A method for generating synthetic hourly solar radiation data for any location in the south west of Western Australia, in a world wide web page

Dean Laslett[a], Chris Creagh[b], and Philip Jennings[c]

Murdoch University Western Australia

(a) School of Engineering and IT, Murdoch University, South Street, Murdoch, WA 6150, Australia
gaiaquark@gmail.com (Corresponding author)
(b) School of Engineering and IT, Murdoch University, South Street, Murdoch, WA 6150, Australia
c.creagh@murdoch.edu.au
(c) School of Engineering and IT, Murdoch University, South Street, Murdoch, WA 6150, Australia
p.jennings@murdoch.edu.au

Abstract

An algorithm was developed to generate synthetic hourly cloudiness data for any time of the year at any location in the south west region of Western Australia (WA). To enable the algorithm to be used for simulation of the power output of both tilted photovoltaic and concentrating solar power systems, a metric of cloudiness was defined which modifies the clear sky beam, diffuse and reflected solar transmittance. Seasonally and positionally adjusted values of daily cloudiness were generated by roughly mimicking the geographic pattern of annual rainfall in WA. Rather than longitude and latitude, distance along the coastline and distance inland from the coast were used as the positional coordinates. Hourly cloudiness data was generated from the daily values using a first order autoregression algorithm with time varying mean and standard deviation. Two years of measured hourly horizontal solar irradiance data from a network of 31 weather stations was used to calibrate the algorithm. The algorithm was simple enough to run inside a world wide web page and has the potential to be adapted to other regions with a similar pattern of declining inland rainfall.

Keywords: Hourly, Solar, Irradiance, Simulation, Western, Australia

1. Introduction

In order to prevent dangerous climate change, there is a need to switch to energy systems with low greenhouse gas emissions, often called renewable energy systems. Systems that utilise solar energy are now a significant way to generate low emission energy. It is insufficient to use an average daily value of solar radiation to simulate these systems because solar radiation reaching the ground can vary significantly within a single day. Energy demand can also vary widely over the course of a day. To investigate the balancing requirements of systems that use significant amounts of solar energy, there is a need to develop algorithms to simulate radiation on finer time-scales. For the purposes of further developing an interactive simulation of renewable energy systems in Western Australia, an algorithm to generate synthetic hourly solar radiation data over a range of

locations, with diurnal and seasonal variations that are a reasonable representation of actual local conditions, was needed and is the aim of this study.

As the simulation is to be used for educational purposes, easy and wide spread accessibility is desirable. Although there are still some incompatibilities, modern web browsers have the potential to provide a universal graphics platform that is common across different operating systems and present on most personal computers in the world. As such, they are a promising tool for making simulations accessible. Maximum accessibility could be achieved if the simulation was available as a stand-alone (capable of running offline) world wide web page, as then there is no need to download data from a remote server or even connect to the internet. The steady increase in code interpretation speed in modern web browsers has improved the plausibility of doing this, although the algorithm should still be computationally as simple as possible and have data storage requirements as small as possible.

Previously, solar radiation data for many locations have only been measured on a horizontal surface. Hence much theoretical work and synthetic data algorithm development has concentrated on horizontal surfaces (e.g. [1] and [2]). In this study also, all references to the word 'horizontal' refer to radiation falling on a horizontal surface rather than the component of radiation travelling in a horizontal direction. These horizontal surface algorithms are generally inadequate for the simulation of most solar power generating devices, because the collecting surfaces are usually not horizontal. Radiation incident on the earth's surface can be divided into three components, and the degree of tilt from the horizontal affects each component differently. The three components are the beam (also called direct) component that has come from the sun; the diffuse component, resulting from radiation that has been scattered in the atmosphere; and the reflected component, resulting from radiation reflected off other surfaces. The diffuse and reflected components are indirect, although there is a dependence on the beam radiation they originate from. Radiation falling on a horizontal surface includes both beam and diffuse components lumped together (in theory, the reflected component is zero for horizontal surfaces in isotropic surroundings). Solar power systems that use mirrors to focus and concentrate radiation utilise only the beam component. Photovoltaic (PV) surfaces utilise all three components.

Under clear skies, each component of radiation can be theoretically estimated from the position of the sun in the sky and variations in Earth's orbit. If clouds, haze, smoke, fumes, or atmospheric pollutants are present, then the beam component will be reduced in a spectrum dependent way, and since the diffuse and reflected components are dependent on the beam component, they will also be affected. The characterisation of these effects over varying time and spatial scales is one of the biggest challenges facing any synthetic

data generation algorithm.

Full scale physics-based meteorological or global circulation models are very computationally intensive and usually require a super computer. However, if measured irradiance data is available, then there is the possibility of using simpler empirical techniques to predict future irradiance. Graham *et al.* [3] posit that any data set, real or synthetic, that captures underlying climatological behaviour and has similar probability characteristics to a long term historical data set, if it existed, should be sufficient to simulate the performance of a solar power generation system. Gazela and Mathioulakis [4] detail methods for constructing a "typical meteorological year" (TMY) database from long term weather data. The database would typically contain values for several variables, including solar irradiation, over an entire year, and ideally, represent the gamut of typical weather patterns for a particular location. However Muneer *et al.* [5] point to a common problem with existing solar radiation data sets: the beam and indirect components are seldom measured independently. Graham *et al.* [3] point to another common problem. To accurately assess the performance of many systems that use solar radiation, radiation data on a time scale that captures transient changes is required. For example, hourly data is necessary to capture changes in the position of the sun throughout the day. Measurements on these fine time scales are not available in many places in the world. It may be possible to get reasonably accurate estimates for missing data by using alternate synoptic information such as pressure, temperature, and cloud cover. Models that can do this are called Meteorological Radiation Models (MRMs).

Generation of fine time scale variation using common statistical distributions has been investigated. Semenov *et al.* [6] found that using a normal distribution to generate synthetic solar radiation values did not match well with the distribution of measured solar radiation in general and at a site in the UK in particular. Punyawardena and Kulasiri [7] compared 17 years of measured daily bright sunshine duration at a meteorological station in Sri Lanka with different statistical distributions and concluded that the best statistical fit was the Weibull distribution. Boland [8] performed a Fourier spectrum analysis of solar radiation data for several sites in Australia (including two in Western Australia), and found that solar radiation followed seasonal and diurnal cycles.

Several algorithms attempt to generate finer time scale radiation data using data from longer time scales. Perhaps the simplest approach was taken by Celik [9], who generated clear sky radiation curves and then modified the amplitude such that the average daily radiation was similar to one of several typical daily values within a particular month. Hence atmospheric conditions such as cloudiness are taken into account in an overall fashion. This approach could be used to generate data on any time scale but does not capture any of the transient variation that may be present on that time scale. Gordon and Reddy [10] used data from widely varying climatic conditions to develop a simple

functional form for the probability density function of daily radiation but found that there is location dependence. More sophisticated autoregressive moving average (ARMA) and Fourier analysis techniques have been developed to synthesise radiation data [2], but these have also been found to be location dependent.

In contrast, Liu and Jordan [11] found much earlier that if clearness index was used instead of radiation as the variable to be analysed, then the probability features were quasi-universal. The clearness index (k_t) is defined as the ratio of global (i.e. total) radiation falling on a horizontal plane at the surface of the Earth (I_h) to the total extraterrestrial radiation falling on a plane horizontal to the surface but at the top of the atmosphere (I_{hex}):

$$k_t = \frac{I_h}{I_{hex}} \tag{1}$$

The symbol K_t is commonly used to represent daily clearness index, the ratio of total daily radiation (H) falling on a horizontal plane at the surface to the total daily extraterrestrial radiation (H_{ex}) falling on a plane horizontal to the surface but at the top of the atmosphere:

$$K_t = \frac{H}{H_{ex}} \tag{2}$$

Graham *et al.* [3] claim that most of the seasonal variation in daily radiation is due to variation in the extraterrestrial radiation, which can be accounted for by using the clearness index. This finding has been prominent in the development of a number of empirical algorithms to synthesise data when measurements are only available on a longer time scale. These algorithms operate over different time scales, such that it is conceivable to use them in cascade, ultimately synthesising data on an hourly or finer time scale, when measurements are only available on a monthly or yearly time scale. Mora-Lopez and Sidrach-De-Cardona [12] proposed a method to generate synthetic hourly values of the clearness index directly from monthly average values of the daily average clearness index. The method involved incorporating a seasonal component as well as a component related to the diurnal sun cycle. However, the authors found the algorithm was not universal.

Both Graham and Hollands [13] and Aguiar and Collares-Pereira [14] developed well known algorithms to synthesise hourly data from daily average clearness index values, with time of day dependent probability distributions. The Aguiar and Collares-Pereira [14] algorithm, called the TAG algorithm, claims better statistical consistency between the synthetic data and measured data for clearer months. Also, since a wider range of locations were used to develop the TAG algorithm, it may perhaps be more widely applicable. Remund *et al*. [15] provided modifications to the TAG algorithm for locations

where the sun might be occluded by the skyline, such as in mountainous regions.

There have been several models using artificial neural networks (ANNs) developed to generate synthetic hourly data from average daily solar radiation. Hontoria *et al.* [2] developed an ANN model based on data from several locations in Spain, and found that their model matched the measured data more closely than either the Graham and Hollands [13] or Aguiar and Collares-Pereira [14] models. However Reikard [16] reported mixed results, with ANN models significantly outperforming autoregressive models only at higher temporal resolutions, in the order of a few minutes. ANN models can also be much more numerically intensive. More recently, satellite remote sensing data has been used to provide estimates of surface radiation. The quantity of data required to comprehensively represent solar radiation behaviour over all seasons and cover an entire region would be large.

A number of shortcomings with the clearness index approach have become apparent. The beam and diffuse components of radiation falling on a horizontal surface are bundled together into measurements of the clearness index. Generation of synthetic radiation values for surfaces oriented at an angle other than horizontal will require individual quantification of these components and also the reflected component. Since both clearness index and the diffuse component are affected by atmospheric processes such as clouds, it was hoped that there might be a close relationship between the two [17]. However, it has been found that the diffuse component can vary widely for very similar values of clearness index. Skartveit *et al.* [18] found that clouds have a complex effect on the diffuse component. For example, in the case of broken or scattered cloudiness conditions, the diffuse component can be so high that the measured clearness index exceeds the theoretical clear sky value. Perez *et al.* [19] identified two limitations of using the clearness index: firstly, there is a dependency not only on atmospheric processes, but also on solar elevation. Secondly, the complexity of the diffuse component relationship means that in some situations it might be impossible to differentiate two different cloud conditions with the same clearness index. The authors proposed the use of a normalised clearness index that is solar elevation independent as a possible solution to the first limitation.

Approaches that use cloudiness metrics rather than clearness index have also been taken. Yang and Koike [20] developed a sky clearness indicator (SCI) coefficient that accounts specifically for cloud effects, modifying the surface global horizontal radiation from its clear sky value. Butt *et al.* [21] developed a method of estimating cloudiness using high frequency (2 minute interval) measurements of global horizontal surface irradiance. Although this method of defining cloudiness was somewhat imprecise, the authors found a straight-line relationship between the ratio of diffuse horizontal irradiance to global horizontal irradiance (the diffuse fraction) and their cloudiness metric for two locations in the Amazon.

Clearness index usually has lower values earlier and later in the day because of the increased air mass sunlight travels through when the sun is low in the sky. In contrast, cloudiness metrics are not expected to exhibit such a strong trend because they modify the clear sky transmissivities rather than directly modifying the solar radiation. As such, they perhaps can be claimed to be a "purer" measure of the local cloud condition at the time. There is a hierarchy here. Direct measurements of radiation contain effects due to variations in Earth's orbit, the position of the sun in the sky, air mass, and cloud condition. Measurements of clearness index contain effects due to the latter two, and cloudiness metrics aspire to capture only the cloud condition.

In this study, the aim was to develop a numerically simple algorithm that can generate synthetic values of radiation data for any location within the South West quarter of Western Australia, with hourly, diurnal and seasonal variations in solar radiation that give a reasonable representation of actual local conditions. Australia lies in the Southern hemisphere, so simply using models developed for Northern hemisphere locations may not achieve this aim. Lanini [22] points out that stratospheric sulphate aerosol content is higher in the Northern hemisphere, increasing the diffuse fraction. Hence a new model was developed that is a combination of the approaches of Yang and Koike [20] and Aguiar and Collares-Pereira [14]. Firstly, equations to estimate the theoretical clear sky values for beam, diffuse, and reflected radiation were obtained from previous studies. Then, to represent the effects of cloud, a coefficient of cloudiness was defined that modifies all three components of clear sky radiation individually, and so can be used to model the performance of both concentrating solar power systems and tilted PV systems on cloudy days. An autoregressive algorithm was developed to generate synthetic hourly values of this cloudiness coefficient, with parameters that are themselves seasonally and locationally dependent and able to be calibrated with locally measured data from Western Australia.

It was considered that using latitude and longitude as the location coordinates would not lead to the numerically simplest operation of the model. Latitude provides information about the position of the sun in the sky, but there is little more direct information about climatic conditions in latitude and longitude values. For example, they do not say whether the location is over land or sea. In Western Australia, the climate is usually drier further inland. There are two areas of higher annual rainfall, one in the South West corner, and one in the tropical North. Since clouds are needed to generate rain, a correlation between cloudiness patterns and rainfall patterns can be expected. So if location coordinates are represented by a distance along the coast from a set starting position, and then a distance inland from this point on the coast, the relationship between seasonal solar radiation and location might take a simpler form. The model was developed to mimic this geographic pattern of annual rainfall.

The development of this new model indicates that the approach of using a pure cloudiness metric and alternate positional coordinates can provide an algorithm that is simple enough to use interactively and provide a viable alternative to using satellite data. There is scope for adaptation to other parts of the world with a similar pattern of declining inland rainfall, such as the West coast of North America, and the West coast of South Africa.

2. Irradiance in clear sky conditions

The first stage in developing a model for irradiance in cloudy conditions is to estimate irradiances for clear sky conditions. This has already been done in previously developed theory (e.g. Kumar *et al.* [23], Gates [24] or Liu and Jordan [11]). Later, in the methods section, new equations for modifying the clear sky irradiances due to the effects of cloud are developed. A table of the symbols used in this paper can be found in Appendix D.

The solar altitude angle, α, the vertical angle of the sun to the surface of the Earth, affects the amount of solar energy reaching the Earth not only because of geometrical effects but also because of the amount of air the light has to travel through. The sine of the solar altitude angle, $\sin\alpha$, can be calculated directly from the day of the year, time of day, and estimates of variation in Earth's orbit (see Appendix A). The air mass ratio, M, is the ratio of the atmospheric path length a beam of light must travel through, compared to the path length if the sun was at the zenith (i.e. the shortest path length). One of the most numerically simple approximation formulas for M as a function of $\sin\alpha$ comes from Young [25]:

$$M = \frac{1.002432\sin^2\alpha + 0.148386\sin\alpha + 0.0096467}{\sin^3\alpha + 0.149864\sin^2\alpha + 0.0102963\sin\alpha + 0.000303978} \quad (3)$$

This approximation has the advantage that the solar altitude angle does not have to be explicitly calculated.

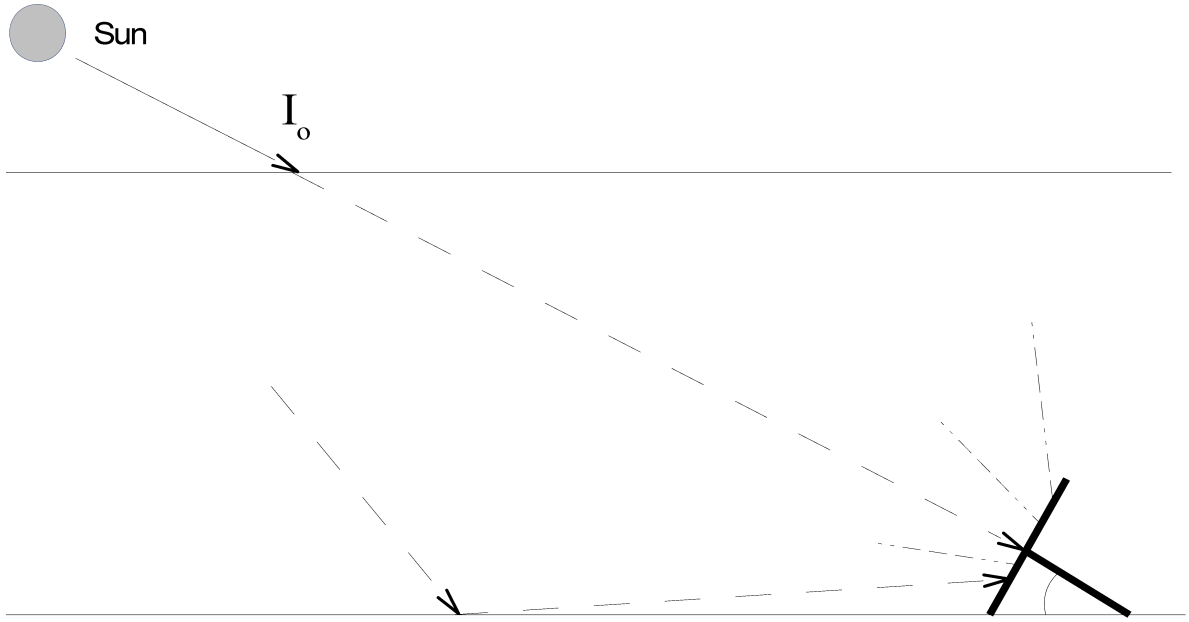

Figure 1. The three types of irradiance falling upon a surface perpendicular to the position of the sun in the sky. I_o is the extraterrestrial beam irradiance. I_b is the beam irradiance falling on the surface. I_d is the diffuse irradiance falling on the surface. I_r is the reflected irradiance falling on the surface. α is the solar altitude angle.

The total (or global) irradiance falling on the Earth's surface at the ground can be divided into three types: the beam irradiance, I_b, the diffuse irradiance, I_d, and the reflected irradiance, I_r (Figure 1). Beam irradiance comes directly from the sun, and is also called direct irradiance. The extraterrestrial beam irradiance, I_o, is the irradiance arriving at the top of the atmosphere on a plane perpendicular to the direction of the sun. The fraction of of incident light that travels through a medium is called a transmittance. The beam irradiance arriving at the Earth's surface, I_b, also on a plane perpendicular to the direction of the sun, can be related to I_o by defining a beam radiation atmospheric transmittance, t_b, such that:

$$I_b = t_b I_o \quad W/m^2 \tag{4}$$

t_{bcs} is the beam atmospheric transmittance under clear sky conditions. At very high elevations with extremely clear air, t_{bcs} may be as high as 0.8, while for a clear sky with high turbidity it may be as low as 0.4 [24]. A simple formula based on Bourghers law [26] can be used to estimate t_{bcs} from M:

$$t_{bcs} = K_{b1} e^{K_{b2} M} \tag{5}$$

where K_{b1} and K_{b2} are constants. K_{b2} is an absorption constant and is negative. If a surface is oriented in a direction other than perpendicular to the sun's position in the sky, then the

incident beam irradiance will be reduced. Many data sets and many studies deal only with measurements of solar radiation falling on a horizontal surface. Therefore expressions for horizontal surface radiation must be developed if comparison with these models is to be made. The beam irradiance falling on a horizontal surface, I_{hb}, will be:

$$I_{hb} = I_o t_b \sin\alpha \quad W/m^2 \tag{6}$$

Diffuse radiation arises from beam radiation that has been scattered by the atmosphere. The diffuse irradiance I_d on a plane at the Earth's surface can be related to the extraterrestrial beam irradiance I_o by defining a diffuse atmospheric transmittance, t_d such that:

$$I_d = 0.5 I_o t_d \sin\alpha \left(1 + \cos\left(tilt\right)\right) \quad W/m^2 \tag{7}$$

where tilt is the vertical angle of the plane compared to the horizontal. This equation implies that for surfaces tilted at angles other than horizontal, the diffuse irradiance will be less because the surface will not "see" the full sky hemisphere [27]. If the surface is vertical, then the diffuse radiation falling on it will be halved, as it only "sees" half of the full sky hemisphere. Lui and Jordan [11] formulated the following relationship between clear sky diffuse transmittance t_{dcs} and the beam atmospheric transmittance under clear sky conditions t_{bcs}:

$$t_{dcs} = 0.271 - 0.294 t_{bcs} \tag{8}$$

This equation implies that a higher beam transmittance means less diffuse radiation, as would be expected. Beam transmittance t_{bcs} for a dust free clear sky typically ranges from 0.4 to 0.8, and the corresponding diffuse transmission coefficient t_{dcs} ranges from 0.153 to 0.037 [24]. The diffuse irradiance falling on a horizontal surface will therefore be:

$$I_{hd} = I_o \left(0.271 - 0.294 t_{bcs}\right) \sin\alpha \quad W/m^2 \tag{9}$$

For a surface under consideration, a part of both the beam and diffuse components of clear sky radiation may be reflected by the surroundings. Based on a formula given by Gates [24], the reflected radiation falling on a surface, I_r, in isotropic surroundings can be related to I_o by:

$$I_r = 0.5 r_g I_o t_r \sin\alpha \left(1 - \cos\left(tilt\right)\right) \quad W/m^2 \tag{10}$$

where t_r is the reflectance transmittance, and r_g is the ground reflectance coefficient averaged over the solar wavelength spectrum (see Monteith and Unsworth [28]). Gates [24] gives a reflectance coefficient for vegetation of 0.2, and also a formula that relates the

clear sky reflectance transmittance t_{rcs} to the clear sky beam transmittance t_{bcs}:

$$t_{rcs} = 0.271 + 0.706\, t_{bcs} \tag{11}$$

such that if clear sky beam transmittance increases, then so does clear sky reflectance transmittance. If the surroundings are not isotropic, then in principle it is still possible to explicitly model reflected radiation, but the equations will be complex and site specific. For a horizontal surface, tilt is zero, and the estimate for reflected radiation from equation 10 is zero. In this case:

$$I_{hcs} = I_o \left(t_{bcs} + t_{dcs} \right) \sin\alpha \;\; W/m^2 \tag{12}$$

where I_{hcs} is the clear sky irradiance falling on a horizontal surface. Note that multiple reflection effects between the Earth's surface and the atmosphere are present even if the tilt is zero. These effects can be modelled or made implicit in the diffuse transmittance estimation. For an example see Suckling and Hay [29]. Some authors have included an explicit term for these effects [30]. It may be more important to have an explicit term if multiple cloud layers are modelled, as for example in Yang and Koike [20], and there is a need to differentiate between intrinsic cloud behaviour and ground-atmosphere reflection effects.

3. Method

So far, theoretical equations for the transmittances t_b, t_d, and t_r have been formulated for clear skies, based on results from previous studies. To effectively model radiation under all weather conditions, there is a need to quantify the behaviour of these transmittances under cloudy conditions. In this study, a new metric of cloudiness was defined, its effects on t_b, t_d, and t_r were quantified, and an algorithm for generating synthetic cloudiness values at any location in the South West corner of Western Australia was developed. The algorithm was calibrated and verified using cloudiness values obtained from hourly horizontal surface radiation data measured at a set of meteorological stations. Use of the word 'horizontal' refers to radiation falling on a horizontal surface.

3.1 Defining cloudiness

Previous studies have attempted to quantify cloudiness from observation (e.g. Muneer *et al.* [5]). However, it is not simply the amount of cloud that matters. Cloud type, height, distribution, and layering will also affect irradiance in an ever-changing manner. To maintain simplicity, this study did not differentiate between these different aspects of cloudiness. Instead the level of cloudiness was quantified by how much beam transmittance is affected. A cloudiness transmission factor, t_c, was defined as a fractional

modifier to the clear sky beam transmittance, t_{bcs}, due to cloudy conditions:

$$t_b = t_c t_{bcs} \tag{13}$$

The "cloudiness"(c) was defined such that the cloudiness transmission factor t_c is reduced by increasing cloudiness until t_c is zero when c is one:

$$t_c = 1 - c \tag{14}$$

How much will cloudiness affect the diffuse transmittance, t_d? A diffuse cloudiness function, $k_{cloud}(c)$, was defined that modifies the clear sky diffuse transmittance t_{dcs}:

$$t_{gh} = t_b + t_d = (1 - c) t_{bcs} + k_{cloud}(c) t_{dcs} \tag{15}$$

where t_{gh} is the global irradiance transmittance for radiation falling on a horizontal surface. The observation by Butt *et al.* [21] that the diffuse fraction, k_d, increases in a straight-line fashion with cloudiness was represented using:

$$k_d = \frac{t_d}{t_{gh}} = k_{dcs}(1 + K_{dr} c) \tag{16}$$

where k_{dcs} is the clear sky diffuse fraction, and K_{dr} is a slope constant. Combining equations 15 and 16 gives

$$k_{cloud}(c) = \frac{(1 - c)(1 + K_{dr} c)}{1 - K_{dr} c \dfrac{t_{dcs}}{t_{bcs}}} \tag{17}$$

If the value of K_{dr} is high enough, then for some values of t_{bcs}, the k_{cloud} function increases with increasing cloudiness before reaching a maximum and then decreasing (Figure 2), and there are two possible values of cloudiness for the same value of k_{cloud}.

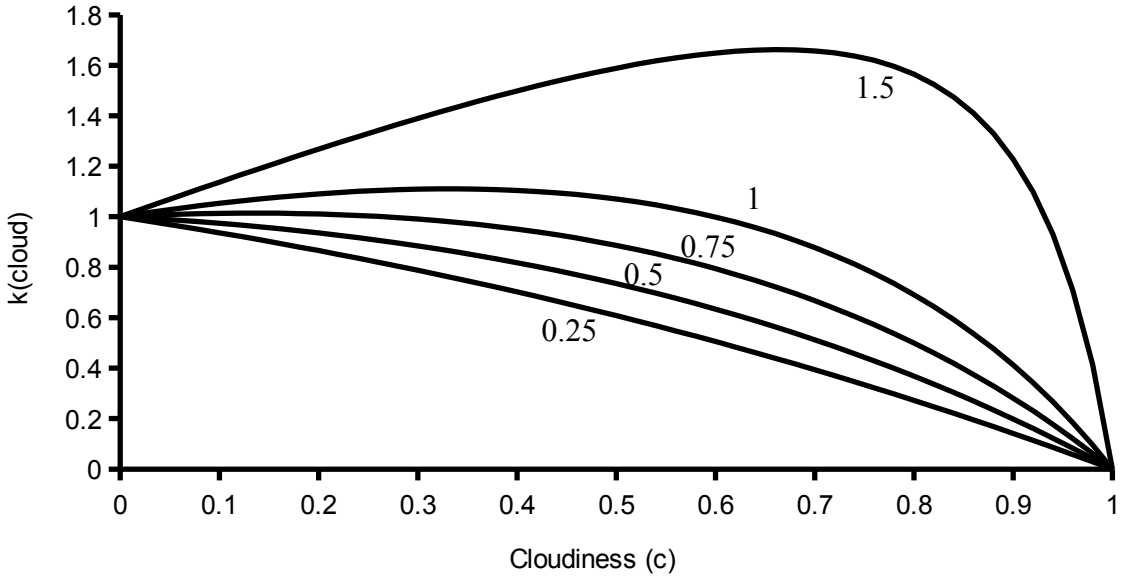

Figure 2. k_{cloud} as a function of cloudiness for $\sin\alpha = 0.1$, M = 9, $t_{bcs} = 0.3$ and $t_{cds} = 0.182$. The 4 curves are for 4 different values of K_{dr}: 0.25, 0.5, 0.75, 1 and 1.5.

How much will cloudiness affect the reflectance transmittance, t_r? Hay [31] implies that the reflected irradiance under cloudy conditions remains proportional to the global horizontal irradiance, and hence the reflectance transmittance will remain equal to the global horizontal transmittance:

$$t_r = t_{gh} = t_b + t_d = (1-c)\,t_{bcs} + k_{cloud}(c)\,t_{dcs} \tag{18}$$

3.2 Estimating cloudiness

To quantify the range and distribution of cloudiness over the South West of Western Australia, a way to estimate cloudiness from measured data was developed. Hourly horizontal solar radiation data from a network of 31 W.A. Department of Agriculture weather stations for the years 2007 and 2008 were obtained. The data were originally stored in units of kJ/m^2 for 1 hour of radiation, so each value was divided by 3.6 to obtain a value for average hourly irradiance in units of W/m^2. Instrumentation error has the greatest relative effect when the sun is near the horizon and measured values of radiation are low. To minimise this error, only the data collected when the sine of the solar altitude angle, $\sin\alpha$, was greater than 0.1 were used.

Values for global horizontal transmittance, t_{ghm}, were obtained from these measurements using

$$t_{ghm} = \frac{I_{hm}}{I_{hex}} \tag{19}$$

where I_{hm} is the measured hourly horizontal solar irradiance, and I_{hex} is the extraterrestrial horizontal irradiance. I_{hex} was calculated using:

$$I_{hex} = S_o OF \sin\alpha \quad W/m^2 \tag{20}$$

where S_o is the solar constant and OF is the orbital correction factor. These can be calculated or estimated from the day of the year, time of day, and location (see Appendix A). The solar constant S_o has been the subject of some debate. A value of 1367 W/m^2 has been widely used and was adopted in this study, but is subject to small variations due to solar activity [32].

The cloudiness was estimated by combining equations 15 and 17 to give:

$$c = \frac{t_{gcs} - t_{ghm}}{t_{gcs} + K_{dr}\left(1 - \frac{t_{gcs}}{t_{bcs}}\right) t_{ghm}} \tag{21}$$

The value of K_{dr} was set to 0.8 so that the denominator of equation 21 was always positive for all values of global horizontal transmittance t_{ghm} obtained from measurements.

The daily average cloudiness c_d was calculated as the average of the hourly cloudiness values for those hours when the sun was above the horizon. If c_d was less than 0.05 and no hourly value of cloudiness was greater than 0.1, then the day was categorised as a clear sky day. A total of 4141 days (out of 21422 days of measured data across the 31 stations) were clear sky days.

A downhill simplex error minimisation algorithm [33] was used to find the optimum values for K_{b1} and K_{b2} (see equation 5) that gave the best fit between the theoretical clear sky irradiance and the measured irradiance on the chosen clear sky days. To prevent negative cloudiness transmission factors, an additional weighting term was introduced within the error function calculation of the simplex algorithm so that if on any hour of any day (not just the clear sky days) the measured cloudiness using equation 21 exceeded 1 ($t_c < 0$), then the error value was dramatically increased.

3.3 Generation of daily cloudiness

The measurements for each day from each station were divided into two data sets, a training set and a test set (also called the calibration set and the validation set). The

training set was used to establish the parameters of the algorithm. The synthetic data generated by the algorithm was then compared to the test data points. Because of the different ways that models from other studies have separated data for calibration and validation, two different methods of separating data were used in this study. In the first configuration, the measured data from each station was split almost equally into calibration data and validation data by assigning each day of data to one of the two groups using a pseudorandom number generator. In the second configuration, all the data from 7 stations (or about one quarter of the total) were reserved for validation data to test the algorithm. All the data from other 24 stations were used as training data to calibrate the algorithm.

Following the lead of Boland [8] concerning seasonal variation in radiation, the monthly average of the mean daily cloudiness, and the monthly standard deviation in mean daily cloudiness were represented as varying sinusoidally with month:

$$cdavmon \approx K_{cd1} + K_{cd2} \sin\left(\frac{\pi}{6}\left(month + K_{cd3}\right)\right) \tag{22}$$

$$c_{dsdmon} \approx K_{cd4} + K_{cd5} \sin\left(\frac{\pi}{6}\left(month + K_{cd6}\right)\right) \tag{23}$$

where month is the month of the year (from 1 to 12). The sinusoid coefficients K_{cdi} (i = 1 to 6) were estimated for each station in the training data set by using the downhill simplex method to minimise the sum of the absolute errors between measured and estimated values of c_{davmon} and c_{dsdmon}.

In order to estimate the value of c_{davmon} and c_{dsdmon} for any location in Western Australia, a geographic relationship for the K_{cdi} coefficients was established. A shape map of the Western Australian Coastline was constructed from the GEODATA COAST 100K 2004 data package published by Geoscience Australia [34]. This data set is based on a 1:100,000 scale map sheet. The shape map consists of a vector map of the coastline and state border in longitude and latitude coordinates. It does not include any of the islands off the coast of Western Australia that are included in the data package. Note that the distance along a coastline will depend on the precision of the coastline representation, with higher precisions generating longer distances. There were 74032 vertices in the constructed shape map, so a global simplification algorithm [35] was used to simplify the map down to a 500 vertex coastline map (Figure 3). The coastline is represented by a straight line segment drawn between each pair of vertices.

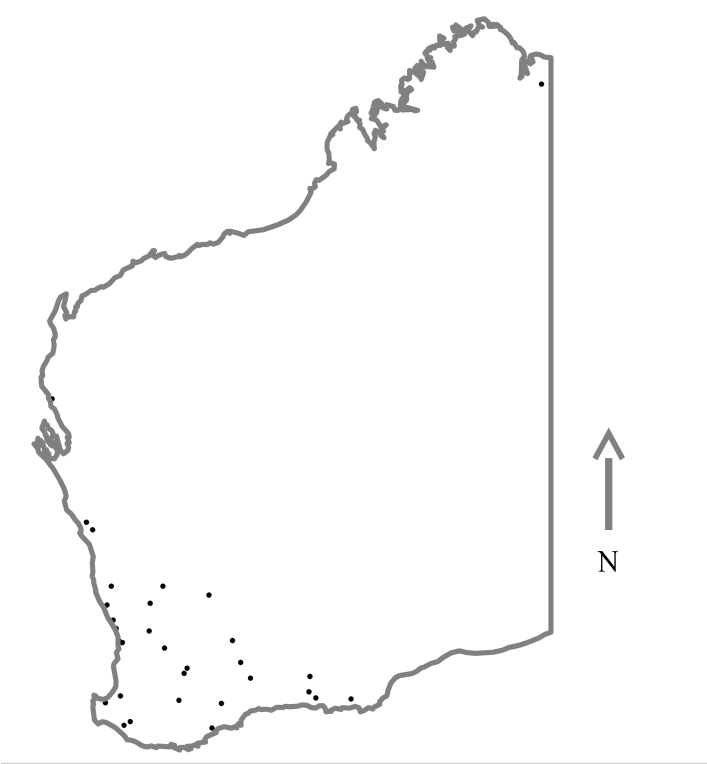

Figure 3. Simplified Coastline of Western Australia and Locations of Meteorological Stations.

The intersection between the coastline and the state border with the Northern Territory in the North East corner was used as the starting point for calculating coastal distance. The position of each meteorological station in (cpos,cdist) coordinates, where cpos is the distance along the coastline and cdist is the shortest distance to the coast, was calculated from the latitude and longitude coordinates using Euclidean geometry.

To establish the geographic relationship, the value of each coefficient K_{cdi} (i = 1,6) was assumed to vary in a way roughly mimicking the rainfall pattern in Western Australia. That is, in a piecewise straight-line fashion along the coastline, and with a combination of initial increase and then exponential decay as distance inland from the coast increases. The value of each coefficient is estimated by interpolation using a number of setpoints positioned along the coast. See Appendix A for the detailed algorithm. The downhill simplex algorithm was used to optimise the number, position, and value of the setpoints such that the overall sum of the absolute errors between estimated and measurement derived values of the K_{cdi} coefficients at each of the meteorological stations in the training data set was minimised.

To establish the frequency distribution of the daily cloudiness, the residual variable y_d was calculated for each day in each month of the training data set using the measured values for c_{davmon} and c_{dsdmon}:

$$y_d = \frac{c_d(measured) - c_{davmon}(mon)}{c_{dsdmon}(mon)} \tag{24}$$

The inverse cumulative frequency distribution, Cf^{-1}, of these values of y_d was approximated for each station in the training data set using:

$$Cf^{-1}(r) \approx K_{cf1} + K_{cf2} r + K_{cf3} r^2 + K_{cf4} r^8 \tag{25}$$

where r is the cumulative frequency. One purpose of using the residual variable y_d was to remove seasonal effects. However, because of the high number of clear sky days at many stations, the residual variables for each month were not normally distributed and had predominantly positive skews. The coefficients K_{cfj} j = 1 to 4 were found to retain a seasonal dependence as well as the expected locational dependence, so they were estimated using the same technique as c_{dmonav} and c_{dmonsd}:

$$K_{cfj} \approx K_{cd(4+3j)} + K_{cd(5+3j)} \sin\left(\frac{\pi}{6}\left(month + K_{cd(6+3j)}\right)\right) \quad for\ j = 1,4 \tag{26}$$

The sinusoid K_{cdi} coefficients, i = 7 to 18, were also adjusted for location using piecewise straight-line interpolation between coastal setpoints and a combination of linear and exponential functions for distance from the coast. As before, for each K_{cdi} the downhill simplex algorithm was used to optimise the number, position, and value of the coastal setpoints, such that the overall sum of the absolute errors between estimated and measurement derived values of the K_{cdi} coefficients at each of the stations in the training set was minimised. In total, 18 K_{cdi} coefficients are generated for each location using this technique (Table 1).

Table 1. Sinusoid Coefficient K_{cdi} Indices

Variable Estimated	Mean	Amplitude	Season Shift
c_{davmon}	1	2	3
c_{dsdmon}	4	5	6
K_{cf1}	7	8	9
K_{cf2}	10	11	12
K_{cf3}	13	14	15
K_{cf4}	16	17	18

Synthetic values of the residual variable y_d were obtained by generating pseudorandom values of r with a uniform frequency distribution between 0 and 1 and then transforming

backwards:

$$y_d \approx Cf^{-1}(r) \tag{27}$$

Synthetic values of mean daily cloudiness, c_d, were calculated from y_d using:

$$c_d = c_{davmon} + c_{dcsdmon}\, y_d \tag{28}$$

3.4 Generation of hourly cloudiness

The measured hourly cloudiness values from every station were lumped together and categorised according to sine of the solar altitude angle sinα and average daily cloudiness c_d. This assumes that for a given c_d, the statistical characteristics of cloudiness are the same for any location in Western Australia. The following function was used to estimate the average hourly cloudiness, c_{hm}, as a function of sinα and c_d:

$$c_{hm}(c_d, sin\alpha) = c_d\left(1 + K_{h0}\frac{(1-c_d)}{(1+K_{h1}c_d^2)}\left(sin\alpha + K_{h2}sin^2\alpha + K_{h3}sin^3\alpha\right)\right) \tag{29}$$

The coefficients K_{h0} to K_{h3} were determined from the training data. Each measured cloudiness value was placed into one of 200 bins arranged in a 2 dimensional grid according to variations in steps of 0.05 for average daily cloudiness c_d and 0.1 for sine of the solar altitude angle sinα. The mean cloudiness for each bin was calculated. The values of K_{h0} to K_{h3} were optimised using a downhill simplex algorithm to minimise the sum of the absolute differences between the mean of the measured hourly cloudiness values in each bin and c_{hm}. See Appendix B for their values. With this function, when c_d is zero (a clear sky day), then c_{hm} is zero regardless of the value of sinα or K_{h0} to K_{h3}. Similarly, if c_d is 1, then c_{hm} is 1.

For each bin, the standard deviation of the measured cloudiness values from the mean hourly cloudiness in each bin was calculated and the following function was used to estimate the standard deviation:

$$\sigma(c_d, sin\alpha) = c_d(1-c_d)\left(\frac{K_{dv0}}{(1+K_{dv1}c_d)} + \frac{K_{dv2}}{(1+K_{dv3}c_d)}sin\alpha + \frac{K_{dv4}}{(1+K_{dv5}c_d)}sin^2\alpha\right) \tag{30}$$

The coefficients K_{dv0} to K_{dv5} were determined by using a downhill simplex algorithm to minimise the sum of the absolute differences between the measured standard deviations in each bin and $\sigma(c_d, sin\alpha)$. See Appendix B for their values. With this function, when c_d is zero (a clear sky day), then σ is zero regardless of the value of sinα or K_{dv0} to K_{dv5}. Similarly, if c_d is 1, then σ is zero.

The hourly cloudiness residual function, y_h, was calculated for every measured hourly cloudiness value using:

$$y_h = \frac{c - c_{hm}(c_d, \sin\alpha)}{\sigma(c_d, \sin\alpha)} \tag{31}$$

It was found that a translated Weibull distribution gave a better fit to the frequency distribution curve for y_h than a normal frequency distribution (Figure 4), although a zero significance value for the Kuiper statistical test [33] indicated that the distribution of y_h could not be considered to be identical to the Weibull distribution.

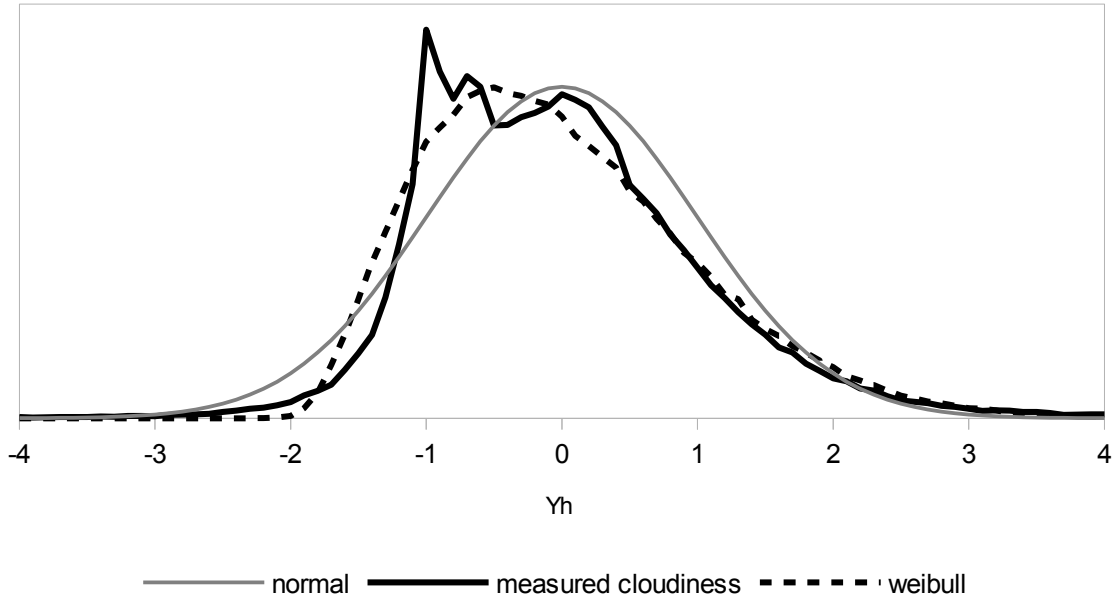

Figure 4. Frequency distribution of measured hourly cloudiness residual y_h, compared to a normal distribution and a translated Weibull distribution.

Therefore a translated Weibull cumulative frequency distribution function was used to generate synthetic values of y_h from the output of a pseudorandom number generator:

$$r_w = \theta_w + \lambda_w\left(-\ln(1-r)\right)^{\left(\frac{1}{\kappa_w}\right)} \tag{32}$$

where r is a psuedorandom number between 0 and 1 with a uniform frequency distribution. r_w has a Weibull frequency distribution with mean of 0 and standard deviation of 1. θ_w is the translation parameter, λ_w is the scale parameter, and κ_w is the shape parameter. See Appendix B for the values of these coefficients.

Although time series analysis of hourly cloudiness values is limited by the short continuous sequences available during daylight hours, calculation of partial autocorrelation coefficients of the residual y_h showed a clear dependence on the cloudiness for the previous hour. Therefore, at each station and for each day, the relative first order autocorrelation between successive y_h values was calculated. These values were placed into 1 of 10 bins according to daily average cloudiness, such that each bin has an average cloudiness range of 0.1. The following function was used to estimate the average first order autocorrelation:

$$\varphi(c_d) = K_{yac}\left(1 - 8\left(c_d - 0.5\right)^3\right) \tag{33}$$

where K_{yac} is the autocorrelation coefficient. K_{yac} was determined by minimising the sum of the absolute errors between the measured average autocorrelation values in each bin and $\varphi(c_d)$. See Appendix B for the value of K_{yac}. Synthetic hourly cloudiness values were generated with a first order regression component and a random component using the following three equations:

$$\sigma_d = \sqrt{\left(1 - \varphi(c_d)^2\right)} \tag{34}$$

$$y_h = \varphi(c_d)\, y_{h-1} + \sigma_d r_w \tag{35}$$

and

$$c = c_{hm}\left(c_d, \sin\alpha\right) + \sigma\left(c_d, \sin\alpha\right) y_h \tag{36}$$

where σ_d is the standard deviation of the random component, and c is the synthetic hourly cloudiness at hour h. For comparison with other models, synthetic values of horizontal irradiance I_h were calculated from the cloudiness using

$$t_{gh} = \left(1 - c\right) t_{bcs} + k_{cloud}\left(c\right) t_{dcs} \tag{37}$$

$$I_{hex} = 1367\, OF\, \sin\alpha \tag{38}$$

and

$$I_h = t_{gh}\, I_{hex} \tag{39}$$

See Appendix A for the complete implementation of the algorithm. The values of the coefficients in Appendix B were calculated using all of the measured data as training data.

4. Results and Discussion

In order to assess the performance of the model, model generated synthetic data and data derived from measurements were compared, and the overall differences quantified in a statistical manner. To capture seasonal and diurnal, as well as hourly, effects, the model was assessed on monthly, daily, and hourly time scales. These results were compared with other models that estimate solar radiation over a region or several locations (see below). The following statistical measures were used for comparison, depending on the study: Root mean square error (RMSE), Mean bias error (MBE), mean absolute percentage error (MAPE), standard error (SE), and mean relative variance (MRV). Definitions are given in appendix C. These measures were calculated for each station and then the average taken over all stations in the test data set. Because the results for the other studies were derived in different ways, two configurations of training and test data sets were assessed in this study (see Section 3.3). To avoid an artificially better ranking due to data set configuration, the results with the highest error were used for comparison.

4.1 Monthly average daily horizontal radiation

For each day of validation data, the daily average horizontal radiation, for both measured and synthetically generated data sets, were obtained from the hourly horizontal radiation measurements. These were in turn averaged over each month. The average RMSE between the measured and synthetic monthly values was 9.9% and the average magnitude of the MBE was 3.9%. The average error according to three measures was within the upper end of the range of several other models (Table 2). Synthetically generated values of monthly average daily radiation were also compared with satellite derived monthly averaged data [36] for 392 locations in the South West of Western Australia, compiled over the period 1990 to 2011. The RMSE was 19.2% (3.9 MJ/m^2/day), and the MBE was -3.7% (-0.75 MJ/m^2/day), indicating that the model generated radiation is conservative.

Table 2. Model comparison of monthly average daily horizontal radiation errors.

Study	Model Type	Region modelled	MAPE(%)	RMSE(%)	SE (MJ/m^2/month)
Mohandes [37]	Neural net	Saudi Arabia	10.1		
This study	Autoregressive	Southern Western Australia	8.3	9.7	55.15
Coops *et al.* [38]	Meteorological	locations in U.S.A., Glasgow, Canberra			54.75
Hutchinson *et al.* [39]	Angstrom,	Australia		5.35	

	Meteorological		
Sozen *et al.* [40]	Neural net	Turkey	5.7
Reddy and Ranjan [41]	Neural net	India	3.0
Mellit *et al.* [42]	Neural net	Algeria	1.2

4.2 Daily average horizontal radiation

For the daily horizontal radiation values obtained in Section 4.1 above, the average RMSE of the model developed in this study was 33.7% and the average magnitude of the MBE was 3.9%. In a straightforward comparison of daily horizontal radiation values, this model has a higher error than several other models (Table 3). However, it must be remembered that these other models use measured daily meteorological information, so their errors are expected to be less.

Table 3. Comparison of daily horizontal radiation.

Study	Model type	Region modelled	RMSE ($MJ/m^2/day$)	Magnitude of MBE($MJ/m^2/day$)
This study	Autoregressive	Southern Western Australia	6.3	0.68
Liu *et al.* [43]	Meteorological	China	3.95	0.04
Fortin *et al.* [44]	Meteorological and Neural net	Eastern Canada	3.74 – 5.45	
Liu and Scott [45]	Meteorological	Australia	2.89 – 3.24	
Lyons and Edwards [46]	Meteorological, measured 3 layer cloud amount	Western Australia	2.16 – 3.28	

4.3 Hourly Irradiance

The average RMSE for hourly horizontal irradiance for each day of validation data was 43.1% and the average magnitude of the MBE was 3.9%. An example of the model generated hourly horizontal irradiance data is given in Figure 5.

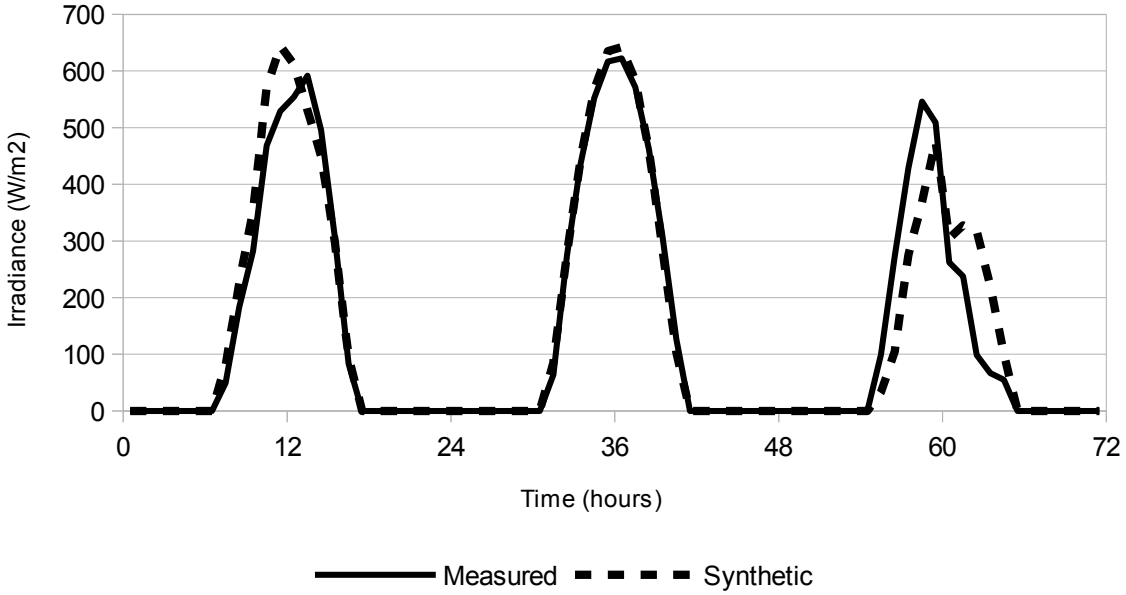

Figure 5. Sample measured and synthetic hourly horizontal irradiance for a 3 day period, starting at midnight on the first day.

The models developed by Yang and Koike [20] and Yang *et al.* [47] use measured hourly meteorological data to derive a sky clearness indicator coefficient (SCI) which is a measure of cloudiness. The SCI is used to estimate horizontal irradiance. To allow for more direct comparison with the Yang models, synthetic hourly cloudiness values were generated by the model developed in this study using measured daily average cloudiness (that is, the average calculated from the measured hourly cloudiness values), not the model generated daily average cloudiness. This will remove the errors present on diurnal and monthly time scales. Horizontal irradiance values were calculated from these hourly cloudiness values. The average RMSE and MBE magnitudes across all test data stations were 112.8 W/m^2 and 6.7 W/m^2 respectively. The RMSE is still higher than the results obtained by Yang and Koike [20] and Yang *et al.* [47], which is expected as these models use hourly measured meteorological data (Table 4).

Reikard [16] compared 6 models designed to forecast the hourly horizontal irradiance at sites in the USA at time periods from 1 to 4 hours ahead of the measured value. Some of the measured data were actually modelled data derived from measured meteorological variables, including cloud cover. Hence the results are also included here. The average Reikard model MAPE ranged from 35.18% for the most accurate model to 51.64% for the reference model. In contrast, this model achieved a lower MAPE of 29.1%, which is surprising given that only the measured average daily cloudiness was used, not previous measured hourly values. This result supports the validity of using an autoregressive approach to modelling hourly cloudiness using the value of average daily cloudiness.

Table 4. Hourly horizontal irradiance comparison of models that use a cloudiness coefficient. Results for the model developed in this study are obtained using daily average of measured hourly irradiance.

Study	Model type(s)	Region modelled	MAPE%	RMSE (W/m^2)	MBE magnitude (W/m^2)
Reikard [16]	Regression, ARIMA, Meteorological, Neural net	USA	35.18 – 51.64		
This study	Autoregressive	South West Western Australia	29.1	112.8	6.7
Yang and Koike [20]	Meteorological	Japan and Islands		96	11
Yang *et al.* [47]	Meteorological	USA and Saudi Arabia		52	14.2

The model developed in this study has errors that are within the middle of the range of several other models that do not use a measure of cloudiness (Table 5). Kambezidis *et al.* [48] reported a high RMSE, but a very low MBE for the model they developed. Hence their model may warrant further development for use in the estimation of mean hourly irradiance. Many of the neural net models achieved very low errors. However, the concerns about neural net models expressed by Reikard [16] of over fitting to noisy data must be kept in mind. The purpose of model building is not to mimic the measured data exactly, but to produce a comprehensive representation of typical measured data behaviour. These results indicate that the model developed in this study is at least comparable to other models in this task.

Table 5. Comparison with other hourly irradiance models. Results for the model developed this study are obtained using daily average of measured hourly irradiance.

Study	Model type	Region modelled	RMSE(W/m^2)	Magnitude MBE(W/m^2)	Monthly average hourly irradiance RMSE%	Magnitude Monthly average hourly irradiance MBE%	Hourly Clearness Index kt MRV
Kambezidis *et al.* [48]	Meteorological, Sunshine duration	Athens, Lisbon	164.4-256.4	1.39-5.56			
Ahmad and Tiwari [49]	Mean hourly irradiance profile	India			13.94	3.36	
Muneer and Younes [50]	Meteorological,	Several locations in	127.7	32.8			

	Sunshine duration	Europe, India, Japan and Middle-East					
Seo *et al.* [51]	Meteorological	Several world locations	120.3	10.5			
This study	Autoregressive	South west Western Australia	112.8	6.7	10.5	7.7	0.6671
Spokas and Forecella [52]	Meteorological	USA and several world locations	111				
Hontoria *et al.* [53]	Neural net	Jaen and Andalucia provinces in Spain					0.1153

Model generated values for the hourly clearness index k_t (using measured daily average cloudiness) were calculated using equation 37. Values of k_t were also generated using the Aguiar and Collares-Pereira TAG algorithm [14]. The distribution of both these synthetic datasets were compared to measurement derived values of k_t (Figure 6). The RMSE between the synthetic values of the standard deviation of the hourly k_t values, calculated for each day of measurement data, was 0.061 for the TAG algorithm and 0.043 for the model developed in this study. These two results indicate that the model did not match the distribution of the measurement values precisely, but performed slightly better than the TAG algorithm on the same Western Australian dataset.

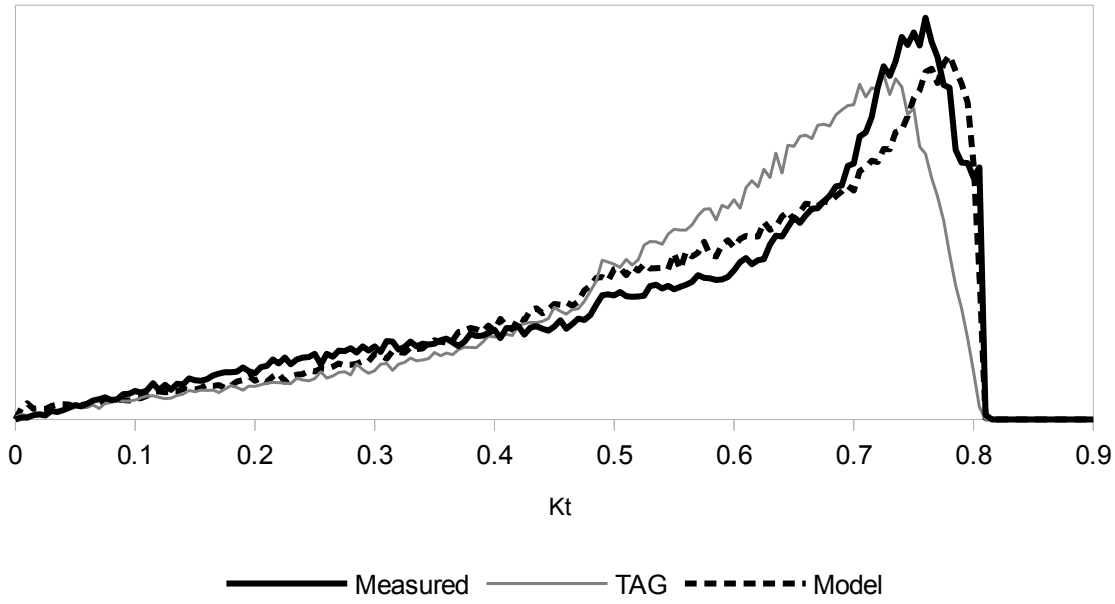

Figure 6. Frequency distribution of clearness index k_t from measured horizontal radiation, synthetically generated using the Aguiar and Collares-Pereira TAG algorithm, and derived from model cloudiness values.

5. Conclusions

This model was designed to simulate solar irradiance on an hourly time scale for any location in the South West region of Western Australia. The algorithm was simple enough to be implemented inside a web page, and the results indicate that it can generate synthetic hourly horizontal radiation data with accuracies within the range of other models developed for wide areas or several locations. Because values of cloudiness are generated rather than global horizontal irradiance, the beam, diffuse and reflected components of irradiance can be obtained from the clear sky values, meaning that this model can be used in simulations of different kinds of solar power devices and different tilt angles to horizontal. This model has direct applicability to both thermal and PV concentrating solar power systems.

The algorithm generated hourly cloudiness data with reasonably similar statistical characteristics to the measured data. However there are likely to be areas where local climatic conditions will produce measurements with significantly different cloudiness characteristics. Detailed measurement and analysis will be required to obtain better accuracy. Data from weather stations concentrated in the Southern and Western region of Western Australia have been used to test the simulation, because this area is more significant for the simulation of broadscale power generation due to the proximity of grid

infrastructure and major load centres. The representation of more remote regions in Western Australia is likely to be poorer, but will be improved by incorporating measurement data from these regions.

The approach of using alternative position metrics to latitude and longitude to simplify the required calculations resulted in a model that uses 337 coefficients, or 837 including the 500 vertex coastline. This is equivalent to the direct representation of just over 2.2 months of hourly data at one location. As such this approach provides a viable alternative to using satellite data and has scope for adaptation to other parts of the world with a similar pattern of declining inland rainfall, such as the West coast of North America, and the West coast of South Africa. The model coefficients are listed in Appendix B. The coordinates for the 500 vertex shape map of the Western Australian Coastline are available upon request.

The intended purpose of developing a synthetic solar radiation data generator simple enough to be suitable for use in web based interactive simulations of broadscale solar power systems in the South Western part of Western Australia was achieved.

Acknowledgements

The assistance of Dr. Ian Foster, Department of Agriculture and Food, Western Australia, Jatin Sala, and Tom Lyons in providing the Department of Agriculture meteorological station data is gratefully appreciated.

References

[1] Mora-Lopez LL, Sidrach-de-Cardona M. Multiplicative ARMA models to generate hourly series of global irradiation. Solar Energy 1998;63:283–91.

[2] Hontoria L, Aguilera J, Zufiria P. Generation of hourly irradiation synthetic series using the neural network multilayer perceptron. Solar Energy 2002;72:441–6.

[3] Graham VA, Hollands KGT, Unny TE. A time series model for Kt with application to global synthetic weather generation. Solar Energy 1988;40:83–92.

[4] Gazela M, Mathioulakis E. A new method for typical weather data selection to evaluate long-term performance of solar energy systems. Solar Energy 2001;70:339–48.

[5] Muneer T, Younes S, Munawwar S. Discourses on solar radiation modeling. Renewable and Sustainable Energy Reviews 2007;11:551–602.

[6] Semenov MA, Brooks RJ, Barrow EM, Richardson CW. Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates. Climate Research 1998;10:95–107.

[7] Punyawardena BVR, Kulasiri D, Biometrics LU (Canterbury, N. Z.) Centre for Computing and. Stochastic Simulation of Solar Radiation from Sunshine Duration in Sri Lanka. Centre for

Computing and Biometrics; 1996.
[8] Boland J. Time-series analysis of climatic variables. Solar Energy 1995;55:377–88.
[9] Celik AN. The system performance of autonomous photovoltaic–wind hybrid energy systems using synthetically generated weather data. Renewable Energy 2002;27:107–21.
[10] Gordon JM, Reddy TA. Time series analysis of daily horizontal solar radiation. Solar Energy 1988;41:215–26.
[11] Liu BYH, Jordan RC. The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. Solar Energy 1960;4:1–19.
[12] Mora-López L, Sidrach-De-Cardona M. Characterization and simulation of hourly exposure series of global radiation. Solar Energy 1997;60:257–70.
[13] Graham VA, Hollands KGT. A method to generate synthetic hourly solar radiation globally. Solar Energy 1990;44:333–41.
[14] Aguiar R, Collares-Pereira M. TAG: A time-dependent, autoregressive, Gaussian model for generating synthetic hourly radiation. Solar Energy 1992;49:167–74.
[15] Remund J, Salvisberg E, Kunz S. On the generation of hourly shortwave radiation data on tilted surfaces. Solar Energy 1998;62:331–44.
[16] Reikard G. Predicting solar radiation at high resolutions: A comparison of time series forecasts. Solar Energy 2009;83:342–9.
[17] Hollands KGT, Huget RG. A probability density function for the clearness index, with applications. Solar Energy 1983;30:195–209.
[18] Skartveit A., Olseth J.A., Tuft M.E. An hourly diffuse fraction model with correction for variability and surface albedo. Solar Energy 1998;63:173–83.
[19] Perez R, Ineichen P, Seals R, Zelenka A. Making full use of the clearness index for parameterizing hourly insolation conditions. Solar Energy 1990;45:111–4.
[20] Yang K, Koike T. Estimating surface solar radiation from upper-air humidity. Solar Energy 2002;72:177–86.
[21] Butt N, New M, Malhi Y, da Costa ACL, Oliveira P, Silva-Espejo JE. Diffuse radiation and cloud fraction relationships in two contrasting Amazonian rainforest sites. Agricultural and Forest Meteorology 2010;150:361–8.
[22] Lanini F. Division of global radiation into direct radiation and diffuse radiation. Masters Thesis, Faculty of Science, University of Bern, 2010.
[23] Kumar L, Skidmore AK, Knowles E. Modelling topographic variation in solar radiation in a GIS environment. International Journal of Geographical Information Science 1997;11:475–97.
[24] Gates DM. Biophysical Ecology. Springer-Verlag; 1980.
[25] Young AT. Air mass and refraction. Appl Opt 1994;33:1108–10.
[26] Kreith F, Kreider JF. Principles of solar engineering. Hemisphere Pub. Corp.; 1978.
[27] Rudy M. Sun and Climate Modeling for Thermal Simulation. Proceedings: Building Simulation'97, International Conference of IBPSA, Prague, 1997.
[28] Monteith JL, Unsworth MH. Principles of Environmental Physics. Butterworth-Heinemann; 1990.
[29] Suckling PW, Hay JE. A cloud layer-sunshine model for estimating direct, diffuse and total

solar radiation. Atmosphere 1977;15:194–207.
[30] Davies JA, Abdel-Wahab M, Howard JE. Cloud Transmissivities for Canada. Monthly Weather Review 1985;113:338–48.
[31] Hay JE. Calculation of monthly mean solar radiation for horizontal and inclined surfaces. Solar Energy 1979;23:301–7.
[32] Gueymard CA. The sun's total and spectral irradiance for solar energy applications and solar radiation models. Solar Energy 2004;76:423–53.
[33] Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical Recipes in C: The Art of Scientific Computing. Second Edition. 1992.
[34] Geoscience Australia 2004. GEODATA COAST 100K 2004 dataset ANZCW0703006621. Available from: http://www.ga.gov.au/meta/ANZCW0703006621.html; 2012.
[35] Visvalingam M, Whyatt JD. Line generalisation by repeated elimination of points. Cartographic Journal, The 1993;30:46–51.
[36] Bureau Of Meteorology. Mean monthly, season and annual solar exposure data. Available from: http://www.bom.gov.au/jsp/ncc/climate_averages/solar-exposure/index.jsp; 2012.
[37] Mohandes M, Balghonaim A, Kassas M, Rehman S, Halawani TO. Use of radial basis functions for estimating monthly mean daily solar radiation. Solar Energy 2000;68:161–8.
[38] Coops NC, Waring RH, Moncrieff JB. Estimating mean monthly incident solar radiation on horizontal and inclined slopes from mean monthly temperatures extremes. Int J Biometeorol 2000;44:204–11.
[39] Hutchinson MF, Booth TH, McMahon JP, Nix HA. Estimating monthly mean valuesof daily total solar radiation for Australia. Solar Energy 1984;32:277–90.
[40] Sözen A, Arcaklıoğlu E, Özalp M, Çağlar N. Forecasting based on neural network approach of solar potential in Turkey. Renewable Energy 2005;30:1075–90.
[41] Reddy K., Ranjan M. Solar resource estimation using artificial neural networks and comparison with other correlation models. Energy Conversion and Management 2003;44:2519–30.
[42] Mellit A, Benghanem M, Arab AH, Guessoum A. A simplified model for generating sequences of global solar radiation data for isolated sites: Using artificial neural network and a library of Markov transition matrices approach. Solar Energy 2005;79:469–82.
[43] Liu X, Mei X, Li Y, Wang Q, Jensen JR, Zhang Y. Evaluation of temperature-based global solar radiation models in China. Agricultural and Forest Meteorology 2009;149:1433–46.
[44] Fortin JG, Anctil F, Parent LÉ, Bolinder MA. Comparison of empirical daily surface incoming solar radiation models. Agricultural and Forest Meteorology 2008;148:1332–40.
[45] Liu DL, Scott BJ. Estimation of solar radiation in Australia from rainfall and temperature observations. Agricultural and Forest Meteorology 2001;106:41–59.
[46] Lyons TJ, Edwards PR. Estimating global solar irradiance for Western Australia, part I.

Archives for Meteorology, Geophysics, and Bioclimatology Series B 1982;30:357–69.
[47] Yang K, Koike T, Ye B. Improving estimation of hourly, daily, and monthly solar radiation by importing global data sets. Agricultural and Forest Meteorology 2006;137:43–55.
[48] Kambezidis HD, Adamopoulos AD, Sakellariou NK, Pavlopoulos HG, Aguiar R, Bilbao J, et al. The "Meteorological Radiation Model." Bull Hell Assoc Chart Mech-Electr Engineers 1998;3:38–42.
[49] Ahmad MJ, Tiwari GN. Study of models for predicting the mean hourly global radiation from daily summations. Open Environmental Journal 2008;2:6–14.
[50] Muneer T, Younes S. The all-sky meteorological radiation model: proposed improvements. Applied Energy 2006;83:436–50.
[51] Seo D, Huang J, Krarti M. Development of Models for Hourly Solar Radiation Prediction. ASHRAE Transactions 2008;114:392.
[52] Spokas K, Forcella F. Estimating hourly incoming solar radiation from limited meteorological data 2009.
[53] Hontoria L, Aguilera J, Zufiria P. An application of the multilayer perceptron: solar radiation maps in Spain. Solar Energy 2005;79:523–30.

Appendix A. The synthetic generation algorithm

The following algorithm for generating synthetic hourly cloudiness at any location was used. Cloudiness mean and variance change throughout the day as a function of average daily cloudiness and solar altitude angle. This algorithm is split into 4 sections. The first section must be computed once per location:

(1.1) From the latitude and longitude of the location (lon, lat), use Euclidean geometry and the coastline shape map coordinate data to calculate the coast position cpos and distance from the coast cdist (both in km). The coastline shape map consists of 500 vertices in longitude and latitude coordinates (lon_i,lat_i), i = 1 to 500. The first vertex is the start of the coastline and is where the coast crosses the Northern Territory border. The coastline is approximated by a set of line segments, each defined by a pair of adjacent vertices (lon_i,lat_i) (lon_{i+1},lat_{i+1}), i = 1 to 499.

(i) Calculate:

$$x = 111.195\left(lon - 129\right)\cos\left(\frac{\pi}{180}lat\right)$$
$$y = 111.195\,lat \tag{A1}$$

(ii) Set cline = 0 and $d_{min} = 10^8$. Starting at i = 0, perform the following iteration steps (iii) to (vi) for each line segment in turn:

(iii) Calculate:

$$x_i = 111.195\left(lon_i - 129\right)\cos\left(\frac{\pi}{180}lat_i\right)$$
$$y_i = 111.195\,lat_i \tag{A2}$$

$$x_{i+1} = 111.195\left(lon_{i+1} - 129\right)\cos\left(\frac{\pi}{180}lat_{i+1}\right)$$
$$y_{i+1} = 111.195\,lat_{i+1} \tag{A3}$$

$$d=\left(x_{i+1}-x_i\right)^2+\left(y_{i+1}-y_i\right)^2$$

$$d_1=\left(x-x_i\right)^2+\left(y-y_i\right)^2$$

$$d_2=\left(x-x_{i+1}\right)^2+\left(y-y_{i+1}\right)^2 \tag{A4}$$

$$f=\frac{1}{2}\left(1+\frac{\left(d_1-d_2\right)}{d}\right)$$

(iv)

$$\begin{gathered} if\ f<0,\ set\ d_m=d_1 \text{ and } f=0 \\ else\ if\ f>1,\ set\ d_m=d_2 \text{ and } f=1 \\ else\ set\ d_m=d_1-f^2 d \end{gathered} \tag{A5}$$

(v)

$$If\ d_m<d_{min},\ set\ d_{min}=d_m \text{ and } cpos=cline+f\times len \tag{A6}$$

(vi) Update cline:

$$cline=cline+len \tag{A7}$$

Go back to (iii) and repeat the calculations for the next line segment, until all are done.

(vii) Calculate:

$$cdist=\sqrt{\left(d_{min}\right)}\ \text{km} \tag{A8}$$

After the calculation has been performed on all line segments, cpos will be the distance from the start to the point along the coastline (in km) that is closest to the location. cdist will be the distance (in km) from this point to the location.

(1.2) Using the coastal setpoints, piecewise straight-line interpolation, and exponential function, calculate the coefficients K_{cdi} i = 1 to 18. For each K_{cdi}, there is a set of N_{kmi} coastal setpoints ($cpos_{i,j}$,$K_{mi,j}$) j = 1 to N_{kmi} (see Appendix B For the values of these coefficients). Find the two adjacent coastal positions $cpos_{i,j}$ and $cpos_{i,j+1}$ such that cpos lies between the two. Using straight-line interpolation, calculate K_{cpi}:

$$K_{cpi}=K_{mi,j}+\left(K_{mi,j+1}-K_{mij}\right)\frac{\left(cpos-cpos_{i,j}\right)}{\left(cpos_{i,j+1}-cpos_{i,j}\right)} \tag{A9}$$

Calculate K_{cdi} using the distance from the coast cdist and two additional distance parameters $K_{mi,Nkmi+1}$ and $K_{mi,Nkmi+2}$:

$$K_{cdi} = K_{cpi}\left(1 + K_{mi,Nkmi+1}\, cdist\right) e^{-K_{mi,Nkmi+2}\, cdist} \tag{A10}$$

(1.3) Calculate the cosine and sine of the latitude:

$$\begin{aligned} coslat &= \cos(lat) \\ sinlat &= \sin(lat) \end{aligned} \tag{A11}$$

where lat is the latitude.

(1.4) For surfaces tilted at angles other than horizontal, calculate the following parameters:

$$\begin{aligned} cosazr &= \cos(azimuth) \\ sinazr &= \sin(azimuth) \\ clca &= coslat\ \ cosazr \\ cics &= sinlat\ \ cosazr \end{aligned} \tag{A12}$$

where azimuth is the horizontal angle between a line perpendicular to the surface and a line running due north, with angles east of north being positive and west of north being negative. For surfaces with dual axis sun tracking, it is not necessary to calculate these parameters. For fixed surfaces, the following two parameters can be calculated here instead of hourly:

$$\begin{aligned} costilt &= \cos(tilt) \\ sintilt &= \sin(tilt) \end{aligned} \tag{A13}$$

where tilt is the vertical angle of the surface from the horizontal.

The second section must be computed at the beginning of each month:

(2.1) Calculate the monthly mean average daily cloudiness and monthly standard deviation of the average daily cloudiness using:

$$c_{davmonth} = K_{cd1} + K_{cd2} \sin\left(\frac{\pi}{6}\left(month + K_{cd3}\right)\right)$$
$$\sigma_{month} = K_{cd4} + K_{cd5} \sin\left(\frac{\pi}{6}\left(month + K_{cd6}\right)\right) \quad \text{(A14)}$$

(2.2) Calculate the 4 cumulative frequency distribution coefficients using:

$$K_{cfj} = K_{cd(4+3j)} + K_{cd(5+3j)} \sin\left(\frac{\pi}{6}\left(month + K_{cd(6+3j)}\right)\right) \; for\; j = 1,4 \quad \text{(A15)}$$

The third section must be calculated at the beginning of each day:

(3.1) Generate a random number r, with a uniform distribution between 0 and 1.

(3.2) Calculate the average daily cloudiness normalised residual value y using

$$y = K_{cf1} + K_{cf2}\, r + K_{cf3}\, r^2 + K_{cf4}\, r^8 \quad \text{(A16)}$$

(3.3) Calculate the average daily cloudiness using:

$$c_d = c_{davmonth} + \sigma_{month}\, y \quad \text{(A17)}$$

(3.4) Calculate the autocorrelation coefficient ϕ using:

$$\varphi = K_{yac}\left(1 - 8\left(c_d - 0.5\right)^3\right) \quad \text{(A18)}$$

See Appendix B for the value of K_{yac}.

(3.5) Calculate the random component standard deviation σ_d using:

$$\sigma_d = \sqrt{\left(1 - \varphi^2\right)} \quad \text{(A19)}$$

(3.6) Calculate the equation of time (EOT) and the time of day at solar noon, tod_{noon}:

$$B = \frac{2\pi}{365}(doy - 81) \ radians$$

$$EOT = \frac{9.87\sin(2B) - 7.53\cos(B) - 1.5\sin(B)}{60} \ hours \qquad \text{(A20)}$$

$$tod_{noon} = 12 - \left(\frac{longitude}{15} - tz\right) - EOT \ \text{hours}$$

where doy is the day of the year and tz is the time zone (+8 hours for Western Australia),

(3.7) Calculate the declination angle d_s and the solar altitude angle constants K_{sas} and K_{sac}:

$$d_s = 0.40928\sin\left(\frac{2\pi}{365}(284 + doy)\right) \ radians$$

$$K_{sas} = \sin(d_s)\, sinlat \qquad \text{(A21)}$$

$$K_{sac} = \cos(d_s)\, coslat$$

(3.8) Calculate the orbital factor OF:

$$OF = 1 + 0.0344\cos(0.0172142\, doy) \qquad \text{(A22)}$$

(3.9) For surfaces that are not horizontal, and are fixed or have vertical axis sun tracking, calculate the following parameters. If the surface has dual axis sun tracking, it is not necessary to calculate these parameters.

$$cikc = \sin(d_s)\, sinlat$$

$$ciks = \sin(d_s)\, csca \qquad \text{(A23)}$$

$$ciss = \cos(d_s)\, sinazr$$

$$cicc = \cos(d_s)\, coslat$$

If the surface is fixed, the following parameters can be calculated here instead of hourly:

$$cik = cikc\ costilt - ciks\ sintilt$$

$$cis = ciss\ sintilt \qquad \text{(A24)}$$

$$cic = cicc\ costilt + cics\ sintilt$$

The fourth section must be calculated for each hour, h, from 1 to 24.
(4.1) Calculate the hour angle h_s:

$$h_s = \frac{\pi}{12}\left(tod_{noon} - h + 0.5\right) \; radians \tag{A25}$$

(4.2) Calculate the sine of the solar altitude angle, sinα:

$$sin\alpha = K_{sas} + K_{sac}\cos\left(h_s\right) \tag{A26}$$

If sinα is greater than zero, then the sun is above the horizon. Otherwise, set all solar irradiances to zero and go to the next hour.

(4.3) Calculate the hourly average cloudiness c_{hm} using:

$$c_{hm} = c_d\left(1 + K_{h0}\frac{\left(1 - c_d\right)}{\left(1 + K_{h1}c_d^2\right)}\left(sin\alpha + K_{h2}sin^2\alpha + K_{h3}sin^3\alpha\right)\right) \tag{A27}$$

See Appendix B for the values of K_{hi}, for i = 0 to 3.

(4.4) Calculate the hourly average standard deviation using:

$$\sigma = c_d\left(1 - c_d\right)\left(\frac{K_{dv0}}{\left(1 + K_{dv1}c_d\right)} + \frac{K_{dv2}}{\left(1 + K_{dv3}c_d\right)}sin\alpha + \frac{K_{dv4}}{\left(1 + K_{dv5}c_d\right)}sin^2\alpha\right) \tag{A28}$$

See Appendix B for the values of K_{dvi}, for i = 0 to 5.

(4.5) Generate a random number r, with a uniform distribution between 0 and 1.

(4.6) Map r to a translated Weibull distribution using:

$$r_w = \theta_w + \lambda_w\left(-\ln\left(1 - r\right)\right)^{\left(\frac{1}{\kappa_w}\right)} \tag{A29}$$

See Appendix B for the values of θ_w, λ_w and κ_w.

(4.7) Generate a cloudiness residual y_h using:

$$y_h = \varphi y_{h-1} + \sigma_d r_w \tag{A30}$$

y_{h-1} is the residual for the previous hour. If h is the first hour on or after sunrise to be calculated then use $y_{h-1} = 0$.

(4.8) Generate the synthetic hourly cloudiness c using:

$$c = c_{hm} + \sigma y_h \tag{A31}$$

(4.9) Calculate the air mass ratio M, and the clear sky beam and diffuse transmissivities t_{bcs} and t_{dcs}:

$$M = \frac{1.002432 \sin^2\alpha + 0.148386 \sin\alpha + 0.0096467}{\sin^3\alpha + 0.149864 \sin^2\alpha + 0.0102963 \sin\alpha + 0.000303978}$$

$$t_{bcs} = K_{b1} e^{K_{b2} M} \tag{A32}$$

$$t_{dcs} = 0.271 - 0.294\, t_{bcs}$$

See Appendix B for the values of K_{b1} and K_{b2}.

(4.10) Calculate the diffuse transmittance cloudiness coefficient $k_{cloud}(c)$:

$$k_{cloud}(c) = \frac{(1-c)(1+K_{dr}\, c)}{1 - K_{dr}\, c \dfrac{t_{dcs}}{t_{bcs}}} \tag{A33}$$

where K_{dr} is set to 0.5.

(4.11) For horizontal surfaces (assuming diffuse radiation can be utilised), calculate the beam and diffuse transmittances t_b, t_d, the global horizontal transmittance t_{gh}, the horizontal beam and diffuse irradiance components I_{hb} and I_{hd}, and the global horizontal irradiance I_h:

$$t_b=(1-c)\,t_{bcs}$$

$$t_d=k_{cloud}(c)\,t_{dcs}$$

$$t_{gh}=t_b+t_d=(1-c)\,t_{bcs}+k_{cloud}(c)\,t_{dcs}$$

$$I_{hex}=1367\ OF\ \ sin\alpha \tag{A34}$$

$$I_{hb}=t_b\,I_{hex}$$

$$I_{hd}=t_d\,I_{hex}$$

$$I_h=t_{gh}\,I_{hex}$$

(4.12) For surfaces at angles other than horizontal, calculate the beam irradiance I_b. Firstly, calculate the extraterrestrial irradiance I_o:

$$I_o=1367\,OF \tag{A35}$$

For vertical axis sun tracking surfaces, calculate the following parameters:

$$costilt=sin\alpha$$

$$sintilt=\sqrt{(1-sin^2\alpha)}$$

$$cik=cikc\ \ costilt\ -\ ciks\ \ sintilt \tag{A36}$$

$$cis=ciss\ \ sintilt$$

$$cic=cicc\ \ costilt\ +\ cics\ \ sintilt$$

For fixed and vertical axis sun tracking surfaces, calculate the beam irradiance I_b:

$$I_b=I_o\,t_b\left(cik+cis\ \sin(h_s)+cic\ \cos(h_s)\right) \tag{A37}$$

For dual axis sun tracking surfaces, calculate the beam irradiance I_b:

$$I_b=I_o\,t_b \tag{A38}$$

(4.13) For photovoltaic surfaces, calculate the diffuse and reflection components of the irradiance, I_d and I_r. For fixed photovoltaic surfaces:

$$I_d = 0.5\, I_o\, t_d\, sin\alpha\,(1 + costilt)$$

$$t_r = t_b + t_d \qquad \text{(A39)}$$

$$I_r = 0.5\, r_g\, I_o\, t_r\, sin\alpha\,(1 - costilt)$$

where r_g is the ground reflectance coefficient. See Appendix B for the value of r_g. For vertical axis and dual axis sun tracking photovoltaic surfaces, costilt = sinα, so:

$$I_d = 0.5\, I_o\, t_d\, sin\alpha\,(1 + sin\alpha)$$

$$t_r = t_b + t_d \qquad \text{(A40)}$$

$$I_r = 0.5\, r_g\, I_o\, t_r\, sin\alpha\,(1 - sin\alpha)$$

(4.13) The global irradiance falling on the surface can now be calculated. For photovoltaic surfaces,

$$I_g = I_b + I_d + I_r \qquad \text{(A41)}$$

For concentrating mirror surfaces,

$$I_g = I_b \qquad \text{(A42)}$$

Appendix B. Model Coefficients

These 334 coefficients were generated using all the measured data from every station as calibration data:

Table B1. Coastal set points and distance parameters.

j	nkmj	Coastal-setpoints	Distance-Parameters
1	7	(0.000121233,0.181881)(4358.68,0.118325)(5855,0.196657)(6107.18,0.308299) (6382.57,0.361306)(7578.43,0.241622)(7704.56,0.112202)	-4.74725e-05,0.00129837
2	9	(0.00211302,0.149049)(5295.77,0.00929281)(6062.87,0.0743396) (6170.42,0.0544799)(6817.07,0.0385803)(6919.46,0.0462027) (7019.46,0.00995161)(7324.11,0.0279123)(7424.38,0.0251248)	0.0126305,0.00709556
3	6	(3557.6,13.8123)(5548.36,8.12282)(6936.52,6.38246)(7133.08,2.9477) (7263.64,14.263)(7626.37,2.83333)	-4.62672e-15,-9.41605e-16
4	5	(183.127,0.140565)(5548.28,0.163597)(6177.15,0.181829)(7276.71,0.194041) (8058.94,0.15176)	8.20061e-05,0.000200941
5	9	(4.82931e-05,0.0619347)(5548.36,0.0382522)(5648.53,0.0125194) (5951.32,0.0278242)(6414,0.00509126)(6544.79,0.0112567)(6644.79,0.021321) (7238.27,0.00511463)(7939.27,0.0704454)	0.0146506,0.00801621
6	8	(1.03457e-05,13.243)(5948.94,8.23257)(6438.31,12.9464)(7155.63,4.14219) (7263.64,13.999)(7377.2,7.63296)(7718.7,1.97949)(9614.86,2.64167)	8.16237e-14,-1.96663e-14
7	8	(759.862,-1.15271)(4450.09,-0.888514)(5855,-1.04833)(6095.42,-1.43222) (6382.57,-1.70468)(7096.81,-1.37403)(7297.2,-1.49405)(7645.29,-0.922309)	-0.000668503,4.28529e-06
8	9	(2804.47,0.374811)(5655.77,0.0806366)(6339.57,0.353919)(6439.57,0.109719) (6680.68,0.182641)(7096.81,0.0446732)(7634.3,0.0769718)(8069.9,0.414361) (9195.84,0.0299513)	0.0266459,0.00849817
9	8	(278.366,8.43906)(4434.12,8.01394)(6059.16,14.3246)(6897.37,12.0145) (7142.72,19.6262)(7333.4,14.2362)(7783.43,16.4985)(8659.08,9.26487)	4.22435e-16,-4.45298e-16
10	8	(1.67366,1.75907)(2590.43,0.168708)(5782.19,0.90459)(5969.6,1.43466) (6069.64,2.28332)(6433.7,2.91901)(7353.76,2.06326)(7821.81,0.298694)	-0.0004577,0.00125533
11	8	(1.85017e-07,1.34352)(5823.34,0.289987)(5923.36,0.778804) (6049.55,0.886153)(6177.76,1.40718)(6721.28,0.682705)(7101.11,0.220369) (7338.59,0.501468)	0.0162485,0.0083404
12	6	(1.56914e-12,2.53601)(5753.42,2.17944)(5934.42,21.3562)(6177.72,19.5243) (7169.75,18.1674)(7269.75,21.8962)	4.9505e-15,7.50947e-15
13	7	(196.582,0.251853)(4694.57,0.613485)(5782.25,1.08065)(6107.18,0.355968) (6463.88,0.397888)(6565.85,0.252028)(6936.44,0.411005)	0.0133546,0.00437936
14	6	(3586.43,0.14879)(5541.96,1.21734)(5641.96,0.779202)(6107.96,0.592607) (7215.47,0.346493)(7429.46,0.540601)	0.00817635,0.00410551
15	7	(3.78325e-05,16.7796)(3556.04,21.1344)(5927.72,20.0101)(6165.06,13.6784) (6863.41,12.7892)(6963.41,21.9553)(7316.56,26.1913)	-2.25303e-10,-5.96227e-12
16	9	(0.00116518,2.29642)(4383.42,2.99919)(6081.62,1.63957)(6186.12,1.45255) (6294.7,1.08039)(6891.98,1.22188)(6991.98,1.83623)(7232.26,1.11177) (7455.14,2.80992)	0.0010551,0.0001511
17	7	(1.92816e-12,1.21493)(5782.25,0.966054)(6058.36,0.763544) (6315.79,0.174561)(6936.45,0.298013)(7364.85,0.143312)(7732.08,0.690034)	0.0172375,0.00741007
18	9	(1096.22,20.1305)(3940.9,19.6107)(6637.71,10.9754)(6761.42,6.77296) (6897.98,15.706)(6997.98,5.88923)(7396.6,9.65186)(7603.63,8.74399) (8422.38,8.22955)	5.81999e-14,2.03159e-15

Table B2. Model cloudiness generation coefficients

Kh_i, i = 0,3	3.53164,6.58553,-2.65914,1.53216
Kdv_i, i = 0,5	2.9409,3.05052,-4.08909,4.88456,1.14796,1.03685
kyac	0.342923
θ_w	-1.82568
λ_w	2.05741
κ_w	1.89893
K_{b1}	0.85295
K_{b2}	-0.114757

r_g	0.2

Appendix C. Formulation of statistical measures

The root mean square error (RMSE) is calculated using:

$$\mathrm{RMSE}=\frac{1}{N}\sqrt{\sum_{i=1}^{N}\left(h_m-h_d\right)^2}$$

where h_d is the measured data value, h_m is the model generated synthetic data value, and N is the number of data points. RMSE can be represented as a percentage value by dividing by the mean value of h_d and multiplying by 100. The mean bias error (MBE) is calculated using:

$$\mathrm{MBE}=\frac{1}{N}\sum_{1}^{N}\left(h_m-h_d\right)$$

MBE can be represented as a percentage value by dividing by the mean value of h_d and multiplying by 100. The mean absolute percentage error (MAPE) is calculated using

$$\mathrm{MAPE}=\frac{1}{N}\sum_{1}^{N}\frac{\left|h_m-h_d\right|}{h_d}$$

To be represented as a percentage value, the MAPE can be multiplied by 100. The mean relative variance (MRV) is calculated using:

$$\mathrm{MRV}=\frac{\sum_{1}^{N}\left(h_m-h_d\right)^2}{\sum_{1}^{N}\left(h_d-h_{\mathrm{dav}}(h)\right)^2}$$

where $h_{dav}(h)$ is the average of the measured hourly values h_d for hour h.

Appendix D. Nomenclature

α	solar altitude angle (radians)	I_o	extraterrestrial solar irradiance falling on a surface perpendicular to the direction of the sun (W/m^2)
σ	synthetic average hourly cloudiness standard deviation	I_r	reflection irradiance on a tilted surface (W/m^2)
θ_w	translated Weibull distribution location coefficient	K_{b1}	clear sky beam transmittance magnitude constant
λ_w	translated Weibull distribution scale coefficient	K_{b2}	clear sky beam transmittance absorption constant
κ_w	translated Weibull distribution shape coefficient	K_{cd}	coastal inland distance coefficient
σ_d	random component standard deviation	K_{cf}	daily average cloudiness cumulative frequency distribution coefficient
σ_{month}	monthly mean standard deviation of daily average cloudiness	k_{cloud}	diffuse radiation cloudiness transmission factor
ϕ	autocorrelation coefficient constant	K_{cp}	coastal position setpoint coefficient
azimuth	horizontal angle between a line perpendicular to the surface and a line running due north, with angles east of north being positive and west of north being negative (degrees)	k_d	diffuse fraction
c	cloudiness	k_{dcs}	Clear sky diffuse fraction
c_d	daily average cloudiness	K_{dr}	diffuse fraction cloud slope coefficient
c_{davmon}	monthly mean of daily average cloudiness	K_{dv}	hourly cloudiness standard deviation estimation coefficient
cdist	distance inland from the nearest part of the coastline (km)	K_h	average hourly cloudiness estimation coefficient
c_{dsdmon}	monthly standard deviation in mean daily cloudiness	K_m	coastal setpoint constant
cf^{-1}	inverse cumulative frequency distribution function	K_{sac}	solar altitude angle coefficient
c_{hm}	synthetic average hourly cloudiness	K_{sas}	solar altitude angle coefficient
cic	tilted surface sun angle calculation coefficient	k_t	irradiance clearness index
cicc	tilted surface sun angle calculation coefficient	K_t	daily or monthly radiation clearness index
cics	tilted surface sun angle calculation coefficient	K_{yac}	autocorrelation coefficient constant
cik	tilted surface sun angle calculation coefficient	lat	latitude (degrees north of equator)
cikc	tilted surface sun angle calculation coefficient	len	distance along one line segment of coastline (km)
ciks	tilted surface sun angle calculation coefficient	lon	longitude (degrees east of Greenwich)
cis	tilted surface sun angle calculation coefficient	M	optical air mass ratio
ciss	tilted surface sun angle calculation coefficient	MAPE	mean absolute percent error
clca	tilted surface sun angle calculation coefficient	MBE	mean bias error
cline	cumulative distance along coastline (km)	MRV	mean relative variance
cosazr	cosine of the azimuth	N	number of data points
coslat	cosine of the latitude	OF	orbital factor

costilt	cosine of the tilt angle	r	random number with uniform frequency distribution between 0 and 1. Often used as a cumulative frequency value.
cpos	distance along the coastline from the Northern Territory border (km)	r_g	ground reflectance coefficient
d	square of distance along one line segment of coastline map (km^2)	RMSE	root mean square error
d_1	square of distance from a location to beginning vertex of coastline line segment (km^2)	r_w	random number with translated Weibull frequency distribution of mean 0 and standard deviation 1
d_2	square of distance from a location to end vertex of coastline line segment (km^2)	$\sin\alpha$	sine of solar altitude angle
d_m	square of distance along coastline line segment to point nearest a location (km^2)	sinazr	sine of the azimuth
d_{min}	square of current minimum distance from coastline to a location ($degrees^2$)	sinlat	sine of the latitude
doy	day of year. 1 = 1st January, 365 = 31st December (366 in leap years)	sintilt	sine of the tilt angle
d_s	declination angle of the Earth's spin axis with respect to the sun (radians)	S_o	solar constant (1367 W/m^2)
EOT	equation of time (hours)	t_b	beam(or direct) radiation atmospheric transmittance
f	length fraction along coastline line segment to point nearest a location	t_{bcs}	clear sky beam radiation atmospheric transmittance
H	global daily radiation falling a horizontal surface (MJ/m^2)	t_c	cloudiness transmission factor
h_d	measured data value	t_d	diffuse radiation atmospheric transmittance
h_{dav}	average of the measured hourly values of h_d	t_{dcs}	clear sky diffuse radiation atmospheric transmittance
H_{ex}	global daily radiation falling a horizontal surface at the top of the atmosphere (MJ/m^2)	t_{gh}	global horizontal radiation transmittance
h_m	model generated synthetic data value	t_{ghm}	measured global horizontal radiation transmittance
h_s	hour angle (radians)	tilt	vertical angle between a surface and the horizontal plane (degrees)
I_b	beam irradiance (W/m^2)	tod	time of day in local time (hours)
I_d	diffuse irradiance (W/m^2)	tod_{noon}	time of day when solar noon occurs (sun is midway between sunrise and sunset) (hours)
I_h	horizontal global irradiance (W/m^2)	t_r	reflected radiation atmospheric transmittance
I_{hb}	horizontal beam irradiance (W/m^2)	t_{rcs}	clear sky reflected radiation atmospheric transmittance
I_{hcs}	clear sky horizontal global irradiance (W/m^2)	t_z	time zone (+8 hours for Western Australia) (hours)
I_{hd}	horizontal diffuse irradiance (W/m^2)	y_d	daily cloudiness residual
I_{hex}	extraterrestrial solar irradiance onto a plane parallel with a horizontal plane on the surface (W/m^2)	y_h	hourly cloudiness residual
I_{hm}	measured horizontal irradiance (W/m^2)		