The Canadian society for engineering in agricultural, food, and biological systems

CSAE SCGR

La société canadienne de génie agroalimentaire et biologique

Paper No. 05-005

A Solar Radiation Model with a Fourier Transform Approach

Yung-Chien Sun

Bioresource Engineering, McGill University, 21,111 Lakeshore Rd., Ste. Anne de Bellevue, Québec, Canada, H9X 3V9, yungchien.sun@mail.mcgill.ca.

Robert Kok

Bioresource Engineering, McGill University, 21,111 Lakeshore Rd., Ste. Anne de Bellevue, Québec, Canada, H9X 3V9, robert.kok@mcgill.ca.

Written for presentation at the
CSAE/SCGR 2005 Meeting
Winnipeg, Manitoba
June 26 - 29, 2005

Abstract *The objective of this project is to build a solar radiation model which will be incorporated in a larger ecosystem model. The method for calculating solar radiation intensity outside of the Earth's atmosphere is well known, but the incoming flux is attenuated considerably while traveling through the atmosphere. The attenuation depends mostly on atmospheric composition and cloud distribution, which are difficult to predict. The model described here is based on historical data. Measured daily overall radiation (DOR, kJ/m2) data for four Canadian cities (Vancouver, Winnipeg, Montreal, Halifax) were acquired from the Canadian Meteorological Service. Annual DORs were then analyzed and decomposed with a Fourier transform procedure. This resulted in sets of descriptive parameters for the four cities from which new sets of annual DORs can be synthesized. The model was tested to ensure that the generated data were statistically sufficiently similar to the original sets.*

Introduction

The objective of this project is to provide suitable solar radiation predictions to a larger ecosystem model. This ecosystem model is not constructed after any existing ecological environment. Instead, it is based on a computerized virtual world where we can designate the system composition and study the model outcomes under a set of forcing functions (Lanphere and Kok, 2004). Currently, there are two forcing functions, temperature and solar radiation. The solar radiation plays the role of the unique energy source in the ecosystem. In this project, we aim to emulate the actual radiation intensity pattern of four Canadian cities: Vancouver, Winnipeg, Montreal, and Halifax. The outputs are expected to be similar to reality, in a statistical manner, so that appropriate amounts of energy are allocated in the virtual ecosystem.

Extraterrestrial solar radiation can be described as a function of the sun-earth distance, the earth's inclination and the sun's zenith angle. Its value is deterministic for any latitude on earth, and for any given set of year, day and time. Nevertheless, atmospheric constituents and water molecules of clouds reflect, scatter and absorb the solar flux when it is traveling through the atmosphere. This attenuation effect depends mostly on atmospheric composition and cloud distribution, which are very difficult to predict. A sample set of one year's daily overall solar radiation measured in Montreal is shown in Figure 1. The real data display a trend similar to that of the theoretical, un-attenuated line. However, the discount is obvious. Daily fluctuation in real data varies dramatically and randomly. On a cloudy day, the sky could become very dark even in the summer.

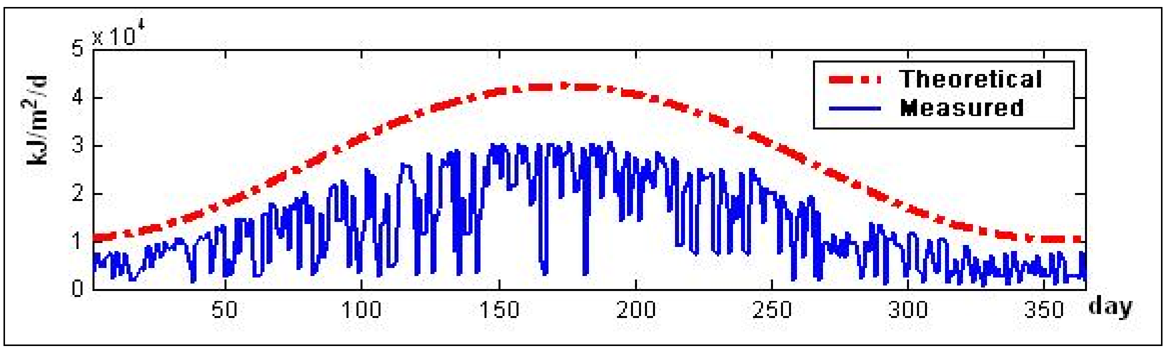

Figure 1 Sample annual daily solar radiation energy, kJ/m^2/d

The proposed model, instead of studying the atmosphere and the clouds, studies the historical on-surface radiation measurements. The variable of interest is the daily overall solar energy, DOR (kJ/m^2/d). One year's DORs are considered as a signal, and one signal is analyzed at a time. Multiple years' signals of one city bear the same statistical characters. The statistical characters are summarized with a set of model descriptors which will be used to synthesize new annual DORs on demand. The synthesized data should be statistically similar to the original set. The output of this model will represent an averaged, typical trend of the history, rather than forecast any future extreme situations.

To fulfill its role in the ecosystem simulation, this model needs to provide unlimited radiation data in an adequate granularity. The reason is that the simulation time period may vary from a few days to several decades, depending on the objective of the study. Hence this radiation model should be able to generate as many DORs as required for the simulation. Moreover, to ensure the similarity between the generated data and the real data, an evaluation series is proposed. First of all, the comparison of model and real data in different time granularities will be examined. The granularities include DOR, weekly average DOR (WDOR), monthly average

DOR (MDOR), and the difference of DOR between two consecutive days (DDOR). Then, the histograms of DOR, WDOR, MDOR and DDOR will be inspected. Finally, statistics of a large number of model data will be checked to ensure their accordance with the real data.

Background

The complex and stochastic nature in modeling solar radiation is well recognized. Most researchers have tackled this problem with a stochastic approach. Deterministic approaches are relatively scarce in the literature.

A deterministic approach considers the atmospheric attenuation as an inevitable consequence of other known causes. Solar Analyst is one such model. Developed by Rich et al. (1994), this model can generate intercepted direct and diffuse radiation from the sky in each direction for a specific geographic context. Among the necessary inputs are the digital elevation models (DEMs), the atmospheric transmittance and the diffuse proportion of incoming solar radiation. The latter two can be acquired by physical measurements or from typical values. Combined, they account for the atmospheric attenuation. Goodin et al. (1999) propose a method to determine solar irradiance, using the daily air temperature as a reference. A daily transmittance coefficient is estimated from the daily temperature range using a Bristow-Campbell (B-C) model. The result is suitable for crop growth studies. Deterministic approaches basically make it possible to utilize the inherent dependency between the surface solar radiation and other prominent factors, such as temperature and atmospheric transmittance. In most cases they are used to estimate missing entries in solar irradiance archives. As the prominent factors are difficult to predict, such models are barely capable of forecasting.

Stochastic approaches study the random components in solar radiation and try to capture the statistical meanings. The statistical characters are then used to generate new time series as practical predictions. Lu et al. (1998) study the clearness index, i.e. the ratio of measured radiation to the clear-sky radiation. In their model, monthly cumulative frequency distributions (CFD) are acquired from historical data. Daily indices can be derived from monthly CFD using a first order autoregressive equation, where the values of day n and n-1 are related with an autocorrelation coefficient and a random term. In the same manner, hourly radiation is calculated by disaggregating the daily radiation value (Graham and Hollands, 1990).

Instead of studying the air transmittance, others examine the on-surface radiation itself. In this sub-category, time series analysis is usually the approach of choice to analyze solar radiation records. Original data series are decomposed as identifiable seasonal trends of daily mean and daily variance, along with a dynamic, unpredictable, random part, usually called noise. Fourier transform is found to be broadly used in simulating recognized trends by fitting them with low frequency sinusoids. The principal granularity in solar radiation time series study resides in the daily or hourly value, which is broken down from the daily sum. Richardson (1981) proposes a climate model in which the daily maximum and minimum temperature and solar radiation are dependent on the amount of precipitation. To generate these weather variables, the precipitation is first generated with a Markov chain-exponential model. The residual series of the other three elements are then re-produced with a multivariate generation method. The output values of temperature and radiation are obtained by first multiplying the residual by the variance and then adding a seasonal mean conditioned by the wet or dry status of the day. Boileau (1983) also uses Markov process to produce the residual series of solar radiation. During data analysis, he notes that for certain sites, the fluctuations are not symmetrically distributed about the daily mean. This asymmetry could not be accommodated with standard procedures which assume a Gaussian probability density. Hence a preliminary transformation is necessary to turn the noise into a more symmetric state. Brinkworth (1977) divides a year into six equal intervals. He finds that, for each interval, the solar radiation segment can be modeled by a linear trend and a daily variance. The variance is thus symmetric about the trend and is suitably described by using an autocorrelation process with a displacement of one day. Time series analysis is also applied on similar natural phenomena. Parrott et al. (1996) present a daily average temperature (DAT) model using a Fourier transform approach. For each annual DAT signal, they identify three principal sinusoids (0, 1, 2 yr^{-1} frequency, trend of the seasonal mean) and one beat sinusoid, which is equivalent to the trend of seasonal variance. The remaining signal is identified as random noise and simulated by 180 higher frequency sinusoids. Model descriptors calculated from the parameter of these sinusoidal curves are used to synthesize new sets of annual DAT.

The Model

This solar radiation model was constructed based on Parrot's temperature model (1996). Fifteen years' worth of physical data was acquired for each city and was divided into annual DOR series. Individual annual DORs were deemed independent signals and were processed one by one. No inter-year correlation was examined. Each original signal was decomposed into seasonal trend of mean and variance as well as noise. These three parts were then modeled with various sinusoids. During data analysis, we found that this approach was highly dependent on the time series in question. To deal with solar radiation, we had to include more steps in order to accommodate the asymmetry of noise. The data analysis and model test were completed in MATLAB

Data analysis

Historical solar radiation records for Vancouver, Winnipeg, Montreal and Halifax were acquired from the Canadian Meteorology Service. The data were recorded as hourly overall energy, in kJ/m^2. For each day, hourly values were summed up to obtain the DOR ($kJ/m^2/d$). Identical analysis procedure was performed for each of the four Canadian cities, resulting in one individual model for each city depicting the statistical characteristics of the local solar radiation pattern. At first, the least square method was applied to identify the three principal sinusoids in the signal. The three principal sinusoids had low frequencies (0, 1, 2 yr^{-1}) and big magnitudes; they represented the seasonal mean trend in annual DORs. This procedure followed the order of 1, 2 and 0 yr^{-1} frequency sinusoids; they were calculated from the signal and were subtracted from it in turn. The resulting time series had a mean of zero and displayed a sinusoidal magnitude outline. To standardize the variance, a beat sinusoid describing the absolute magnitude was identified from the resulting time series, which was then divided by it (Parrott et al. 1996). Again, the beat sinusoid was calculated with the least square method. This gave us a new data series which was recognized as random noise. The procedure up to this point is illustrated in Figure 2.

We intended to decompose the noise signal further with higher frequency sinusoids at this stage. However, before proceeding, an asymmetry was observed in the signal: there were more positive noise values, while the negative noise values came with bigger amplitude in average (Figure 3). The asymmetry indicated that applying a Fourier transform directly was inappropriate. The Fourier transform would further break down the noise into a spectrum of sinusoids. Being symmetric in nature, these sinusoids can hardly retain this asymmetric feature.

A step we called symmetry transformation was introduced here. This transformation was similar to the one that Richardson (1981) suggested to remove periodic means and standard deviations. Basically, positive and negative noises were treated separately. As shown in equations (2) and (3), the daily mean and standard deviation of positive and negative noises were calculated from the 15-year real data. The transformation turned the noise into its residual, for $i = 1 \sim 365$,

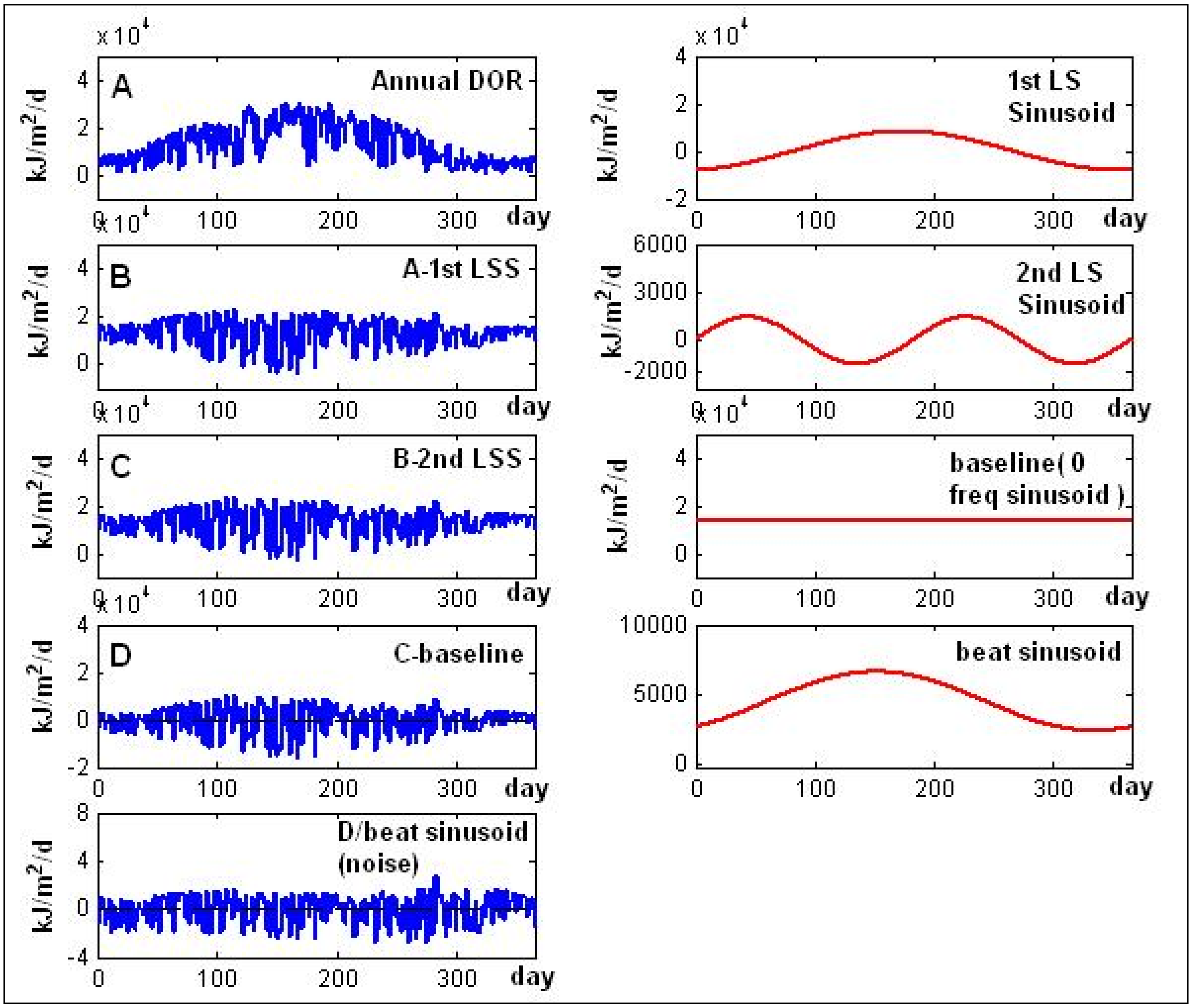

Figure 2 Data analysis: Annual DOR signal, three principal sinusoids, beat sinusoid and noise

$$residual_i = \frac{noise_i - \overline{X_i^p}}{\sigma_i^p} \qquad noise_i \geq 0 \qquad (2)$$

$$= \frac{noise_i - \overline{X_i^n}}{\sigma_i^n} \qquad noise_i < 0 \qquad (3)$$

where $noise_i$ refers to the noise of day *i*; $\overline{X_i^p}$ and σ_i^p represent the mean and STD of positive noises for day *i*, while $\overline{X_i^n}$, σ_i^n stand for negative noises. Because the positive and negative noises have different counts of occurrence, we also calculated the averaged percentage of positive noise occurrence for the 15 years, noted as *PNOc*.

The residual was more symmetric than the noise in terms of closer counts of positive and negative points and smaller gap in their average amplitudes (Figure 3). Fast Fourier Transform (MATLAB function FFT) was then applied to the residual, which was found to consist of 183 sinusoids with frequency from 0 to 182 yr^{-1}.

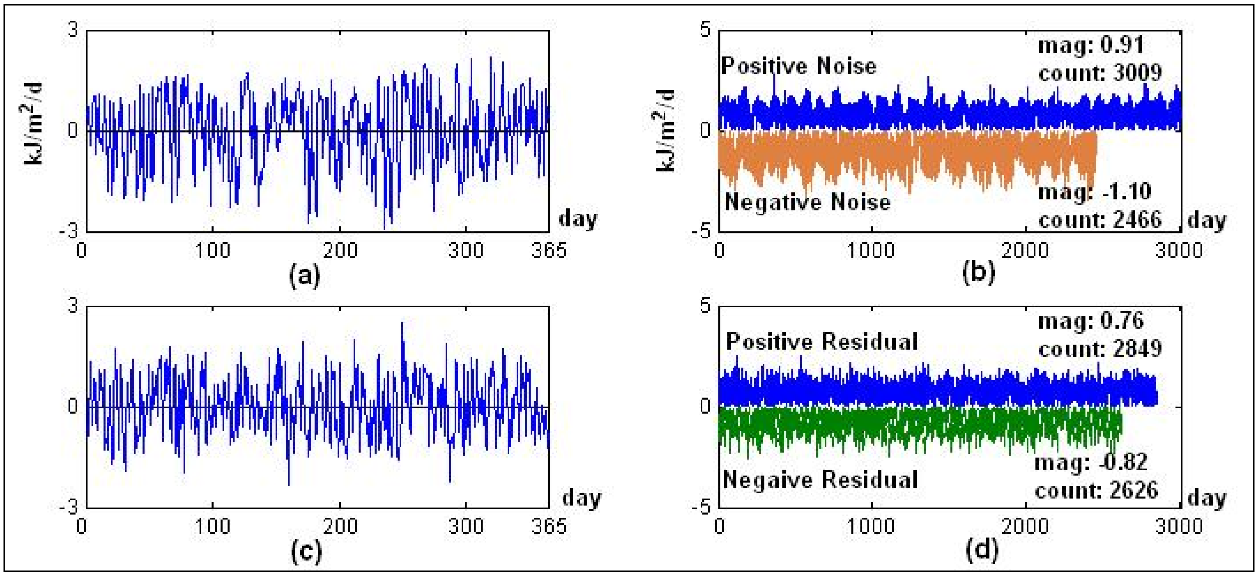

Figure 3 Noise asymmetry illustrations. (a) Sample one year noise signal. (b) Dissection of 15-year noise into positive and negative parts. (c) One year residual after transformation. (d) Dissection of 15-year residual into positive and negative residuals.

$$Radiation = a_0 + a_1 \sin(\tau+a_2) + a_3 \sin(2\tau + a_4) \quad \text{Principal sinusoids}$$

$$+ [b_0+b_1 \sin(\tau+b_2)] \times Noise \quad \text{Beat Sinusoid} \times \text{Noise}$$

$$Noise = T^{-1}_{sym}(Residual, PNOc, \overline{X^p}, \sigma^p, \overline{X^n}, \sigma^n)$$

$$= T^{-1}_{sym} [\sum_{0}^{182} c_{1i} \sin(\tau \times i+c_{2i}) , PNOc, \overline{X^p}, \sigma^p, \overline{X^n}, \sigma^n]$$

Figure 4 The solar radiation model

The data analysis for one year was completed as discussed. In summary, the annual DOR signal was modeled as 187 sinusoids (three principal sinusoids, one beat sinusoid and 183 for the residual) plus the asymmetry property. Figure 4 gives a mathematical expression of this model where T^{-1}_{sym} denotes the reverse operation of the symmetry transformation.

Model descriptors

The identical procedure was repeated for each of the remaining 14 years. As a result, the nine parameters which specified the principal and beat sinusoids each had a sample size of 15: one from each year. They were tested to have a normal distribution at a significance level of 0.1. Their mean and standard deviation were calculated and became the first 18 elements of the model's descriptor set.

As for the 183 sinusoids which form the residual, the distribution of their magnitudes was tested over the 183 frequencies. 167 out of 183 (91%) were found to bear a normal distribution at a significance level of 0.1. The three principal sinusoids and the beat sinusoids were modeled individually in complete mathematical form. The residual, on the other hand, was modeled by how its magnitude and phase angle change with frequency. For the magnitude, the mean and STD of each of the 183 sinusoids were calculated. Their movement over frequency 0 to 182 yr^{-1} is fitted with 5^{th} order, least square polynomials. The phase angle was assumed to be uniformly distributed between π and $-\pi$. The asymmetric feature introduced the mean and STD of daily positive and negative noises. Likewise, their mean and STD over 365 days were fitted with 5^{th} order, least square polynomial lines. Each 5^{th} order polynomial requires six coefficients. They totaled 36 coefficients and were also integrated into the model descriptor set (Figure 5).

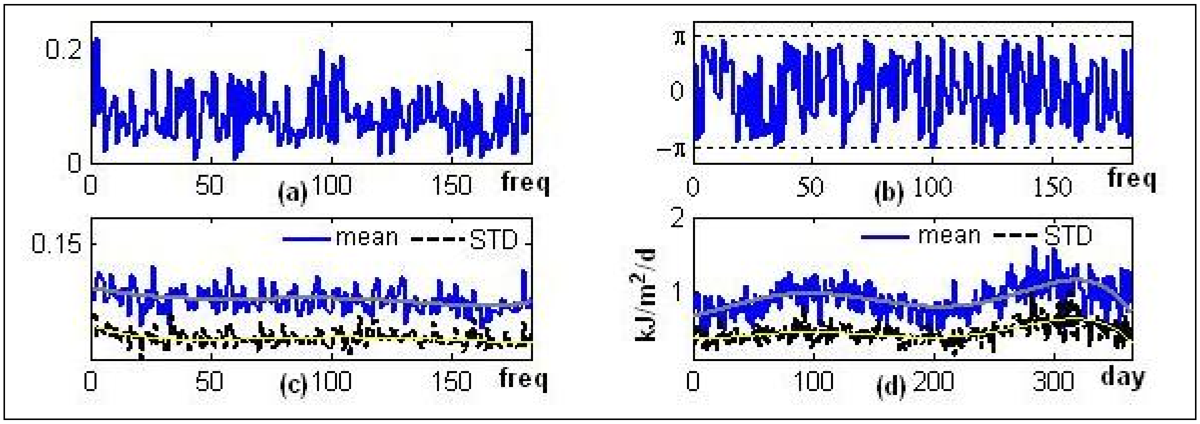

Figure 5 For Montreal, (a) sample magnitude spectrum for one year's residual. (b) Sample phase angle distribution for one year's residual. (c) Mean and STD of 15 years' worth of magnitude over frequency 0 ~ 182, and their 5^{th} order polynomial fittings. (d) Mean and STD of positive noises and their 5^{th} order polynomial fittings.

Among the nine parameters, a correlation study found generally strong correlation between two sets of parameters for all four cities: (a_0 , a_1) and (b_0 , b_1). Their correlation levels were included into the model in order to include the inter-variable dependency. We had to include the minimum DOR in real data. This number will be used later as a hard constraint during new DOR generation. Any generated DOR falling below this limit will be truncated. The maximum limit was not taken because the chance of exceeding it was minor. The final number of descriptors is 57. This set of descriptors represents the radiation model of interest. With this descriptor set, we are able to generate new annual DORs.

New Annual DORs synthesis

The model is capable of generating one annual DOR signal at a time. The procedure to generate new annual DORs follows a reverse order of the data analysis procedure. At first, random parameters were generated with a general form as

$$VALUE_f = MEAN_f + STD_f \times \min([\max(N(0,1), BOUND_1)], BOUND_2) \qquad (4)$$

where N(0,1) is a random number normally distributed with mean of zero and standard deviation of 1. A pair of lower and upper bounds were introduced to prevent eccentric values. For this model, (-2.0, 2.0) were chosen as the bounds. After the random parameters were created, we started with the calculation of the residual. For each frequency (0 to 182 yr^{-1}), the associated magnitude was calculated by obtaining the associated mean and STD from the polynomials, then by following the equation (4). Phase angles were generated from a uniform distribution between (π, $-\pi$). The resulting 183 time-domain sinusoids were then added up to form the residual. The next step involved rebuilding the noise from the residual. The asymmetric feature was reproduced at this stage. A new *PNOc* was first created. Next, by assuming that the chance of positive noise for a specific day is a uniformly distributed random variable whose value falls between (0, 1), a time series of daily positive chance (U (0, 1)) could be generated with a random number generator. For each day, noise was assigned as positive if its chance of being positive was greater than the new *PNOc*, and negative otherwise. Following equations (2) or (3), we could compute the noise from the residual conditioned by the sign. The values of mean and STD of positive and negative noises were retrieved from the 5^{th} order polynomial fittings. The rest of the procedure is identical to that of Parrott's DAT model and will not be repeated here.

Test Results

The goal is to generate annual DORs which are statistically similar to the physical data. To test how this goal was met, we conducted a series of tests on the following items:

1. DOR
2. Weekly average DOR (WDOR) and monthly average DOR (MDOR)
3. Difference between DOR for two consecutive days (DDOR)
4. Histograms of DOR, WDOR, MDOR and DDOR
5. Statistics from a large number of model generated data

Figure 6 shows typical annual DOR signals of real data and model generated data. Figure 7 compares the outputs of 15, one-year model runs with real data in three different granularities (daily, weekly and monthly) and the day-to-day differences. Diagrams in the left column compare the radiation value in kJ/m^2/d. Those on the right give the corresponding histograms indicating the relative abundance.

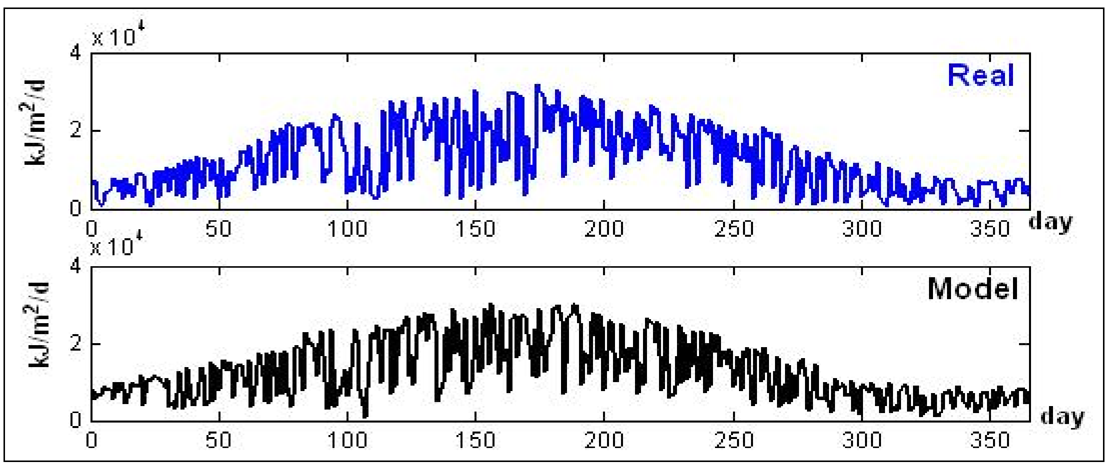

Figure 6 Typical annual DOR from real data and model

The last stage of tests involved 100 sets of 15, one-year model runs. The results should represent a big enough sample of the model outputs to show a good statistical accordance to the original set. The 100 sets of 15 annual DORs were fed back to the data analysis procedure and resulted in 100 sets of model parameters. We compared the mean and STD of the nine major parameters (a_0 ~ *PNOc*). Results for Halifax, Montreal and Vancouver are listed in Table 1. Differences for some parameters were observed. Particularly big discrepancies lie in a_2's STD, a_4 and *PNOc* for most cities (ratios refer to model statistics over real statistics). Nonetheless, their values are not significant compared to other statistics, and differences here should not compromise the quality of the model. The possible cause for the difference is that their distributions are not perfectly normal. This result is also acceptable considering the small sample size of 15 data points per parameter.

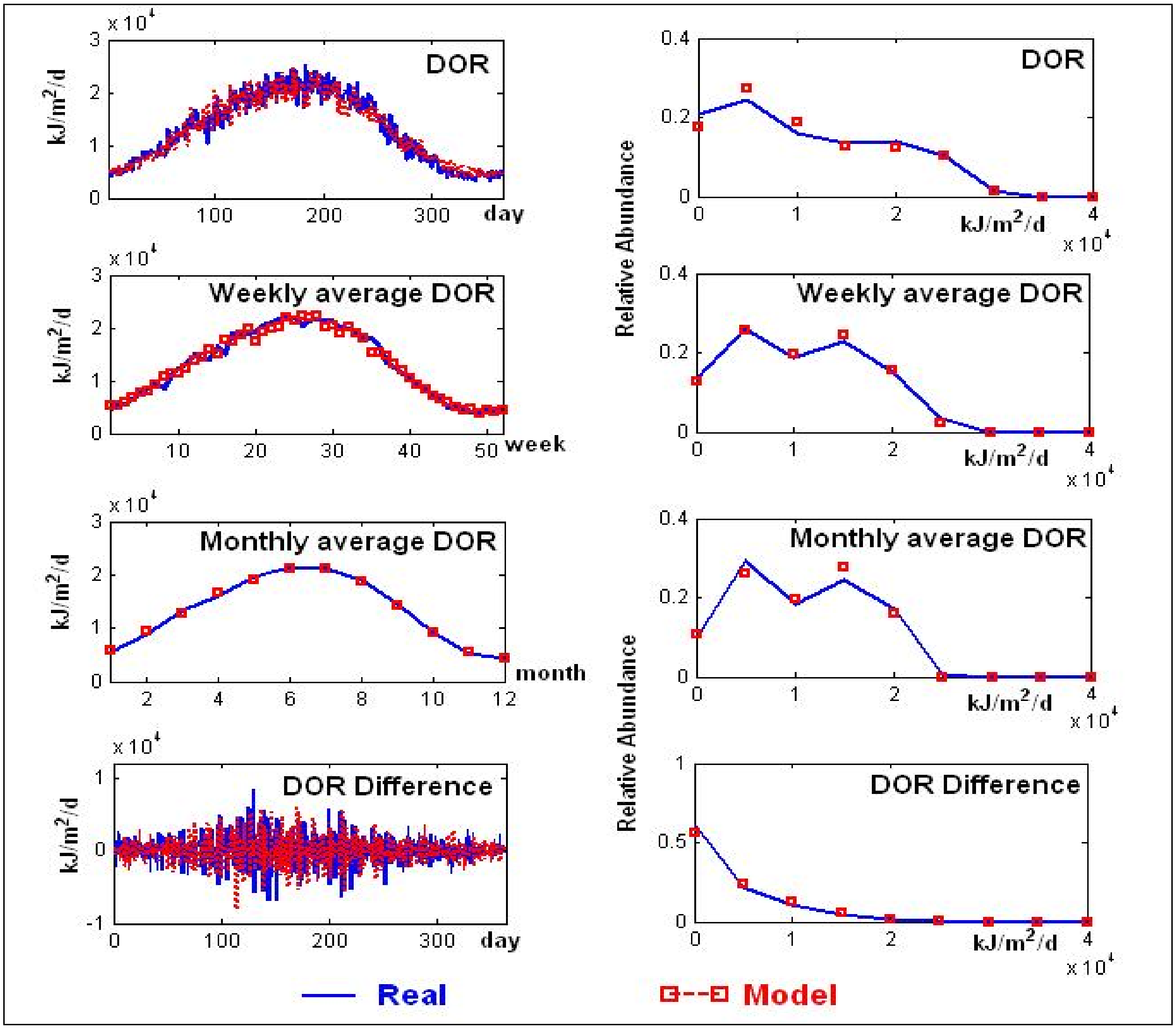

Figure 7 Model tests of 15, one-year runs, for different granularities

Conclusion and Discussion

We proposed a mathematical model for simulating annual daily overall radiation. The model contains one set of descriptors for each of the four Canadian cities. The goal is to generate reasonable solar radiation values based on model descriptors for future use in an ecosystem simulation context. Test results showed the goal was well fulfilled. It is possible to create as many annual DORs as required which have similar statistical characteristics to the original data set.

We intended to build up a parsimonious model, which needs the least number of model descriptors delivering adequate information. This model needs 57 descriptors, which is 28 more than Parrott's temperature model. Twenty-six descriptors are attributed to the asymmetric nature in annual DOR, which appears in all four cities in our study. It is necessary to address this problem by looking at the difference of occurrence and magnitude between positive and negative noises. The additional 26 descriptors stand for the needed asymmetric property to complete the model, which otherwise would produce unrealistic data. For the 15 annual DORs

Table 1 Comparison of parameters obtained from real data and 100 sets of 15, one-year model runs

City		Halifax			Montreal			Vancouver		
Parameter		Data	Model	ratio	Data	Model	ratio	Data	Model	ratio
A0	mean	12262.31	12159.87	0.99	13120.28	13047.04	0.99	12066.49	11872.91	0.98
	STD	527.47	514.79	0.98	456.60	438.18	0.96	469.95	453.44	0.96
A1	mean	8002.97	7657.65	0.96	8829.81	8499.09	0.96	10629.61	9945.18	0.94
	STD	577.38	687.52	1.19	632.28	707.11	1.12	630.47	690.37	1.10
A2	mean	-1.44	-1.42	0.98	-1.35	-1.34	0.99	-1.43	-1.40	0.98
	STD	0.05	0.08	1.54	0.06	0.07	1.21	0.08	0.09	1.06
A3	mean	866.24	912.55	1.05	1037.23	1001.96	0.97	792.46	858.22	1.08
	STD	411.36	397.53	0.97	382.06	380.26	1.00	375.53	367.09	0.98
A4	mean	-0.12	-0.33	2.77	-0.06	-0.12	1.92	0.68	0.75	1.11
	STD	0.53	0.94	1.78	0.85	1.00	1.17	1.00	1.17	1.17
B0	mean	4940.00	4930.79	1.00	4447.53	4486.40	1.01	4107.73	4190.50	1.02
	STD	216.70	226.55	1.05	185.07	211.92	1.15	237.23	254.52	1.07
B1	mean	2705.70	2752.21	1.02	2295.44	2410.72	1.05	2507.83	2632.97	1.05
	STD	184.80	232.57	1.26	208.39	241.37	1.16	374.01	314.12	0.84
B2	mean	-1.25	0.00	0.00	-1.14	0.00	0.00	-1.26	0.00	0.00
	STD	0.10	0.00	0.00	0.10	0.00	0.00	0.09	0.00	0.00
PNOc	mean	0.53	0.53	1.01	0.55	0.55	1.01	0.53	0.53	1.00
	STD	0.02	0.02	1.19	0.01	0.02	1.50	0.02	0.02	1.37

being described with 57 descriptors, the compression rate is about 1% (365×15/57=0.0104).

The output of daily overall energy is useful for ecosystem simulations. From DOR, we could also derive solar intensity at a finer interval, such as every 10 minutes. With interpolation, solar intensity for any time of day becomes available through simple program routines. This is actually what this solar radiation model does in the ecosystem simulation stated at the outset. The result has been satisfactory in terms of the values generated and the speed of model execution.

References

Boileau E. 1983. Use of some simple statistical models in solar meteorology. *Solar Energy*, 30(4): 333-339

Brinkworth, B. J. 1977. Autocorrelation and stochastic modeling of sequences. *Solar Energy* 19:343-347

Fu P., P. M. Rich. 2002. A geometric solar radiation model with applications in agriculture and forestry. *Computers and Electronics in Agriculture* 37: 25-35

Goodin D. G., J.M.S. Hutchinson, R. L. Vanderlip, M. C. Knapp. 1999. Estinating solar irradiance for crop modeling using daily air temperature data. *Agronomy Journal* 91:845-851

Graham V. A., K. G. T. Hollands. 1990. A method to generate synthetic hourly solar radiation globally. *Solar Energy*, 44(6): 333-341

Lanphere Tania R., R. Kok. 2004. A generic ecosystem model for developing ecosystem engineering theory. *2004 ASAE/CSAE annual international meeting, Ottawa, Canada. Aug. 1-4*

Lu Z., R. H. Piedrahita, C. D. S. Neto. 1998. Generation of daily and hourly solar radiation values for modeling water quality in aquaculture ponds. *Trans. ASAE* 41(6): 1853-1859

Parrott L., R. Kok , R. Lacroix. 1996. Daily average temperatures: modeling and generation with a Fourier Transform approach. *Trans. ASAE* 39(5): 1911-1922

Richardson C. W. 1981. Stochastic simulation of daily precipitation, temperature and solar radiation. *Water Resources Research*, 17(1): 182-190