Site-adaptation methods

1. Linear Regression (multiplicative)

This method can be referred to as the ratio method [30], wherein the data are corrected by scaling or multiplying by the calculated ratios of the ground observation and satellite-modeled data. This is a simple technique, but it only removes the mean bias [30] as shown in our initial results. This method can be used to correct the mean and maximum irradiance values [18]; in this study, we applied it to time series data using the following steps.

The slope *k* was determined by fitting a linear line to the observed and modeled data points that followed the following relationship:

$$\mathrm{O}_{\mathrm{train}} = k \times M_{train} \tag{S1}$$

Then, the test dataset was corrected using the obtained slope (*k*) as the correction factor:

$$\mathrm{M}_{\mathrm{test,corrected}} = k \times M_{test} \tag{S2}$$

2. Polynomial Fit

The training dataset was fitted using third-order polynomial regression:

$$\mathrm{y(x)} = \mathrm{p}_1 x^3 + p_2 x^2 + p_3 x \tag{S3}$$

where x denotes the satellite-modeled GHI and y is the observation GHI with no intercept [20].

Then, the obtained coefficients, p1, p2, and p3, were applied to the test dataset using the following equation:

$$\mathrm{M}_{\mathrm{test,corrected}} = \mathrm{p}_1 x^3 + p_2 x^2 + p_3 x \tag{S4}$$

where x denotes the satellite modeled GHI from the long-term dataset.

3. Multilinear Regression

Multilinear regression (MLR) can be defined as regression to a certain response variable using several predictor variables selected from an exhaustive list of variables [22,23]. For GHI, the response variables include the measured clearness index (k_t), and regression is performed to determine the best combination of these variables, i.e., modeled clearness index, modeled clear-sky index, relative air mass, and solar elevation angle. For DNI, it is based on the predicted value of DHI; then, the predicted DNI is calculated using the following equation: $\mathrm{GHI} = \mathrm{DNI}\cos(SZA) + \mathrm{DHI}$. The observed diffuse ratio (k_d) is the response variable for DHI MLR; its predictor variables include the relative air mass (m), solar elevation angle (θ_e), modeled clear-sky index, and modeled normalized clearness index.

The best combination of predictor variables is selected by the Akaike information criterion (AIC), which is used as a model to produce site-adaptation results. It should be noted that owing to this step, the model used for MLR site adaptation can be different for each unique condition (i.e., different solar irradiance components, sky conditions, site locations, or time periods).

The relative air mass can be calculated using the formula proposed by [31], which was then corrected by [32]; it represents a function of the solar zenith angle (*z*) and altitude (*h*) of the considered site. Moreover, $\mathrm{k}_{\mathrm{t,m}}$ was designed to diminish the solar zenith angle dependence of $\mathrm{k_t}$ by normalizing it with respect to a standard clear-sky GHI profile (normalized to 1 for a relative air mass of 1) [33]. The equations used for MLR are listed below, where TOA represents the top-of-atmosphere solar irradiance on the same plane ($\mathrm{S_0}.\cos z$) and CGHI is the clear-sky GHI, which comes from the UASIBS-KIER model.

$$\mathrm{k_t} = \frac{GHI}{TOA} \tag{S5}$$

$$k_c = \frac{GHI}{CGHI} \tag{S6}$$

$$k_d = \frac{DHI}{GHI} \tag{S7}$$

$$\mathrm{k_{t,m}} = \frac{k_t}{1.031\ \exp\left[-\frac{1.4}{0.9+\frac{9.4}{m}}\right] + 0.1} \tag{S8}$$

$$\mathrm{m} = \frac{1}{\cos z + 0.50572\cdot(96.07995 - z)^{-1.6364}}\cdot\exp\left(-\frac{h}{8434.5}\right) \tag{S9}$$

4. Quantile Mapping

Quantile mapping (QM) is based on the inverse transform [19,24] and briefly described here. From the general empirical CDF, formulated as

$$\mathrm{y} = F(x) \tag{S10}$$

where x is the sorted irradiance value and y is the probability of finding an irradiance whose value is less than or equal to x, we can perform QM using the following equation:

$$x_{sat,test,corrected} = F^{-1}_{obs,train}\left[F_{sat,train}\left(x_{sat,test}\right)\right] \tag{S11}$$

This method has several variations, such as ECDF mapping or various quantiles that select the number of quantiles (a free parameter defined by the user). In this study, we used three different quantities of quantiles, i.e., 5 (few), $\mathrm{N}^{0.5}$ (some), and N/5 (many), where N denotes the total number of data points [9]. Another variation of this method includes KDE mapping, where the CDF is obtained by integrating the probability density function (PDF) obtained from kernel density estimation [24].

5. Quantile Delta Mapping

Quantile delta mapping (QDM) is similar to the QM technique; however, the satellite-derived solar irradiance of the testing dataset is not transformed into the probabilistic profile of the training dataset. Instead, the probabilistic profile of the testing dataset is directly inversely transformed into bias-corrected solar irradiance values; subsequently, it is multiplied by the relative change between the solar irradiance of the testing and training datasets.

$$x_{sat,test,corrected} = F^{-1}_{obs,train}\left[F_{sat,test}\left(x_{sat,test}\right)\right] \times \Delta_r \tag{S12}$$

$$\Delta_r = \frac{F^{-1}_{sat,\,test}\left(y_{sat,test}\right)}{F^{-1}_{sat,\,train}\left(y_{sat,test}\right)} = \frac{x_{sat,test}}{F^{-1}_{sat,\,train}\left(y_{sat,test}\right)} \quad \text{(S13)}$$

Compared with the standard QM, QDM preserves the relative changes between variables in two different time periods; this is to ensure that the long-term trends in all quantiles are preserved. However, in the QM method, they can be artificially deteriorated (owing to the transformation of the probability profile of the testing dataset). Although the realistic nature of long-term trends related to natural or climate changes (such as temperature variability) depends on the model, the QDM technique helps in preventing artificial alterations in the trend of variable of interest during site-adaptation.

6. Polynomial Fit of ECDF

In this method, polynomial fitting is used, and the difference between the fitting of observed and modeled ECDFs is subtracted to correct satellite-modeled data [21]. The steps of this method are given below:

a. Transform the observed and satellite-modeled irradiance of the train dataset to ECDF, $y_{obs}(x_{obs})$ and $y_{sat}(x_{sat})$, where x is the sorted irradiance value and y is the probability of finding an irradiance whose value is less than or equal to x. Then, fit the ECDF points with 4-degree polynomial function.

$$x_{obs} = a_0 + a_1 y + a_2 y^2 + a_3 y^3 + a_4 y^4 \quad \text{(S14)}$$

$$x_{sat} = b_0 + b_1 y + b_2 y^2 + b_3 y^3 + b_4 y^4 \quad \text{(S15)}$$

b. Subtract the difference between the coefficients to obtain the correction factor:

$$\mathrm{D_i} = a_i - \mathrm{b_i} \quad \text{(S16)}$$

c. Transform the satellite-modeled irradiance of the testing dataset to ECDF and then fit the points with the fourth-degree polynomial function:

$$x_{sat,test} = s_0 + s_1 y + s_2 y^2 + s_3 y^3 + s_4 y^4 \quad \text{(S17)}$$

d. The corrected coefficients would be:

$$\mathrm{c_i} = s_i + \mathrm{D_i} \quad \text{(S18)}$$

e. Finally, the corrected long-term satellite-modeled irradiance can be expressed using the following equation:

$$x_{sat,test,corrected} = c_0 + c_1 y + c_2 y^2 + c_3 y^3 + c_4 y^4 \quad \text{(S19)}$$

Result of ranking procedure for GHI

Table S1. Statistical metrics and ranking for estimated GHI in clear-sky condition. The rank is provided for each metric (r_{1-6}) and overall rank.

Method	MBE (%)	r_1	MABE (%)	r_2	RMSE (%)	r_3	KSI (%)	r_4	OVER (%)	r_5	CPI (%)	r_6	Overall Rank
Initial	-3.99	17	6.96	9	12.6	12	98.67	17	23.84	10	36.93	15	14
LIN	1.35	7	8.34	10	12.72	17	78.13	8	21.98	9	31.39	9	9
POLY	2.86	9	8.6	11	11.96	9	79.48	9	14.96	8	29.59	8	8
MLR	1	4	4.21	1	5.56	1	33.83	2	3.08	3	12.01	2	2
ECDF	3.45	14	8.98	15	12.61	14	88.24	13	27.51	14	35.24	13	15
QM few	5.46	18	10.69	18	14.07	18	133.38	18	60.98	18	55.62	18	18
QM some	3.51	15	9.04	16	12.66	16	88.62	14	27.26	12	35.3	14	16
QM many	3.45	13	8.98	14	12.6	13	87.9	12	27.37	13	35.12	12	12
KDE	3.27	11	8.82	12	12.45	10	84.22	11	26.56	11	33.92	10	10
QDM	3.75	16	9.11	17	12.61	15	93.45	16	32.76	16	37.86	16	17
Polyfit	3.2	10	8.86	13	12.54	11	81.28	10	29.54	15	33.98	11	11
MLR-ECDF	1.26	6	4.52	6	6.14	6	43.66	6	4.7	6	15.16	6	6
MLR-QM few	3.36	12	6.19	8	8.71	8	89.93	15	47.58	17	38.73	17	13
MLR-QM some	1.36	8	4.62	7	6.33	7	44.67	7	5.13	7	15.62	7	7
MLR-QM many	1.26	5	4.52	5	6.14	5	43.26	5	4.63	5	15.04	5	5
MLR-KDE	0.99	3	4.26	2	5.67	2	34.74	3	4.53	4	12.65	3	3
MLR-QDM	0.99	1	4.3	3	5.74	4	38.09	4	2.04	1	12.9	4	4
MLR-Polyfit	0.99	2	4.32	4	5.73	3	33.05	1	2.25	2	11.69	1	1

Table S2. Same as Table S1 but for estimated GHI in cloudy-sky condition.

Method	MBE (%)	r_1	MABE (%)	r_2	RMSE (%)	r_3	KSI (%)	r_4	OVER (%)	r_5	CPI (%)	r_6	Overall Rank
Initial	-6.75	16	26.46	17	36.09	17	134.44	16	56.26	16	65.72	16	17
LIN	-11.48	18	26.56	18	36.14	18	189.83	18	112.28	18	93.6	18	18
POLY	-7.54	17	24.64	15	33.31	10	115.69	15	54.31	15	59.16	15	16
MLR	-3.4	7	22.57	8	30.68	6	153.6	17	76.29	17	72.81	17	12
ECDF	-4.65	8	23.97	11	33.69	12	82.9	10	25.75	10	44.01	10	11
QM few	-6.17	15	23.58	9	33.2	9	109.58	14	33.09	13	52.27	14	14
QM some	-4.66	10	23.98	12	33.7	13	81.69	8	25.07	8	43.54	8	9
QM many	-4.65	9	23.97	10	33.69	11	82.69	9	25.65	9	43.93	9	8
KDE	-4.75	12	24.42	14	33.97	14	77.04	7	23.57	7	42.14	7	10
QDM	-5.06	14	24.67	16	34.26	16	95.43	12	36.41	14	50.09	13	15
Polyfit	-4.76	13	24.39	13	34.14	15	86.71	11	30.08	11	46.27	11	13
MLR-ECDF	-2.72	2	22.19	3	30.58	3	51.98	3	4.52	4	29.42	4	3
MLR-QM few	-4.75	11	22.2	4	30.31	1	102.76	13	30.88	12	48.56	12	7
MLR-QM some	-2.77	3	22.17	1	30.55	2	51.27	1	4.3	2	29.17	1	1
MLR-QM many	-2.72	1	22.19	2	30.58	4	51.56	2	4.47	3	29.3	2	2
MLR-KDE	-2.82	4	22.4	7	30.63	5	52.04	4	4.19	1	29.38	3	4
MLR-QDM	-3.16	5	22.34	6	30.77	8	59.72	6	6.45	5	31.93	5	6
MLR-Polyfit	-3.39	6	22.31	5	30.72	7	59.35	5	7.74	6	32.14	6	5

Result of ranking procedure for DNI

Table S3. Same as Table S1 but for estimated DNI in all-sky condition.

Method	MBE (%)	r_1	MABE (%)	r_2	RMSE (%)	r_3	KSI (%)	r_4	OVER (%)	r_5	CPI (%)	r_6	Overall Rank
Initial	13.51	18	37.17	10	62.82	10	292.57	17	203.63	17	155.46	17	17
LIN	-12.81	17	31.8	7	51.87	2	233.56	15	155.85	16	123.29	16	13
POLY	-8.6	15	31.28	2	49.52	1	233.57	16	153.27	15	121.47	15	9
MLR	11.13	16	45.78	12	62.93	11	366.47	18	294.27	18	196.65	18	18
ECDF	-2.54	5	31.54	5	53.15	6	91.34	5	17.82	4	53.86	4	4
QM few	1.79	2	32.97	9	55.14	9	77.12	2	14.04	2	50.36	2	3
QM some	-2.59	7	31.54	6	53.17	7	91.07	4	18.12	5	53.88	5	6
QM many	-2.55	6	31.53	4	53.14	5	90.88	3	17.69	3	53.71	3	2
KDE	-2.68	8	31.11	1	52.67	3	103.51	6	22.3	6	57.79	6	5
QDM	-4.2	13	32.2	8	54.75	8	116.09	7	50.29	7	68.97	7	7
Polyfit	-2.02	3	31.48	3	52.7	4	72.61	1	6.7	1	46.18	1	1
MLR-ECDF	-2.69	9	47.49	17	68.44	17	145.24	10	58.25	10	85.09	10	14
MLR-QM few	-2.48	4	43.14	11	64.33	12	151.72	11	68.43	11	87.2	11	8
MLR-QM some	-2.77	11	47.44	16	68.41	15	143.91	8	57.75	8	84.62	8	10
MLR-QM many	-2.69	10	47.49	18	68.44	16	144.86	9	57.92	9	84.91	9	12
MLR-KDE	-0.11	1	46.34	14	66.78	13	175.59	14	90.63	13	99.95	13	11
MLR-QDM	-6.39	14	47.15	15	68.87	18	174.56	13	92.99	14	101.32	14	16
MLR-Polyfit	-3.35	12	46.28	13	67.07	14	158.4	12	78.51	12	92.76	12	15

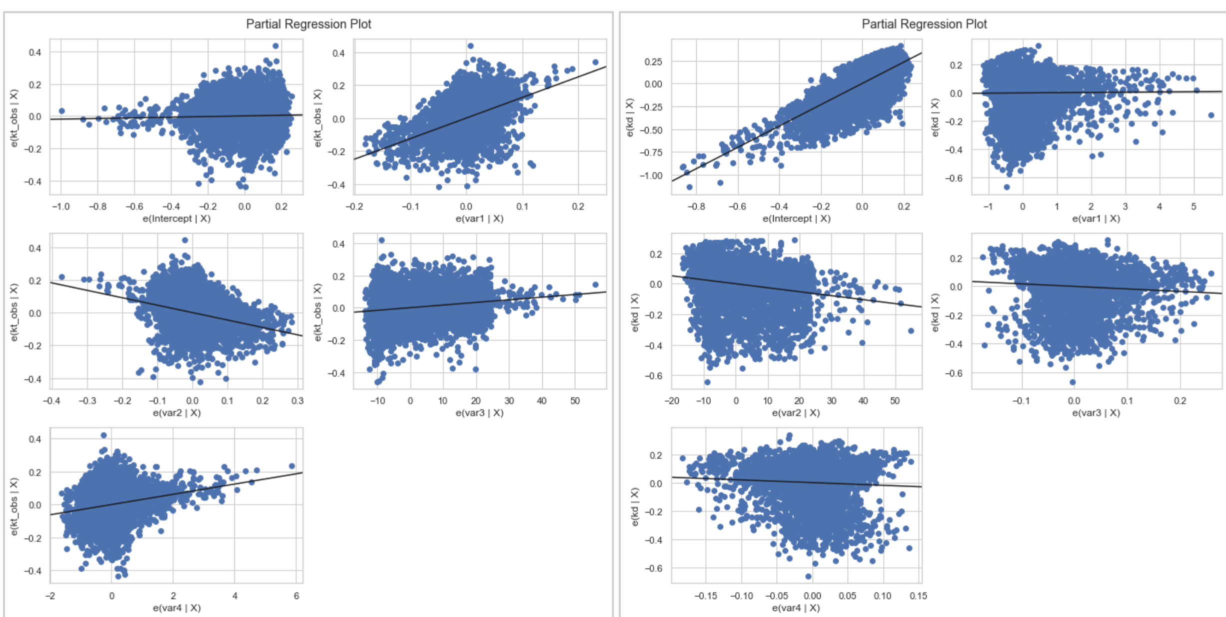
Table S4. Same as Table S1 but for estimated DNI in clear-sky condition.

Method	MBE (%)	r_1	MABE (%)	r_2	RMSE (%)	r_3	KSI (%)	r_4	OVER (%)	r_5	CPI (%)	r_6	Overall Rank
Initial	6.67	18	23.67	4	35.99	10	194.56	16	125.62	16	98.04	16	17
LIN	-6.15	17	23.96	9	34.31	2	165.56	14	97.55	15	82.93	15	12
POLY	-3.3	14	23.72	5	32.99	1	340.88	18	253.44	18	165.07	18	14
MLR	5.09	16	24.24	10	35	3	285.88	17	222.33	17	144.55	17	16
ECDF	1.09	9	23.88	7	35.51	8	66.65	4	7.76	4	36.36	4	6
QM few	2.12	12	23.53	1	35.07	4	76.12	6	16.59	7	40.71	6	7
QM some	1.06	6	23.89	8	35.53	9	65.69	2	7.84	5	36.15	2	4
QM many	1.09	8	23.88	6	35.51	7	66.33	3	7.66	3	36.25	3	2
KDE	1.07	7	23.58	2	35.29	6	83.39	7	16.27	6	42.56	7	5
QDM	-0.54	4	24.73	11	37.06	11	56.86	1	3.64	1	33.66	1	1
Polyfit	1.28	10	23.65	3	35.11	5	68.91	5	7.47	2	36.65	5	2
MLR-ECDF	0.24	3	30.28	16	41.19	17	124.63	10	53.42	11	65.1	10	10
MLR-QM few	3.52	15	27.64	12	38.22	12	168.53	15	84.38	14	82.34	14	18
MLR-QM some	0.15	1	30.22	15	41.16	15	123.73	8	52.65	9	64.68	8	8
MLR-QM many	0.24	2	30.28	17	41.19	16	124.37	9	53.17	10	64.98	9	9
MLR-KDE	1.7	11	29.54	13	40.47	13	138.88	12	55.46	12	68.82	12	13
MLR-QDM	-2.53	13	30.77	18	41.79	18	130.86	11	47.5	8	65.48	11	15
MLR-Polyfit	-0.65	5	29.77	14	40.47	14	149.44	13	65	13	73.85	13	11

Table S5. Same as Table S1 but for estimated DNI in cloudy-sky condition.

Method	MBE (%)	r_1	MABE (%)	r_2	RMSE (%)	r_3	KSI (%)	r_4	OVER (%)	r_5	CPI (%)	r_6	Overall Rank
Initial	50.59	18	110.35	10	206.8	12	250.27	17	157.12	16	205.25	16	17
LIN	-48.93	17	74.28	8	135.5	5	200.03	13	157.03	15	157.01	12	14
POLY	-37.31	15	72.24	2	127.32	1	202.1	15	154.45	14	152.8	11	11
MLR	43.84	16	162.47	18	213.71	18	413.67	18	339.52	18	295.15	18	18
ECDF	-22.26	8	73.02	6	135.59	8	140.13	7	71.92	9	120.81	5	6
QM few	0.03	1	84.13	9	156.35	9	145.24	9	67.46	5	131.35	7	5
QM some	-22.36	10	73	5	135.57	7	139.68	5	71.89	8	120.68	3	4
QM many	-22.31	9	72.98	4	135.51	6	139.98	6	71.86	7	120.71	4	3
KDE	-23.02	11	71.93	1	133.45	2	144.76	8	69.31	6	120.25	2	2
QDM	-23.98	12	72.7	3	134.94	3	152.62	10	81.74	10	126.06	6	7
Polyfit	-19.88	7	73.87	7	135.36	4	122.49	4	60.98	4	113.55	1	1
MLR-ECDF	-18.55	4	140.75	17	212.01	16	122.08	3	48.6	3	148.68	10	10
MLR-QM few	-35.01	14	127.12	11	202.76	10	246.51	16	181.39	17	208.36	17	16
MLR-QM some	-18.6	6	140.71	15	211.99	15	121.5	1	48.1	1	148.4	8	8
MLR-QM many	-18.56	5	140.75	16	212.01	17	122.07	2	48.59	2	148.67	9	9
MLR-KDE	-9.95	2	137.39	14	204.83	11	200.78	14	130.98	13	185.36	15	13
MLR-QDM	-27.29	13	135.88	13	210.86	14	170.37	11	96.83	11	172.23	13	15
MLR-Polyfit	-17.96	3	135.76	12	206.99	13	186.64	12	114.86	12	178.87	14	12

Partial regression plot for cloudy-sky condition



(**a**) (**b**)

Figure S1. Partial regression plot under cloudy-sky condition for (a) GHI and (b) DNI.