Database (CLIDB), an Oracle relational database. This is maintained by the National Institute of Water and Atmospheric Research Ltd. (NIWA) in Wellington, New Zealand. There is an extensive network of currently-open climate stations throughout New Zealand. For example, rainfall is currently measured at around 680 sites; screen observations (i.e. maximum and minimum air temperature, and dry- and wet-bulb temperature) at around 240
sites; wind speed or wind run at around 140 sites; and solar radiation or sunshine hours at around 113 sites. These numbers have varied significantly since the early 1970’s (for example, hourly global solar radiation was only measured at 4 locations in 1972 compared with 98 sites in 2006, Figure 2).

The solar radiation sensors, mostly LiCor LI200 pyranometers, are periodically calibrated at a NIWA testing facility. Most sensors have been calibrated between two and four times, and between these times they have been redeployed to different climate stations. The calibration history shows an average coefficient of variation of 3.5%. In the data analysed for this study, time averaging will reduce the effect of any random measurement error, but calibration error will bias the results for the duration of an instrument’s deployment. Thus the 3.5% provides a minimum uncertainty for the site values to be fitted by spatial interpolation.

For the purposes of input into and calibration of a pasture growth model, it was necessary to estimate daily solar radiation for a number of decades into the past. This was not possible using measurements of global solar radiation alone, due to the paucity of data prior to the late 1990s. However, there have been measurements of daily sunshine hours at many locations throughout New Zealand back to 1972, although the number of sites is now declining (Figure 2). Thus it was decided to augment the daily global solar radiation measurements with estimates of solar radiation made from daily sunshine hours data. The estimation procedure is by way of the Ångström equation (Ångström, 1924):

\[
\frac{I}{I_{\text{max}}} = a + b \cdot \frac{S}{S_{\text{max}}}
\]

Equation 1

where \( I \) is the measured global solar radiation (or irradiance); \( I_{\text{max}} \) is the maximum possible global surface solar radiation at the given latitude and day of year; \( S \) is the measured bright sunshine hours; \( S_{\text{max}} \) is the maximum possible surface sunshine hours at the given latitude and day of year; and \( a \) and \( b \) are constants.
The maximum global surface irradiance $I_{\text{max}}$ is taken as the clear sky value. Although instantaneous surface irradiance under scattered bright cloud with unobscured sun can exceed the clear sky value by tens of percent, this effect is more than compensated in hourly integrals by the large reduction in irradiance for the times that the sun is obscured. 

There are 22 sites in New Zealand which have recorded at least five years of coincident daily solar radiation and sunshine hours data over the period 1971–2000. These data, divided by their respective calculated maximum possible surface values, were plotted as a scatterplot (Figure 3) and a linear regression model was applied to calculate the intercept and slope of the line of best fit (coefficients $a$ and $b$ respectively in Equation 1). The regression coefficients from this analysis were 0.3822 (coefficient $a$) and 0.5485 (coefficient $b$). The $r$-squared value from the regression was 0.622 with a standard error of 0.04, indicating there is a significant relationship between the two parameters (F-test for significance gives $p < 0.01$). The scatter about the best fit line shown on Figure 3 is predominantly caused by regional-scale variability.
Figure 3: The relationship between the average daily global solar radiation divided by the average maximum daily global solar radiation (y-axis) versus the monthly sunshine hours total divided by the maximum monthly sunshine hours total (x-axis) at the 22 climate stations around New Zealand with at least 5 years of overlapping radiation and sunshine data during the period 1971–2000. Also shown are the regression relationship and $r^2$-squared value.

Table 2 shows the results of regression analyses performed on each of the 22 sites with coincident daily solar radiation and sunshine hours data. It can be seen that the $r^2$-squared value for the individual analyses is generally higher than the combined value and that the slope and intercept of the best fit line varies slightly between sites. It was therefore decided to interpolate the 22 slope and intercept values using a simple inverse-distance weighting bilinear interpolation to provide estimates of these values anywhere in the country, rather than use the coefficients from the single multi-site regression analysis. These location-specific estimates of the Ångström equation coefficients were then used to estimate daily global solar radiation at sites with measured daily sunshine hours data using a rearrangement of Equation 1 with the measured global solar radiation as the predictand.

As a final step before the interpolations were initiated, the input data were given a weighting (or a relative variance, as it is referred to in the ANUSPLIN
Table 2: Regression slope, intercept, and r-squared value for average daily global solar radiation divided by the average maximum daily global solar radiation versus the monthly sunshine hours total divided by the maximum monthly sunshine hours total at each of the 22 climate stations around New Zealand with at least 5 years of overlapping radiation and sunshine data during the period 1971–2000. The highest r-squared value is bolded and the lowest is italicized.

<table>
<thead>
<tr>
<th>Station ID</th>
<th>Latitude (°S)</th>
<th>Longitude (°E)</th>
<th>Elevation (m)</th>
<th>Slope</th>
<th>Intercept</th>
<th>R-squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>1024</td>
<td>35.1</td>
<td>173.3</td>
<td>80</td>
<td>0.48</td>
<td>0.39</td>
<td>0.69</td>
</tr>
<tr>
<td>1041</td>
<td>35.1</td>
<td>173.3</td>
<td>85</td>
<td>0.53</td>
<td>0.36</td>
<td>0.68</td>
</tr>
<tr>
<td>1340</td>
<td>36.3</td>
<td>174.8</td>
<td>27</td>
<td>0.61</td>
<td>0.39</td>
<td>0.46</td>
</tr>
<tr>
<td>1768</td>
<td>38.1</td>
<td>176.3</td>
<td>281</td>
<td>0.45</td>
<td>0.43</td>
<td>0.31</td>
</tr>
<tr>
<td>1962</td>
<td>37.0</td>
<td>174.8</td>
<td>33</td>
<td>0.60</td>
<td>0.38</td>
<td>0.65</td>
</tr>
<tr>
<td>2101</td>
<td>37.8</td>
<td>175.3</td>
<td>40</td>
<td>0.41</td>
<td>0.42</td>
<td>0.55</td>
</tr>
<tr>
<td>2109</td>
<td>37.8</td>
<td>175.3</td>
<td>66</td>
<td>0.42</td>
<td>0.44</td>
<td>0.70</td>
</tr>
<tr>
<td>2282</td>
<td>39.0</td>
<td>174.2</td>
<td>27</td>
<td>0.57</td>
<td>0.37</td>
<td>0.83</td>
</tr>
<tr>
<td>2807</td>
<td>38.7</td>
<td>178.0</td>
<td>4</td>
<td>0.56</td>
<td>0.36</td>
<td>0.81</td>
</tr>
<tr>
<td>3145</td>
<td>40.9</td>
<td>175.0</td>
<td>5</td>
<td>0.64</td>
<td>0.33</td>
<td>0.90</td>
</tr>
<tr>
<td>3206</td>
<td>40.2</td>
<td>175.4</td>
<td>48</td>
<td>0.58</td>
<td>0.39</td>
<td>0.80</td>
</tr>
<tr>
<td>3277</td>
<td>40.7</td>
<td>175.3</td>
<td>46</td>
<td>0.63</td>
<td>0.38</td>
<td>0.62</td>
</tr>
<tr>
<td>3385</td>
<td>41.3</td>
<td>174.8</td>
<td>125</td>
<td>0.69</td>
<td>0.31</td>
<td>0.81</td>
</tr>
<tr>
<td>3909</td>
<td>42.7</td>
<td>171.0</td>
<td>39</td>
<td>0.65</td>
<td>0.32</td>
<td>0.73</td>
</tr>
<tr>
<td>4241</td>
<td>41.3</td>
<td>173.2</td>
<td>2</td>
<td>0.54</td>
<td>0.38</td>
<td>0.62</td>
</tr>
<tr>
<td>4310</td>
<td>41.5</td>
<td>174.0</td>
<td>4</td>
<td>0.65</td>
<td>0.31</td>
<td>0.78</td>
</tr>
<tr>
<td>4507</td>
<td>42.4</td>
<td>173.7</td>
<td>108</td>
<td>0.63</td>
<td>0.36</td>
<td>0.66</td>
</tr>
<tr>
<td>4630</td>
<td>44.0</td>
<td>170.5</td>
<td>1027</td>
<td>0.51</td>
<td>0.43</td>
<td>0.55</td>
</tr>
<tr>
<td>4843</td>
<td>43.5</td>
<td>172.5</td>
<td>37</td>
<td>0.66</td>
<td>0.32</td>
<td>0.70</td>
</tr>
<tr>
<td>5397</td>
<td>45.9</td>
<td>170.2</td>
<td>1</td>
<td>0.61</td>
<td>0.37</td>
<td>0.62</td>
</tr>
<tr>
<td>5576</td>
<td>45.3</td>
<td>169.4</td>
<td>141</td>
<td>0.53</td>
<td>0.39</td>
<td>0.52</td>
</tr>
<tr>
<td>5814</td>
<td>46.4</td>
<td>168.3</td>
<td>0</td>
<td>0.59</td>
<td>0.39</td>
<td>0.58</td>
</tr>
</tbody>
</table>

The relative variance is used to place an emphasis on some stations over others in the interpolation (the smaller the relative variance the greater the weight). Often the inverse of the length of the data record is used when interpolating long-term mean surfaces, which gives more weight to stations with longer records. For the daily solar radiation interpolations, each station has the same period (one day), but some stations have estimates (from sunshine hours data) while others have actual measurements. Thus, it was decided to assign the stations with measured solar radiation a relative variance of zero, while the stations with estimated solar radiation were weighted with the variance (i.e. the square of the standard deviation of the residuals) from the regression analysis shown on Figure 3 (equal to 0.002).
4. Comparison of mean daily solar radiation surfaces

Figure 4a shows a map of the long-term mean daily solar radiation, based on the period 1972–2006. The map is produced from interpolated mean daily solar radiation values recorded at 118 climate stations plus estimated (from mean daily sunshine hours values) at 95 climate stations. The stations have varying lengths of record over the 35-year period. The interpolation model is a second order derivative trivariate (latitude, longitude, and percent cloud-free skies) thin plate spline model. Stations were included in the input data set if they had a minimum of three years of solar radiation or sunshine hours data during the period (see Figure 4a for the locations of these 213 stations), and the inverse of number of years of data at each station was used as a relative variance for the interpolation.

Figure 4b shows the percent difference from Figure 4a of the mean daily solar radiation derived from averaging all the daily solar radiation interpolations over the same period, 1972–2006 (rather than from a single interpolation of the mean values, as is mapped in Figure 4a). It can be seen that for most of New Zealand the difference is plus or minus 5 percent, with a tendency towards higher positive differences of up to 15 percent (indicating the mean of the daily interpolations is greater than the interpolated mean values) in areas of complex terrain.

These differences are caused by a combination of two main factors. First, there are fewer input climate stations available on a particular day for the daily interpolations (approximately 100 on any day) compared with the mean daily interpolation (213). Second, the interpolation of mean climate data is generally more robust (particularly when using a mean cloudiness surface in the spline model) and hence more accurate than the interpolation of daily values, due to the greater spatial variability exhibited in daily values associated with the passage of short-duration weather events.

Despite these sources of difference between the two surfaces, it is argued that the percent difference shown in Figure 4b is not great and indicates that the interpolation of daily solar radiation data produces reasonable values for
Figure 4: a) A map of the mean daily global solar radiation for the period 1972–2006, interpolated from the 213 (118 radiation plus 95 sunshine hours) climate stations shown on the map; and b) A map of the percent difference from Figure 4a of the average of the daily solar radiation interpolations for the same period.
the majority of the grid points covering New Zealand, particularly those in the lower elevations. High elevation areas show a larger difference, however due to the paucity of climate stations at high elevations in New Zealand it is likely that neither of the interpolations (daily or the long-term mean) are particularly accurate in these areas, so the comparison is difficult to assess.

5. Error estimation

The accuracy of the daily solar radiation interpolations was further examined by selecting 20 validation sites where interpolated daily values were compared with actual daily values, and the average root mean square error (RMSE) for every month was calculated. These RMSE values were then averaged over the 20 sites. The validation sites all have at least 10 years of solar radiation data during the period 1972–2006 (the average record length for the 20 sites is 21.7 years). Figure 1 shows the location of the validation sites, which are well spread around the country. For the validation run, the data from these 20 sites were not used in the daily interpolations.

Figure 5 shows the average RMSE for each month, as well as the average measured daily solar radiation for each month at the validation sites. The line on Figure 5 shows the relative error for each month, calculated as the average RMSE divided by the average daily solar radiation, and expressed as a percentage. It can be seen that the average RMSE varies between around 4 MJ/m²/day in the summer months to around 1.5 MJ/m²/day in winter. The average daily solar radiation also varies seasonally, from a maximum in summer of around 21 MJ/m²/day to a minimum in winter of around 5 MJ/m²/day. Thus, the relative error varies from 18 – 20% in summer to 27 – 30% in winter.

Two important results should be emphasised here. Firstly, the relative error throughout the primary growing season (October – March) is consistently
between 18 and 20%. This is the period when estimating water loss through evapotranspiration and biomass production through photosynthesis is most crucial. Secondly, the removal of the 20 validation sites from the daily interpolations has likely reduced the overall interpolation accuracy, as these sites all have long-term solar radiation data. The actual daily interpolations (i.e. not the validation run) include these data, thus the true error is likely to be less than that shown in Figure 5.

Mahmood and Hubbard (2002) report relative errors of between 20 and 36% from the best of the daily solar radiation models using other meteorological variables they evaluated (this was the Bristow and Campbell, 1984 model using maximum and minimum air temperature data). The interpolation method presented here is comparable with this method based on the relative error, but is superior in that the estimates are not restricted to locations with measured data.
6. Summary and conclusion
The interpolation of daily solar radiation onto a regular 0.05° latitude / longitude grid for all of New Zealand from 1972–2006 (which is currently being updated to the present) has yielded a valuable dataset which has many potential applications. The spline interpolation scheme uses an annual mean percentage cloudiness surface derived from satellite data which has been shown to yield the lowest model prediction standard error of a number of candidate surfaces.

A regression-based method is used to extend the number of stations at which daily solar radiation measurements are made to those sites at which daily sunshine hours is measured. The solar radiation estimates are then weighted lower than the actual values in the interpolation procedure, which places more emphasis on the actual values while significantly extending the number of input data sites.

A comparison of mean daily solar radiation surfaces; one derived from interpolated mean daily values and the other from averaging all the daily interpolations, showed that for the majority of low elevation areas throughout New Zealand the difference is between plus or minus 5 percent. High elevation areas showed a larger difference, which indicates that the estimates in these areas should be used with some caution, particularly due to the paucity of climate stations at high elevations in New Zealand.

A more detailed error analysis has shown that the average RMSE of the daily solar radiation interpolations for the validation sites varies between about 4 MJ/m²/day in summer and 1.5 MJ/m²/day in winter. The average relative error (the RMSE divided by the average daily solar radiation) is between 18 and 20% throughout the important growing season months (October – April), but increases to 27 – 30% over winter when the average daily solar radiation at the validation sites is only around 5 MJ/m²/day. The true error of the daily interpolations is likely to be less than these values however, as a significant amount of good quality data was removed from the validation run interpolations.
In conclusion, the interpolated daily solar radiation dataset is regarded as a new and valuable resource for New Zealand. Together with daily rainfall, temperature, vapour pressure, wind speed, and potential evapotranspiration estimates, the data are already being used operationally for pasture growth modelling. Further applications, such as household solar energy assessments, are currently being tested. Readers of this paper may contact the authors for more information regarding the acquisition of these data for their own purposes, if desired.

Acknowledgements

The authors would like to thank Michael Uddstrom, John McGregor, and Greg Bodeker (all of NIWA), Val Snow (AgResearch Ltd.), and Cameron Clark (DairyNZ Ltd.) for their assistance in this study. We would also like to thank New Zealand’s Foundation for Research Science and Technology for supporting this work under contract C01X0010.

References


Submitted to *Weather and Climate* February 2009, Revised May 2009