Comparison of Direct Radiation Split Algorithms for Energy Simulation of Buildings

Giorgio Lupato – University of Trieste – s232454@stud.units.it
Marco Manzan – University of Trieste – manzan@units.it
Stefano Cirilli – University of Trieste – cirilli@units.it

Abstract

Direct normal radiation (DNI) has great importance for both energy building simulations and solar energy systems. The data is seldom available from measurements but usually is recovered from global radiation data using split algorithms. The present paper analyses the performance of 33 different split radiation models and the error which arises when applied to building energy simulations using generated hourly weather files.

The split models have been applied to an observed dataset composed by 525888 points, which comprises global and diffuse radiation on the horizontal plane, related to -year measurements, starting from 2001 with 10-minute time steps.

The generated weather files have been employed as input for energy simulations with EnergyPlus on a building generated using DesignBuilder software. We investigated the impact of the weather files in building energy simulation highlighting the performances of four models selected among the 33 models by means of statistical indicators, during different periods of the dataset, since its amplitude allowed us to decompose and analyse 10 different years.

1. Introduction

Dynamic building energy simulations, usually carried out at least on an hourly basis, require detailed environmental data such as temperature, humidity, wind velocity and direction usually available from a number of climatic stations. However, in order to compute solar loads, direct normal radiation and diffuse horizontal radiation are also required. Unfortunately, continuous records of DNI are scarce due to the cost of the equipment: the monitoring stations equipped with solar trackers are very rare. An intermediate solution is to record diffuse and global irradiance, but global insolation is the unique parameter monitored in numerous locations around the world, therefore a great number of climatic data report only this value.

Starting from the work by Liu and Jordan (1960), many efforts were undertaken to develop separation models to estimate the diffuse horizontal irradiation component and, by subtraction, the direct horizontal component. Thereafter, the direct radiation is obtained by dividing it by the cosine of the zenith angle, properly averaged on the interval monitored.

In literature, more than 150 models have been developed with numerous comparison papers. Nevertheless, the definition of a universal model able to provide the best possible result at any specific location is still very complicated, because algorithms are usually expressed in terms of first or fourth degree polynomial functions, empirically derived from site-specific measurements. This technique usually tends to suffer for excessive model localization and/or overfitting which implies that one model can lead to accurate results for a location, while performing poorly for a different one.

This paper describes the performance of different split algorithms using a dataset containing global and diffuse irradiance measurements recorded in Trieste. The availability of a large number of climatic data, which spans for a period of ten years, allows for a deep comparison among split methods. Therefore, the former part of this paper is focused on the statistical analysis of the models while the latter explores the effect of the different split methods on building energy simulations.

Part of
Pernigotto, G., Patuzzi, F., Prada, A., Corrado, V., & Gasparella, A. (Eds.). 2018. Building simulation applications BSA 2017. bu,press.
https://doi.org/10.13124/9788860461360

2. Dataset

The data used is composed of 10-year records of Trieste (45°,65°,13.76°) collected by the Meteorology and Oceanography Laboratory of the University of Trieste containing global and diffuse horizontal irradiation measurements with 10-minute interval detection. The total number of available points is 525888 and includes the following exogenous measurements: dry bulb temperature, relative humidity, wind speed and wind direction too.

Solar position has been considered in the middle of the measurement interval, shifting the time detection back 5 minutes for all the datasets, since the row data were originally recorded as mean solar measurement in Wh/m^2 reported at the end of the interval.

The global horizontal irradiation has been used as the input data for the application of split methods, while the diffuse, and hence the direct horizontal data, have been used as the reference value.

3. Model Selection

Since the first split model proposed by Liu and Jordan in 1960, many models were developed in literature. According to Lannini (2010), three different types of models can be considered: polynomial models, exponential models, and logistic models. All of these categories use predictors, intended as a measurement, or an evaluated variable, required for applying the model. In all the proposed models of this paper clearness index *kt*, defined in Equation 1, is used as a predictor in order to obtain the diffuse fraction *kd*, defined in Equation 2.

$$kt = \frac{GLO}{E_{0h}} \quad (1)$$

$$kd = \frac{DIF}{GLO} \quad (2)$$

The extraterrestrial solar radiation was calculated with Spencer Fourier series expansion.

Other predictors can be used as well. They can be grouped in *kt* class predictors and exogenous predictors as dry bulb temperature, dew point temperature or relative humidity. Table 1 presents the list of the models chosen and analyzed in this article.

4. Quality Control

The posteriori quality control of the measured data was followed as described in Gueymard et al. (2016) and summarized in Table 2. With the application of this quality check, the number of valid points was reduced from the original 525888 to 239594, taking into consideration night hours, too.

Table 1 – Models analyzed with total number of required predictors and number of exogenous predictors in brackets

Id	Model	# predictors
1	Orgill and Hollans	1
2	Reindl1	1
3	Reindl2	2
4	Hawlader	1
5	De Miguel	1
6	Karatasou	1
7	Erbs	1
8	Chandrasekaran	1
9	Oliveira	1
10	Soares	1
11	Lam Li	1
12	Furlan 1	1
13	Lee	1
14	Maxwell	2
15	Macagnan	2
16	Boland 2001	1
17	Louche	2
18	Spencer	1
19	Jacovides	1
20	Boland 2008	1
21	Reindl3	4(2)
22	Perez	4(1)
23	Ulgen Hepbash	1
24	Ruiz-Arias	2
25	Chikh	1
26	Engerer2+Bird	5
27	Paulescu and Blaga 1	1
28	Paulescu and Blaga 3	1
29	Paulescu and Blaga 4	4(2)
30	Paulescu and Blaga 5	2
31	Elminir	1
32	Al Riahi	1
33	Torres	1

The PSA algorithm (Blanco-Muriel et al., 2001) was used to calculate the position of the sun.

Finally, the maximum allowable value of clearness index *kt* has been forced to 1, but this condition had to be enforced 3 times only.

Table 2 – Quality checks applied and number of eliminated points (Npe) for each rule

Id	Limit	Npe
a	Z<85°	284736
b	GLO>0 and DIF>0 and DIR≥0	267586
c	DIF<0.95*E0*cos^1.2*Z+50	95
d	GLO<1.5*E0*cos^1.2*Z+100	71
e	DIF/GLO<1.05 and Z<75°	934
f	DIF/GLO<1.10 and Z>75°	1120

An additional correction has been adopted by limiting minimum and maximum values of the estimated diffuse fraction to values of 0 and 1 respectively in order to prevent unphysical results. This quality check avoids negative diffuse irradiation or diffuse irradiation greater than the global irradiation. It is worth noticing that the diffuse radiation lower limit of 0 represents an unreal value too, since even with the clearest sky condition, the diffuse fraction should be present, too.

5. Statistical Indicators

Three statistical errors were used: mean bias deviation, mean absolute deviation, and root mean square deviation, defined in Equations 3–5.

$$MBD = \frac{\sum_1^n (DIFe_i - DIFm_i)}{n} \qquad (3)$$

$$MAD = \frac{\sum_1^n | DIFe_i - DIFm_i |}{n} \qquad (4)$$

$$RMSD = \sqrt{\frac{\sum_1^n (DIFe_i - DIFm_i)^2}{n}} \qquad (5)$$

Subscript *e* indicates the estimated value, subscript *m* the measured values, and *n* is the number of valid points after the quality check.

The purpose of this paper is to investigate the error that occurs in building energy simulations because of the split model selection in defining climatic datasets. The statistical analysis could help the designer in making a conscious choice between the proposed models. In this paper, four models have been selected for a subsequent energy simulation analysis.

6. Results: Statistical Indicators

Tables 3 to 5 present for each statistical indicator the results obtained for the top five performing models considering the whole dataset and separate heating and cooling periods.

Table 3 – Models with lowest mean bias deviation

	MBD [Wh m^{-2}]		
Model	year	Heating	Cooling
Oliveira	0.2	4.23	8.86
Maxwell	2.6	-3.37	26.25
Ulgen Hepbasli	3.6	-3.04	12.78
Torres	4.1	6.38	15.30
Perez	4.5	-1.37	8.35

Table 4 – Models with lowest mean absolute deviation

Model	MAD [Wh m^{-2}]		
	year	Heating	Cooling
Perez	26.0	16.12	32.59
Louche	27.7	14.94	35.99
Spencer	27.8	15.23	35.61
Soares	28.3	14.54	35.55
Ruiz-Arias	28.6	14.06	37.25

Table 5 – Models with lowest root mean square deviation

Model	RMSD [Wh m^{-2}]		
	year	Heating	Cooling
Perez	45.6	30.51	53.33
Oliveira	46.6	24.81	54.54
Erbs	46.6	26.94	56.66
Torres	46.8	26.71	56.17
De Miguel	47.3	27.71	58.57

From the analysis of the results, the Perez model shows remarkable results in the cooling period, but the performance degrades for heating. Since the number of day hours is greater in the cooling period, the Perez model performs better than the other models on an annual basis. Taking into account the full results, of which Tables 3, 4, and 5 are just a summary, the Perez, Oliveira and Torres models were chosen to carry out building energy simulations.

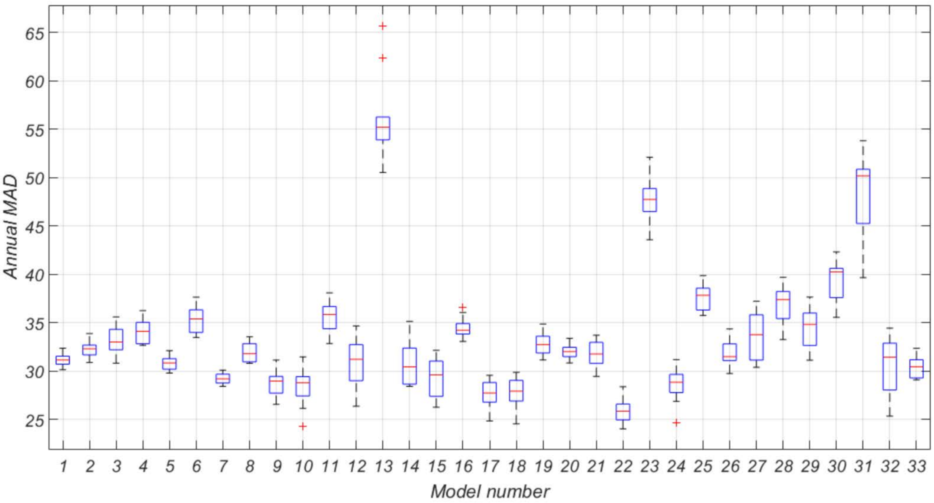

Fig. 1 – Distribution of MAD computed for every year of the dataset and each model

It is interesting to consider that the Oliveira and Torres models were obtained by linear regression applied to datasets collected in San Paulo, Brazil, and Pamplona, Spain, respectively. These models perform well also with the present dataset of Trieste because those locations have similar solar irradiation.

Fig. 1 presents the performance of each model by applying the statistical indicator for each year of the dataset. Therefore for each model a distribution of ten values is obtained. To compare the performance of the complete set of models, the results are reported as box plots, where the box extends from the first to the third quartile, the upper and lower whisker represent the minimum and maximum values of the set, while the line drawn into the box represents the median of the distribution. In this way an idea of the performance distribution can be obtained at a glance.

The inspection of Fig. 1 shows which model performs better compared to the others, not only in an absolute manner, but also considering the dispersion of the results: one model can perform well for one year but can give unsatisfactory results for other situations with large dispersed values.

Due to this consideration, we focused on the Al Riahi model. As can be seen in Fig. 1, the Torres model (model n. 33) and the Al Riahi model (model n. 32) have comparable MAD over the entire dataset (red line within the box). However, we can also notice that the amplitude of Al Riahi box is remarkably greater than the Torres box. The Al Riahi model has therefore been added to the selected models: the goal is to observe how the performance discrepancy between different years influences the building energy simulation. In detail, the Al Riahi model has its lowest MAD for 2003 (MAD 25.4), and its highest for 2004 (MAD 34.5). This can be explained through the inspection of Fig. 2: the year 2003 is catheterized by high mean global and direct irradiation, with a low diffuse fraction; on the contrary, the year 2004 shows low global and direct solar radiation paired with high diffuse fraction. The last condition is different from Trieste's solar irradiation and from the original Al Riahi dataset, monitored in Baghdad. For the same reasons also the Torres model, collected in Pamplona, Spain, has its lowest MAD in 2003, while the highest is recorded in 2004.

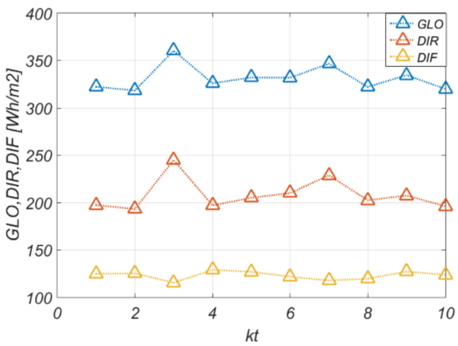

Fig. 2 – The mean yearly values of irradiance: Global, Direct and Diffuse on the horizontal plane

7. Model Choice

Following the observation explained in the previous paragraph, we selected four split models: Oliveira, Perez, Al Riahi, and Torres models. According to Lannini et al. (2010), except the Perez model that is classified as an exponential model, the three others are polynomial models. Fig. 3 presents the dispersion of diffuse fraction over the clearness index for the Perez model, while Fig. 4 compares the same dataset with the Oliveira, Al Rhiai, and Torres models.

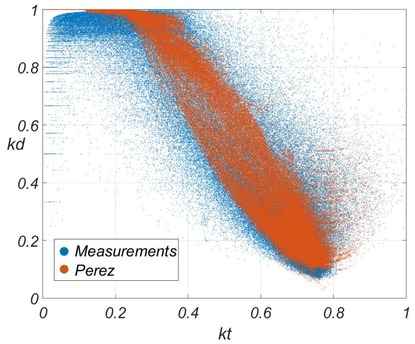

Fig. 3 – The Perez model

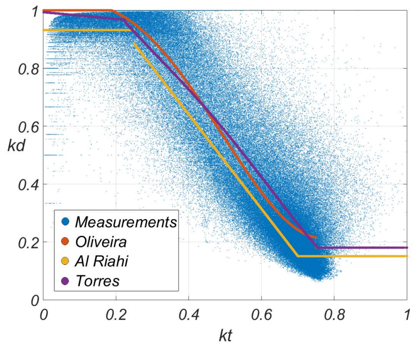

Fig. 4 – The Oliveira, Al Riahi, and Torres models

8. Building Envelope

The selected models were to generate climatic data in order to check the effect on building simulations. According to Pernigotto et al. (2016), a simple test building was simulated. It consists of a prism shape building with a square base, oriented to the main cardinal directions. Its internal floor area is 100 m² with 3 m internal height. All the opaque constructions are composed by two-layer structures with external isolation and 30 cm of concrete internal structure. Thermal bridges are neglected. The material characteristics are reported in Table 6. The insulation thickness, reported in Table 7, was set in order to obtain the reference thermal transmittance, according to the Italian regulations for climatic zone E. Windows have thermal transmittance U_w = 1.8 W m^{-2} K^{-1} with g = 0.35.

Table 6 – Layer material thermal proprieties

Layer	λ	Cp	ρ
Concrete	0.13	1.88	399
Insulation	0.04	1.47	40

Table 7 – Thermal transmittance and insulation layer thickness

Construction	U	s [cm]
External wall	0.30	9.41
External roof	0.25	12.2
Ground floor	0.30	9.25

Wall solar absorbance was set to 0.3, except for the internal floor and the external roof where a value of 0.6 was set. Ground temperature at building surface and at ground deep temperature was considered as constant, respectively of 18 °C and 14 °C.

In summary, the test building is composed of a well-insulated envelope with external windows on the east, south, and west facade, each one of 9.707 m² (29.12 m² of total area). This last choice was made to analyze a building with a wide window size that, as highlighted by Pernigotto et al. (2016), is critical for possible high cooling loads.

Internal gains are constant equal to 4 W/m² with radiant and convective fraction of 0.5. Ventilation air change was set at 0.3 vol/h. Heating and cooling systems were simulated as an ideal system, 100 % convective and with a unitary coefficient of performance. Heating and cooling systems are available according to Table 8.

Table 8 – Beginning, end and setpoint temperature related to heating and cooling systems

System	Begin	End	Setpoint [°C]
Heating	1/10	31/03	20
Cooling	1/04	30/09	26

9. Weather Files

EnergyPlus simulation files require input variables, all monitored or derived from the dataset such as the DNI. Except for the Perez model that calculates DNI directly, in the other cases it was estimated through Equation (6).

$$DNI = DIR / \cos(Z) \quad (6)$$

Due to the uncertainty of the measuring instrument at high zenith angles, Equation (6) can return extremely high values (that were previously eliminated from the quality control, Equation 3). This can lead to unphysical values, even higher than extraterrestrial solar radiation, which can cause simulation errors if used in building simulation codes. The issue was solved, as was recommended by Spinelli F. (personal communication, 2016, ENEA - Italian National agency for new technologies, Energy and sustainable economic development) with the substitution of out of range values. The problematic points were identified by applying the check reported in Equation (7) (Gueymard et al., 2016), with the DNI values obtained using the Bird clear-sky model according to Sengupta and Gotseff (2013).

$$DNI < 1100 + 0.03 \cdot Elev \quad (7)$$

The application of Eqn. 7 resulted in 95 points replacement in the dataset, 499 in Oliveira, 533 for Al Riahi, and 499 for Torres, and none for the Perez model. For all the variables previously described, the dataset was reduced from 10-minute detection to hourly weather files by averaging.

10. Simulation Results

Ten-year measurements and four selected models were used to obtain 40 simulations related to the split models. The results were compared to a reference simulation with weather files obtained using the original dataset. In the following paragraphs, the energy required for heating and cooling are analysed. As previously explained, we focused mainly on two different years, 2003 and 2004.

Tables 9 and 10 report the error in seasonal cooling or heating energy, as defined in Equation (8), where subscripts "*se*" and "*sm*" represent respectively the simulation result with estimated and measured weather files.

$$Error = \frac{Energy_{se} - Energy_{sm}}{Energy_{sm}} \quad (8)$$

Table 9 – Cooling energy errors for the various models

Year	Sim [kWh]	Oliveira [%]	Perez [%]	Al Riahi [%]	Torres [%]
2001	2678	+1,9	+0,5	+1,9	+1,6
2002	2539	+1,7	+0,2	+1,5	+1,4
2003	3852	+1,8	+0,4	+1,4	+1,4
2004	2585	+2,4	+0,5	+2,2	+1,9
2005	2377	+2,9	+1,3	+2,6	+2,6
2006	2718	+3,2	+0,8	+3,5	+2,7
2007	3001	+3,6	+1,4	+3,3	+3,1
2008	2910	+2,7	+1,8	+2,4	+2,5
2009	3190	+3,9	+2,5	+3,8	+3,6
2010	2580	+3,6	+2,2	+3,3	+3,3

Table 10 – Heating energy errors for the various models

Year	Sim [kWh]	Oliveira [%]	Perez [%]	Al Riahi [%]	Torres [%]
2001	2516	-0,3	+3,0	-2,9	+0,1
2002	2381	-1,5	+2,0	-3,9	-1,1
2003	2234	-1,3	+2,8	-4,4	-0,6
2004	2706	-1,8	+1,4	-4,4	-1,3
2005	2839	-2,4	+1,4	-5,1	-2,0
2006	2403	-1,5	+2,6	-4,2	-0,9
2007	2040	-2,7	+2,1	-5,8	-2,3
2008	2429	-2,9	+1,0	-5,5	-2,4
2009	2544	-3,6	+0,6	-6,0	-3,3
2010	3249	-3,0	-0,4	-5,0	-2,8

As can be seen from Table 9, the Perez model in estimating the cooling energy preforms remarkably better than the others, with errors ranging from 0.2 % to 2.5 %. However, when considering heating energy, the Perez model shows the same performance of the Oliveria and Torres models. The result confirms the statistical analysis presented in Tables 3 to 5 where the Perez model showed the best results in the cooling period, but with lower performance during the heating one.

Furthermore, the numerical simulation confirms the distribution of the statistical error with the box plots presented in Fig. 1: a model can perform very differently if applied to different years on the same location. Considering Table 9, the Torres and Al Riahi models present the same error for the year 2003, while the latter shows consistently higher errors for different years.

Finally Fig.s 5 and 6 graphically show the errors trend reported in Tables 9 and 10, showing the good performance of the Perez model in cooling simulations, and the comparable performance of the Torres and Oliveira models for heating.

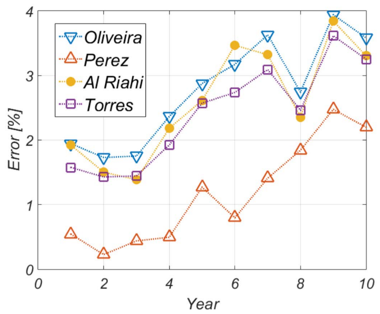

Fig. 5 – Cooling energy, percentage error between models and measured data

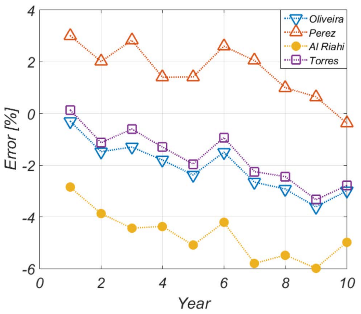

Fig. 6 – Heating energy, percentage error between models and measured data

11. Conclusions

33 split models were implemented in order to study their impact in building energy simulations. This analysis was developed both with statistical indicators and with simulation results, through EnergyPlus software. Considering the whole 10-year dataset, the Perez model performs significantly better than the others, both for MAD and RMSD statistic parameters, but with remarkable differences between cooling and heating periods. It shows also consistency if applied to different years of the same location. Regarding the polynomial models, they can perform with satisfying accuracy if the current dataset has similar climatic characteristics to the one originally used to obtain the model. Those that behave better in this case are Oliveira and Torres.

Simulation results, applied on a test building with a well-insulated envelope, have shown that the Perez model performs extremely well for what concerns the cooling energy. On the other hand, the model shows higher energy needs during the heating period, but with a performance comparable to the other models.

The Al Riahi, with a greater range of variability between different years, for this reason included in the set of tested models, showed a trend similar to the Torres and Oliveira, but with more errors for the selected years, and critical results for heating due to its constant underestimation of the required energy.

Finally, the Perez model shows the best overall performance. Nevertheless, the issue of lower performance in simulations concerning the heating period, especially for insulated buildings, cannot be neglected. Therefore, further investigation is required to identify the different behavior of the Perez model between heating and cooling periods.

Regarding the influence of split models in simulation applied to insulated buildings, the choice can affect the simulation error from 0.2 % up to 4 %, this result is remarkable and has to be considered when dealing with building detailed energy analysis.

Nomenclature

Symbols

Cp	Thermal capacity (kJ kg^{-1} K^{-1})
DIF	Diffuse horizontal solar radiation (Wh m^{-2})
DNI	Normal solar radiation, (Wh m^{-2})
DIR	Horizontal direct solar radiation, (Wh m^{-2})
E0	Extraterrestrial solar radiation, (Wh m^{-2})
E0h	Extraterrestrial solar radiation, horizontal (Wh m^{-2})
Elev	Elevation (m)
GLO	Global horizontal solar radiation, (Wh m^{-2})
kd	Diffuse fraction (-)

kt	Clearness index (-)
U	Thermal transmittance (W m^{-2} K^{-1})
Z	Zenith angle (°)
λ	Thermal conductivity (W m^{-1} K^{-1})
ρ	Density (kg m^{-3})

References

Al-Riahi, M., N. Al-Hamdani, K. Tahir. 1992. "An empirical method for estimation of hourly diffuse fraction of global radiation". *Renewable Energy* 2(4): 451-456. doi:10.1016/0960-1481(92)90079.

Bird, R.E., R.L. Hultrom. 1981. *A Simplified Clear Sky Model for Direct and Diffuse Insolation on Horizontal Surfaces, SERI/TR-642-761*. Golden, U.S.A.: Solar Energy Research Institute.

Blanco, M.J., D.C. Alarcon-Padilla, T. López-Moratalla, M. Lara-Coira. 2001. "Computing solar vector". *Solar Energy* 70(5): 431-441. doi:10.1016/S0038-092X(00)00156-0.

Boland, J., B. Ridlay, B. Brown. 2007. "Models of diffuse solar radiation". *Renewable Energy* 33(4): 575-584. doi:10.1016/j.renene.2007.04.012.

Chikh, M., A. Mahrane, M. Haddadi. 2012. "Modeling the diffuse part of the global solar radiation in Algeria". *Energy Procedia* 18: 1068-1075. doi:10.1016/j.egypro.2012.05.121.

Elminir, H. 2005. "Experimental and theoretical investigation of diffuse solar radiation: data and models quality tested for Egyptian sites". *Energy* 32(1): 73-82. doi:10.1016/j.energy.2006.01.020.

Engerer, N.A. 2015. "Minute resolution estimates of the diffuse fraction of global irradiance for southeastern Australia". *Solar Energy* 116: 215-237. doi:10.1016/j.solener.2015.04.012.

Furlan, C., A.P. Oliveira, J. Soares, G. Codato, J.F. Escobedo. 2011. "The role of clouds in improving the regression model for hourly values of diffuse solar radiation". *Applied Energy* 92: 240-254. doi:10.1016/j.apenergy.

Gueymard, C., J.A. Ruiz-Arias. 2016. "Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance". *Solar Energy* 128: 1-30. doi:10.1016/j.solener.2015.10.010.

Jacovides, C.P., F.S. Tymvios, V.D. Assimakopoulos, N.A. Kaltsounides. 2006. "Comparative study of various correlations in estimating hourly diffuse fraction of global solar radiation". *Renewable Energy* 31(15): 2492-2504. doi:10.1016/j.renene.2005.11.009.

Lannini, F., S. Wunderle, M. Lehning, C. Fierz. 2010. "Division of global radiation into direct radiation and diffuse radiation". *Master thesis*. Bern, Switzerland: University of Bern.

Lee, K., H. Yoo, G.J. Levermore. 2013. "Quality control and estimation hourly solar irradiation on inclined surfaces in South Korea". *Renewable Energy* 57: 190-199. doi:10.1016/j.renene.2013.01.028.

Liu, B.Y.H., R.C. Jordan. 1960. "The interrelationship and characteristic distribution of direct, diffuse and total solar radiation". *Solar Energy* 4(3): 1-19. doi:10.1016/0038-092X(60)90062-1.

Macagnan, M.H., E. Lorenzo, C. Jimenez. 1993. "Solar radiation in Madrid". *International Journal of Solar Energy* 16(1): 1-14. doi: 10.1080/01425919408914262.

Perez, R., P. Ineichen, E. Maxwell, R. Seals, A. Zelenka. 1992. "Dynamic Global-to-Direct Irradiance Conversion Models". *ASHRAE Transactions* 98(1):354-369.

Pernigotto, G., A. Prada, P. Baggio P., A. Gasparella, A. Mahdavi. 2016. "Solar irradiance modelling and uncertainty on building hourly profiles of heating and cooling energy needs". In: *Proceedings of the IV International High Performance Buildings Conference*. West Lafayette, U.S.A.: Purdue University.

Reindl, D.T., W.A. Beckman, J.A. Duffie. 1990. "Diffuse fraction correlations". *Solar Energy*, 45(1): 1-7. doi:10.1016/0038-092X(90)90060-P.

Sengupta, M., P. Gotseff. 2013. *Evaluaton of Clear Sky models for satellite-based irradiance estimates, Technical Report NREL/TP-5D00-60735*. http://purl.fdlp.gov/GPO/gpo53784

Spencer, J.W. 1982. "A comparison of methods for estimating hourly diffuse solar radiation from global solar radiation". *Solar Energy* 29(1):19-32. doi:10.1016/0038-092X(82)90277-8.

Torres, J.L., M. De Blas, A. Garcia, A. de Francisco. 2010. "Comparative study of various models in estimating hourly diffuse solar irradiance". *Renewable Energy*, 35(6): 1325-1332. doi:10.1016/j.renene.2009.11.025.