Check for updates

Evaluation and performance comparison of different models for global solar radiation forecasting: a case study on five cities

Mohamed A. Ali[1,2,3] · Ashraf Elsayed[2,4] · Islam Elkabani[2,4] · M. Elsayed Youssef[1] · Gasser E. Hassan[1,3]

Abstract

Recently, solar energy has emerged as the most promising renewable energy source to meet the world's energy demands. However, to harness the potential of solar energy, accurate data on solar radiation are crucial. It is considered the first step in assessing solar resources for various applications and achieving energy sustainability goals. Due to the unavailability of solar radiation measurements in many parts of the world, several models have been developed to predict global solar radiation (GSR) at these locations. Thus, this study aims to evaluate the proficiency of several GSR models at five new locations and determine the most suitable one for GSR prediction. The study has further developed solar radiation models for these new locations, as well as general ones for the entire region, which does not have any GSR models despite the existence of many planned solar energy facilities in this area. Additionally, the study investigates the effect of changing the length of the validation dataset on models' performance and accuracy, as well as assesses the introduced models' generalization capability. To achieve these objectives, the observed data of GSR for approximately 37 years at studied locations are used to develop and validate the proposed models. The study's findings reveal that Model 1 provides the best performance at all locations, with accuracy, coefficient of determination (R^2), ranging from 95 to 98%, except for the coastal location, where it is from 91 to 95%. The remaining performance indicators of the best models, such as RMSE, MABE, MAPE, and r, are good, and their values range from 0.7863 to 1.9097 (MJ m^{-2} day^{-1}), from 0.6430 to 1.7060 (MJ m^{-2} day^{-1}), from 3.4319 to 10.0890 (%), and from 0.9914 to 0.9981, respectively. The length of the validation dataset has a slight effect on the models' performance, ranging from about 1% to 2%. Therefore, Model 1 is the recommended solar radiation model, which can provide precise and rapid estimates of global solar radiation. This approach could be used in the design and performance evaluation of many solar applications. The primary benefit of this approach in the current investigation is that temperature data are continuously and effortlessly recorded for various purposes.

Keywords Solar energy resources · Solar radiation prediction · Empirical models · Temperature-based models · Statistical indicators · Egypt

Extended author information available on the last page of the article

Springer

1 Introduction

1.1 Overview and related work

In recent years, renewable energy sources have become increasingly important. The production of clean energy is essential for preventing climate change and mitigating its most severe consequences. The use of these clean energy sources can also promote sustainability goals, as well as support other sustainable concerns such as human resource development (Eesley, 2016; Piwowar-Sulej et al., 2023; Rao et al., 2023; Zhao et al., 2022). As the world grows more reliant on oil (Dagher & El Hariri, 2013), concerns about climate change and the concentration of mitigation actions on the energy sector have led to the implementation of energy conservation programs in many nations (Dagher & Yacoubian, 2012). Environmental pollution is one of the most significant barriers and hazards to sustainable development. Due to the conflicting effects of economic activity, overuse of natural resources, and environmental sustainability, environmental pollution has gained international attention (Alvarado et al., 2022; Chen & Dagestani, 2023a; Chen et al., 2023; Cihan et al., 2021; Dagestani et al., 2022, 2023; Shen et al., 2023; Su et al., 2023; Zhang et al., 2023). The use of sustainable energy sources, such as solar energy, can contribute to sustainable development and the reduction of the carbon footprint (Chen & Dagestani, 2023b; Kumar et al., 2023). Furthermore, the utilization of renewable energy sources may help improve the economy, which has suffered significantly in recent years due to various factors, including the spread of certain diseases (Dagestani, 2022; Guru et al., 2023; You et al., 2022), among other ways (Agarwal et al., 2023; Cheng et al., 2023; Dagar & Malik, 2023; Dagestani & Qing, 2022; Dagher, 2012; Tao et al., 2023; Zhou et al., 2023).

Currently, solar energy is regarded as one of the most potential renewable energy sources for meeting a significant portion of the world's energy needs (Ali et al., 2022; Chanchangi et al., 2022; Dagher & Ruble, 2010, 2011; Eroğlu, 2022; Feng et al., 2019; Tlili, 2015). Thus, accurate solar radiation information is considered the initial step in determining the availability of solar energy (Doorga et al., 2019; Guermoui et al., 2020; Hassan et al., 2016a, 2016b; Karaman et al., 2021; Makade et al., 2021; Mawire et al., 2021; Saleh et al., 2021; Yıldırım et al., 2018; Zaaoumi et al., 2021). Also, it is the fundamental input for many solar energy applications (Chand et al., 2021; Chaudhary & Yadav, 2021; He et al., 2020; Quej et al., 2016). Because solar radiation data are not available in many regions, many solar radiation models have been presented to forecast the global solar radiation (Nwokolo & Ogbulezie, 2018; Tao et al., 2021; Yorukoglu & Celik, 2006; Youssef et al., 2016). Much research has been conducted to investigate the usefulness of various solar radiation models in measuring solar radiation availability at various places worldwide. The most regularly validated and widely used solar radiation models are those based on meteorological factors (Besharat et al., 2013; Bounoua et al., 2021; Samuel Chukwujindu, 2017). These models rely on empirical correlations and meteorological data, such as sunlight duration, temperature, cloud cover, and relative humidity, which are the most often employed meteorological variables to predict Global Solar Radiation (GSR). The fundamental sunshine-based model was provided by Ångström (1924). Prescott (1940) adapted Ångström (1924) model that has become the most extensively used one to estimate solar radiation at various sites throughout the world (Almorox et al., 2005; Besharat et al., 2013). Al-Mostafa et al. (2014) studied the efficacy of fifty-two solar radiation models based on sunshine data to predict the monthly mean GSR on a horizontal plane in Jouf region, KSA. The finding demonstrated

that certain models are completely unsuited for usage in this region (Jouf), while others perform differently. Similarly, the best solar model for evaluating the monthly GSR on a horizontal ground at six sites in Algeria was presented by Mecibah et al. (2014). Despotovic et al. (2015) reviewed and analyzed different GSR models depending on sunshine data. They sought to analyze several models on a worldwide scale, which may be useful in selecting the most appropriate sunshine-based model. Barbaro et al. (1978) model was modified by Robaa (2009) to calculate GSR in Egypt. Hassan et al. (2018) investigated the effectiveness of several sunshine-based GSR models in the case study of New Borg El-Arab city, Egypt. The collected findings demonstrated that the general form of Robaa model (Robaa, 2009) performs well in calculating GSR. Ajayi et al. (2014) introduced a solar radiation model to predict daily values of GSR in Nigeria, which provided high agreement between estimated values and observed data. El-Metwally (2005) conducted a research on sunshine and GSR forecasting in several Egyptian sites, and he sought to create a simple nonlinear method to evaluate relative sunshine period and GSR.

Similarly, Hargreaves and Samani (1982) devised a simple model for predicting solar radiation that relied just on the lowest and maximum temperatures. Annandale et al. (2002) adapted Hargreaves and Samani (1982) model to calculate the effects of decreasing altitude and air thickness on GSR. As well, Allen (1997) provided a self-calibrated model based on Hargreaves and Samani model (1982) to predict monthly mean GSR. El-Metwally also suggested three simple new models for calculating GSR horizontally in Egypt depending on the lowest and maximum temperature and cloud cover (El-Metwally, 2004). Quej et al. (2016) investigated the accuracy and application of thirteen empirical-based models for calculating GSR in the Yucatan Peninsula, Mexico. The findings illustrate that the new model that was developed based on temperature and other meteorological variables provides the best forecast of GSR at all sites. Mghouchi et al. (2016) studied the proficiency of three models under all sky conditions in Tetuan, Morocco, to assess various parts of solar radiation flux on a horizontal plane. Youssef et al. (2016) evaluated the efficacy of more than thirty models, which do not depend on sunshine data, for calculating the GSR on horizontal ground. Their findings indicated that the models which are depend on solar declination angle and extra-terrestrial radiation, Türk Togrul and Onat (1999) and Ertekin & Yaldiz (1999) models, gave the most accurate prediction. Almorox et al. (2011) investigated the challenge of calibrating seven current solar radiation models and establishing a new model to predict GSR using temperature obtained from seven stations in Madrid, Spain. The collected findings demonstrated that empirical models that rely on temperature provide good predictions at any location if their coefficients are properly adapted. Furthermore, for all sites, the newly proposed model delivers the highest accurate forecast. Besharat et al. (2013) assessed the proficiency and application of existing empirical models for evaluating the monthly average daily GSR on a horizontal plane in Yazd, Iran. The findings illustrate that all developed correlations yield good predictions, with El-Metwally model (2005) providing the highest accurate prediction.

Generally, regression-based approaches to solar radiation estimation often have the benefit of having readily available equipment for measuring the climatic variables that are utilized as model input variables. As a result, most sites' data are often available. These regression models have demonstrated the ability to construct a strong and positive relationship between observed meteorological variables and solar radiation over a wide range of locations. They are also quite straightforward to perform and fairly precise. The equipment needed to measure sunshine and cloud data is frequently more expensive and not usually available at meteorological stations, in contrast to the equipment needed to

monitor temperature. This presumption suggests that knowledge about clouds or sunshine is typically unattainable (Babatunde et al., 2023).

1.2 Contribution statement

In general, sunshine-based models outperform other meteorological parameter-based models (Al-Mostafa et al., 2014). Sunshine data, on the other hand, is not as publicly available as other weather variables such as air temperature that are gathered at conventional weather stations (Li et al., 2014). As a result, sunshine-based solar models are challenging to apply in locations where sunshine information is absent (El-Metwally, 2005; Kumar et al., 2015). Owing to the absence of sunshine data at the majority of stations throughout the world, Hassan et al. (2016a, 2016b) proposed many new temperature-based models to forecast GSR as a substitute to the frequently employed sunshine-based models. To estimate the monthly average daily GSR horizontally, these new models are built, evaluated, and compared with three additional temperature-based models obtained from the literature. The findings display that the local model of the best model from these newly suggested temperature-based models (Hassan et al., 2016a, 2016b Model 6) provides accurate GSR forecasts at different sites. As well, they concluded that an accurate and quick prediction of global solar radiation can be provided using this newly suggested model.

In this sense, the goal of this work is to study the proficiency of several empirical models at five new locations to evaluate the strengths and weaknesses of these models and determine which one is the most suitable for GSR prediction. Furthermore, solar radiation models are developed for these new locations as well as general ones for the whole region, the Suez Canal Zone, which does not have any GSR models even though there are several proposed solar energy projects in this area. The Egyptian Ministry of Electricity, NREA, suggests a number of sites with solar energy potential for photovoltaic installations around Egypt. One such site is "ZAAFRANA Solar Park," which is situated in the investigated region (the Suez Canal's Zone). An essential first step in evaluating the feasibility and effectiveness of such solar energy application operations is evaluating estimates of solar radiation. Thus, Global Solar Radiation, GSR, observed data over 37 years at the studied locations are utilized, as a case study, to construct and evaluate the developed models in this article. The following points can be regarded as some of the novelties and contributions of this study:

- Development of exact GSR models for the researched areas as well as the entire region, which currently lacks GSR models despite the existence of many planned solar energy facilities.
- Investigate performance for three models at five new locations, especially at coastal ones, for accurate GSR forecasting on a horizontal surface.
- Study the effect of changing the length of the validation data set on models' performance and accuracy.
- Assessing performance for one of the best Temp-Based Models, Model 1 (Hassan et al., 2016a, 2016b), which has been presented recently, to estimate GSR at these new sites and after a period of years.
- Evaluation and performance comparison for these selected GSR models in evaluating Monthly Average Daily GSR on a horizontal plane (MADGSR).

- Present local formulas for estimating MADGSR at these cities, as well as general ones for the whole region (Suez Canal's Zone) due to its importance and the existence of different solar energy projects such as Zaafrana Solar Park.
- Perform a comparative study of the GSR models that have been identified as the most effective based on the results obtained from two recently proposed research studies, to evaluate the strengths and weaknesses of these models and determine which one is the most suitable for GSR prediction.

This comprehensive research provides important information for designers, engineers, and others interested in solar energy usage at the study sites. As mentioned above, these studied region has been planned to have several proposed solar energy projects such as "ZAAFRANA Solar Park." In addition, providing this essential information for different solar energy projects in this region will help in promoting sustainable development and economy, as well as the reduction of the carbon footprint.

1.3 Paper organization

To provide more explanation, Fig. 1 depicts the overall flowchart of the work and all of its phases, including data processing, model construction, performance evaluation, performance comparison, and, at the end, the major findings of the work. The remainder of the paper is arranged as follows: the global solar radiation models are described in Sect. 2.1. Then, Sect. 2.2 presents more information about the used indicators for evaluating models' performance and accuracy. Data collection and extra-terrestrial solar radiation calculation method are explained in Sect. 2.3. Section 3 introduces both experimental results and discussions for the studied models, including performance comparison for these developed models using different validation data sets. Finally, the conclusion and future work are presented in Sect. 4.
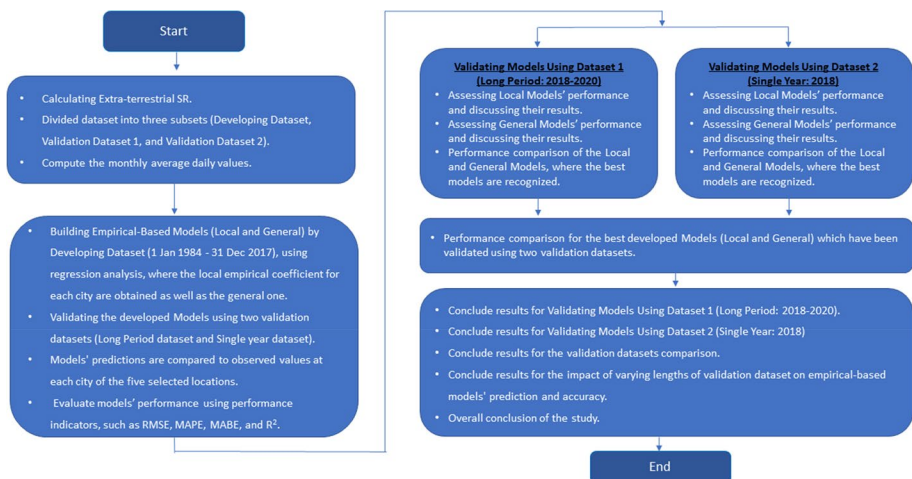

Fig. 1 General flowchart of the present study with a clarification for its various stages

2 Materials and methods

2.1 Solar radiation modeling

In general, the developed solar radiation models rely on linear and nonlinear relations. They present a correlation between incident solar radiation and other climatic parameters like cloud cover, relative humidity, temperature (minimum, ambient, or maximum), and sunshine period (Khalil & Shaffie, 2013a). Generally, the GSR models which depend on temperature can be considered one of the best models because temperature data are readily available compared with the other meteorological parameters, and it is already recorded very easily and continuously for other purposes (Li et al., 2014).

Hassan et al. (2016a, 2016b) proposed novel temperature-based models as another option to the frequently utilized ones, sunshine-based methods. The results showed that the local formula for the most accurate model, from these novel temperature-based models, provided excellent prediction for GSR at various sites. Also, this best model outperforms the two best models which depend on the sunshine parameter in the literature. In another paper (Youssef et al., 2016), they studied how thirty-one non-sunshine-based solar models perform in GSR prediction on a horizontal plane. These models are classified into six categories (only based on temperature, only based on cloud cover, only based on relative humidity, only based on alien solar energy and solar declination angle, only based on temperature, and only based on relative humidity and air temperature). The most accurate models which showed excellent performance within temperature-based group and temperature-relative humidity-based group will be selected and compared with the former one, the novel temperature-based model which has been presented recently for estimating global solar radiation (Hassan et al., 2016a, 2016b). These models are defined as follows:

Model 1 (Hassan et al., 2016a, 2016b):

$$G/G_0 = aT^b G_0 + c \tag{1}$$

Model 2 (Youssef et al., 2016):

$$G = a + bG_0 + cT \tag{2}$$

Model 3 (Youssef et al., 2016):

$$G = \left(aT_{\mathrm{Max}} + bT_{\mathrm{Min}} + cRH\right) \times G_0 + d \tag{3}$$

where a, b, c, and d represent the coefficients, and T, T_{Max}, T_{Min}, RH, G, and G_0 are the monthly mean values of daily air temperature (°C), maximum and minimum temperature (°C), relative humidity (%), GSR on a horizontal ground (MJ/m^2 day^{-1}), and extra-terrestrial solar radiation on a horizontal plane (MJ/m^2 day^{-1}), respectively.

For the sake of simplicity, temperature data are used in this study since, to a great extent, it is easier to get in most climate stations than other weather characteristics. Concerning its dependence on solar radiation, it is evident that the behavior of the surface of the earth in response to solar radiation obtained from the sun can shed light on the tight link between solar radiation and the surrounding temperature (Dincer et al., 1996). Shortwave electromagnetic radiation from the sun is absorbed by the earth's surface, warming the atmosphere as a result. Some of the absorbed energy is released as longwave radiation by the heated earth, warming the nearby ambient air. Rather than being heated directly by sunlight, the ambient air is heated indirectly through contact with the planet's surface. There is a phase delay in this process

between the cycles of temperature and solar radiation (Atkins, n.d.). In general, radiation balance and air mass advection can have an impact on variations in air temperature. The local air temperature and radiation balance are affected by the kind of surface cover, the percentage of clouds, the time of day, and the season. The annual temperature cycle, which demonstrates a strong correlation between solar radiation and surrounding temperature, would be a good way to track the regular variations in solar radiation that occur during the course of a year (Dincer et al., 1996). Thus, the primary focus of this study is to develop unique correlations between temperature and solar radiation, which is considered a major claim in the present study.

In fact, the main concept for developing the new economic zones is to create new societies providing new jobs to attract young people to this newly developing area. Moreover, developing these new generations of cities based on renewable energy resources such as solar energy creates more extra jobs in the field of renewable energy, which enhances the job market related to renewable and green energy. Furthermore, evaluating the solar energy resources in this new region using the developed numerical model increases the utilization of solar energy as a renewable source of energy reducing the carbon footprint for any economic activities in the region supporting the concept of environmentally friendly societies. This motivates the cultural shift of society toward the new concept of sustainable green cities for future development trends for new cities and societies. Thus, the current study attempts to develop an accurate model, which can be considered as a simple tool to evaluate solar energy resources based on temperature data available in a certain location. The predictions from this simple tool can be used as input to the feasibility studies related to any future development plans for solar energy projects at these locations. This enables the decision-makers to evaluate the economic feasibility and payback period for any future project related to solar energy.

2.2 Evaluating Models' Performance

The developed models' performance in this study is examined and evaluated using the most frequently applied indicators such as the coefficient of determination (R^2), mean percentage error (MPE), mean absolute percentage error (MAPE), mean bias error (MBE), root-mean-square error (RMSE), correlation coefficient (r), and mean absolute bias error (MABE) (Besharat et al., 2013; Cihan et al., 2021; Hassan et al., 2016a, 2016b; Youssef et al., 2016). The accepted values of these errors, MPE, MAPE, RMSE, MBE, and MABE, are between $\pm 10\%$, and the values of R^2 and r are between 0 and 1 ($0 \leq R^2, r \leq 1$), where their ideal values which close to the unity (Ajayi et al., 2014; Hai et al., 2020; Khorasanizadeh & Mohammadi, 2013a). The equations for these indicators are defined as:

$$\text{MBE} \;=\; \frac{1}{n} \sum_{i=1}^{n} \left(G_{i.c} \;-\; G_{i.m} \right) \tag{4}$$

$$\text{MPE} \;=\; \frac{1}{n} \sum_{i=1}^{n} \left(\frac{G_{i.c} - G_{i.m}}{G_{i.m}} \right) \;\times\; 100 \tag{5}$$

$$\text{RMSE} \;=\; \left[\frac{1}{n} \sum_{i=1}^{n} \left(G_{i.c} - G_{i.m} \right)^2 \right]^{1/2} \tag{6}$$

$$\text{MABE} \ = \ \frac{1}{n} \sum_{i=1}^{n} \left| \left(G_{i.\,c} \ - \ G_{i.\,m} \right) \right| \tag{7}$$

$$\text{MAPE} \ = \ \frac{1}{n} \sum_{i=1}^{n} \left| \left(\frac{G_{i.\,c} \ - \ G_{i.m}}{G_{i.\,m}} \right) \ \times \ 100 \right| \tag{8}$$

$$r \ = \ \frac{\sum_{i=1}^{n} \left(G_{i.\,m} \ - \ \overline{G_{m}} \right) \left(G_{i.c} \ - \ \overline{G_{c}} \right)}{\left[\sum_{i=1}^{n} \left(G_{i.m} \ - \ \overline{G_{m}} \right)^{2} \sum_{i=1}^{n} \left(G_{i.\,c} - \overline{G_{c}} \right)^{2} \right]^{1/2}} \tag{9}$$

$$R^{2} \ = \ 1 - \frac{\sum_{i=1}^{n} \left(G_{i.\,m} - G_{i.\,c} \right)^{2}}{\sum_{i=1}^{n} \left(G_{i.\,m} - \overline{G_{m}} \right)^{2}} \tag{10}$$

$$e \ = \ \left(\frac{G_{i,\,c} - G_{i,\,m}}{G_{i,\,m}} \right) \ \times \ 100 \tag{11}$$

where n is the observation number, $G_{i.c}$ and $G_{i.m}$ are the ith predicted and calculated values, and $\overline{G_{m}}$ and $\overline{G_{c}}$ are the average values of the observed and predicted values.

Coefficient of determination values, R^2, donate knowledge about the goodness of fit between the predicted and the measured values. Coefficient of determination and correlation coefficient values range from 0 to 1, with the greatest number being the intended value. Mean bias error (MBE) statistics provide info about the long-term model's proficiency. While the MBE's negative values point to underestimation, the positive values refer to overestimation, and the small values consider the desired ones. RMSE's value is always positive and gives feedback about the short-term model's efficacy. Smaller RMSE values imply a high model's accuracy, whereas zero values represent the optimum value. The agreement between the value of the measured and the calculated G for each month can be expressed by relative percentage error (e), and its preferable values lie between -10 and $+10$ (%).

2.3 Data collection

The study utilized weather data collected over a period of 37 years, from January 1, 1984 to December 31, 2020, to develop and test models that predict monthly average daily GSR on a horizontal plane. The data include temperature (minimum, ambient, and maximum), relative humidity, and GSR, and was obtained from NASA's Surface Meteorology and Solar Energy power data access, a source that has been used in several other studies (M. A. Ali et al., 2016; Fadare, 2009; Hassan et al., 2018; Hassan et al., 2016a, 2016b; Hassan et al., 2016a, 2016b; NASA Surface Meteorology and Solar Energy, n.d.; Youssef et al., 2016). As mentioned above, the study aims to verify the applicability of the established models for predicting the monthly average daily GSR for selected cities and the Suez Canal zone as a whole. Thus, the developed models' performance is evaluated and compared to observed

Fig. 2 Distribution of the selected locations in the Suez Canal's zone, Egypt (I-cias, n.d.)

Table 1 Geographical information for the five studied cities

#	Location	Latitude. N	Longitude. E
1	Port Said	31° 15 `	32° 18 `
2	El Kantara	30° 51 `	32° 18`
3	Ismailia	30° 35 `	32° 16 `
4	Fayid	30° 18 `	32° 18 `
5	Suez	29° 58 `	32° 33 `

data of GSR at the five selected sites. Furthermore, the suggested models' generalizability is investigated for the Suez Canal's zone and examined at the five studied locations too, namely Port Said, El Kantara, Ismailia, Fayid, and Suez. These selected cities are distributed over the Suez Canal's zone as demonstrated in Fig. 2, as well as their geographical information is illustrated in Table 1.

Concerning the NASA weather dataset, in general, it facilitates the study of climate and climatic phenomena by acting as an extensive and ongoing record of climatic data obtained from satellite observations. The NASA database is noteworthy due to its worldwide breadth and chronological consistency. It is a trustworthy source of data since it has been shown to be accurate enough in comparison with traditional observations made on the ground. As such, the NASA record may be used with effectiveness in situations when there are insufficient or no ground records. Despite the widespread belief that observations taken on land are more accurate than those taken from space, the necessary meteorological variables for

this investigation were derived using NASA data. There were two key considerations in this choice. First of all, the NASA collection contains all of the pertinent weather parameters for the regions that are being studied. Second, there are issues with the precision of on-site observations, which are frequently brought on by procedural mistakes, deficiencies in the collected datasets, or calibration problems that might result in inaccurate data collection (Abdo & El-Shimy, 2013; Fadare, 2009). On the other side, numerous measures made on the ground have been used to assess the accuracy of NASA data, which comes from satellite observations. NASA calculates that the Mean Bias Error (MBE) ranges from -2% to $+0.7\%$, despite the fact that the validity of these ground-based observations themselves is not usually carefully evaluated (NASA, n.d.)

Besides, a custom program developed in C# programming language is used to calculate the values of declination angle and extra-terrestrial solar radiation, as well as the monthly average daily values for all the mentioned parameters. Extra-terrestrial solar radiation, G_0, is the solar radiation that originates outside the earth's atmosphere, and it is defined as (Ayodele & Ogunjuyigbe, 2015; Jiang, 2009):

$$G_o = \frac{24 \times 3600 G_{sc}}{\pi} k \left[\left(\frac{\pi \omega}{180} \right) \sin(L) \sin(\delta) + \cos(L) \cos(\delta) \sin(\omega) \right] \quad (12)$$

G_{sc} donates the solar constant ($G_{sc} = 1367$ W/m^2) (Camacho et al., 2012; Khalil & Shaffie, 2013b), k refers to the earth's orbit eccentricity correlation factor, ω points to the hour angle at sunset (degree), L indicates to the latitude angle (degree), and δ is the declination angle (degree). k, δ, and ω are defined as (Karakoti et al., 2012):

$$k = \left[1 + 0.033 \cos \left(\frac{360N}{365} \right) \right] \quad (13)$$

$$\delta = 23.45 \sin \left[\frac{360}{365} (284 + N) \right] \quad (14)$$

$$\omega = \cos^{-1} \left[-\tan(L) \tan(\delta) \right] \quad (15)$$

N refers to the year-day number beginning from January 1.

3 Results and Discussion

The observed data of GSR, temperature (minimum, ambient, and maximum), and relative humidity are separated into three subgroups and averaged to provide the monthly average daily values. The first subset (January 1, 1984–December 31, 2017) is utilized for building models using regression analysis (Ajayi et al., 2014; Besharat et al., 2013; Khorasanizadeh & Mohammadi, 2013b). The coefficients' values matching the observed data in selected locations are acquired and summarized in Table 2. Similarly, the general empirical coefficients for the whole region of the Suez Canal's zone for three models are calculated and represented in Table 2.

On the other hand, the validation process is performed using two different validation data sets. The first one, from January 2018 to December 2020, is the average data of three years. The second validation data set, from January 2018 to December 2018, is the average data for one year, 2018. These two validation data sets will be utilized to assess and

Table 2 Empirical coefficients for the selected five cities and for the Suez Canal's Zone, Egypt

City	Model	a	b	c	d
Port Said	Model 1	0.00034	0.84062	0.50640	−2.52989
	Model 2	−7.66733	0.76573	0.20876	
	Model 3	0.02120	−0.01367	0.00708	
El Kantara	Model 1	0.00101	0.51049	0.47280	0.32544
	Model 2	−6.48772	0.76239	0.12309	
	Model 3	0.02966	−0.02460	0.00298	
Ismailia	Model 1	0.00101	0.51062	0.47281	0.32544
	Model 2	−6.48772	0.76239	0.12309	
	Model 3	0.02966	−0.02460	0.00298	
Fayid	Model 1	0.00089	0.53876	0.48038	1.36865
	Model 2	−6.22812	0.75362	0.12586	
	Model 3	0.03081	−0.02667	0.00178	
Suez	Model 1	0.00099	0.49708	0.50472	−0.07195
	Model 2	−6.09726	0.76975	0.12687	
	Model 3	0.03567	−0.03248	0.00321	
Suez Canal's zone formula	Model 1	0.00076	0.58719	0.48948	−0.05853
	Model 2	−6.51256	0.76026	0.14161	
	Model 3	0.03186	−0.02671	0.00338	

validate the established models in this study. Also, they will be used to investigate the effect of changing the length/period of the validation data set on models' performance and accuracy, using these long and short period of validation data set (average data of three years and average data of one year). The prediction of the three proposed models, Eqs. (1–3), are compared with the measured values of the monthly average daily global solar radiation, and the most common statistical indicators, named RMSE, MBE, MPE, MAPE, MABE, e, r, and R^2, are computed using equations Eqs. (4–11).

The following sections introduce the obtained results from the validation process using each validation data set (Validation Data Set 1: three-year data average, Validation Data Set 2: one-year data average) and their discussion. For each validation data set, firstly, the obtained results of the local formulas are discussed followed by the general formulas' results discussion, and hence a comparison between both formulas (Local and General) is presented. Besides, a performance comparison between these developed models (Model 1, Model 2, and Model 3) at each city is explained. Finally, the revealed results from these two validation data sets (Data Set 1 and Data Set 2) are compared and discussed, too.

3.1 Validation using Data Set 1 (2018–2020)

This validating data set employed the average data of three years, from January 1, 2018 to December 31, 2020. The measured data for all used parameters as well as the extra-terrestrial solar radiation data at each city are averaged to get the monthly mean values. The models' prediction is compared to the observed data at each city, where all statistical indicators (MBE, RMSE, MPE, MAPE, e, MABE, R^2 and r) are computed using Eqs. (4)–(11) for the developed models and summarized in Table 3. Based on the received results, the best model at each city is recognized after comparing the obtained statistical indicators

together, and it will have the largest R^2 value (Ajayi et al., 2014; Hassan et al., 2016a, 2016b; Hassan et al., 2016a, 2016b; Khorasanizadeh & Mohammadi, 2013a; Li et al., 2010). The greatest models are highlighted in bold as illustrated in Table 3.

Based on the comparison with the observed data of GSR, the developed models at the selected sites showed good performance, with the coefficient of determination (R^2) ranging from 94 to 97% at all locations except for Port Said city, it is larger than 91%. This fall in R^2 value can be justified by different climatic variables, particularly at coastal sites, like Port Said city, which is located very close to the Mediterranean Sea as displayed in Fig. 1 (El-Metwally, 2004; Khorasanizadeh & Mohammadi, 2013a). While the models' prediction is considerably declined at Port Said city, their prediction is very good at the four remaining cities. Despite this low performance at Port Said city, the values for all statistical errors (MBE, MPE, RMSE, MABE, and MAPE) are within the accepted limit of $\pm 10\%$, and R^2 values are greater than 91%, which refers to good fitting between predicted and measured values.

Overall, it can be noted that all models have excellent RMSE and MBE values which range from 0.9 to 1.8 (MJ m^{-2} day^{-1}) and from 0.7 to 1.7 (MJ m^{-2} day^{-1}), successively. Also, Model 1, Eq. (1), provides the best performance among all developed models. Despite Model 1 and Model 2 being based on the same input parameters, ambient temperature and extra-terrestrial solar radiation, Model 1 has the most accurate estimation. Its R^2 values at the five selected cities, Port Said, El Kantara, Ismailia, Fayid, and Suez, are 0.9196%, 0.9664%, 0.9664%, 0.9664%, 0.9664%, and 0.9727%, respectively. Model 2 ranked second, and its performance is significantly close to Model 1. Furthermore, although Model 3 depends on relative humidity as one of its input parameters, it does not improve its performance compared with the other models, even at coastal sites such as Port Said City. For more clarification, the best model's estimations are compared with the observed data at the five selected cities as demonstrated in Fig. 3. Similarly, Fig. 4 shows the statistical indicators graph for the local models at the five selected locations using validation Data Set 1 (data of three years average, 2018–2020).

General formulas, on the other side, all measured data of used parameters (T, T_{Max}, T_{Min}, RH, G, and G_0) at the five selected cities of the Suez Canal's zone (as displayed in Fig. 1) are averaged and utilized to assess the generalizability capacity of the developed models throughout the whole Suez Canal zone in Egypt. The general coefficients are calculated from the averaged data using regression analysis (Ajayi et al., 2014; Besharat et al., 2013; Khorasanizadeh & Mohammadi, 2013b) and given in Table 2. The estimation of the developed models' general formulae is compared to the observed data of the five cities. Table 4 summarizes the obtained statistical errors for the general formulas' prediction at each city.

According to the results, what stands out from Table 4 is that, whereas the performance for both Model 1 and Model 2 at Port Said city is improved considerably by about 4% (R^2 values increased from about 91% to more than 95%), the performance of Model 3 is dramatically deteriorated (R^2 values dropped from about 91% to about 78%) and its MPE values also exceeded the acceptable range $\pm 10\%$, equals 12.4784 (MJ m^{-2} day^{-1}). This decline may be returned to this model, Model 3, which depends on relative humidity as one of its input parameters, and when the five cities' data were averaged, it led to a significant change in RH values especially at this coastal site.

Additionally, Model 1 still ranked first in three cities, Port Said, Fayid, and Suez, and second in the remaining two cities, El Kantara and Ismailia, with performance very close to the first one, Model 3, with about 0.003% difference in R^2 values, 0.9801% and 0.9835%, respectively. Also, the models' performance at Suez City is slightly decreased by about 2% compared with the performance of the local formulas, and their R^2 values larger than

Table 3 Statistical errors for the local formulas of the proposed models using validation Data Set 1 (2018–2020)

Type	Site	Model	MPE	MBE	RMSE	MAPE	MABE	r	R^2	Rank
Local Formulas	Port Said	Model 1	**9.3092**	**1.6925**	**1.7998**	**9.3092**	**1.6925**	**0.9975**	**0.9196**	**1**
		Model 2	9.1813	1.7078	1.8371	9.1813	1.7078	0.9964	0.9163	2
		Model 3	9.6164	1.7298	1.8880	9.6164	1.7298	0.9941	0.9116	3
	El Kantara	Model 1	**−5.6514**	**−1.0118**	**1.1207**	**5.6514**	**1.0118**	**0.9979**	**0.9664**	**1**
		Model 2	−5.6833	−1.0009	1.1669	5.6833	1.0009	0.9961	0.9636	2
		Model 3	−6.3213	−1.2073	1.3454	6.3213	1.2073	0.9953	0.9516	3
	Ismailia	Model 1	**−5.6512**	**−1.0118**	**1.1207**	**5.6512**	**1.0118**	**0.9979**	**0.9664**	**1**
		Model 2	−5.6833	−1.0009	1.1669	5.6833	1.0009	0.9961	0.9636	2
		Model 3	−6.3213	−1.2073	1.3454	6.3213	1.2073	0.9953	0.9516	3
	Fayid	Model 1	**−5.6506**	**−1.0151**	**1.1219**	**5.6506**	**1.0151**	**0.9979**	**0.9664**	**1**
		Model 2	−5.6687	−1.0008	1.1647	5.6687	1.0008	0.9961	0.9637	2
		Model 3	−7.1951	−1.3792	1.4817	7.1951	1.3792	0.9961	0.9413	3
	Suez	Model 1	**−4.4699**	**−0.8082**	**0.9823**	**4.8015**	**0.8989**	**0.9971**	**0.9727**	**1**
		Model 2	−4.5183	−0.7952	1.0110	4.9156	0.8991	0.9961	0.9710	2
		Model 3	−5.8544	−1.1077	1.2686	5.9269	1.1275	0.9952	0.9544	3

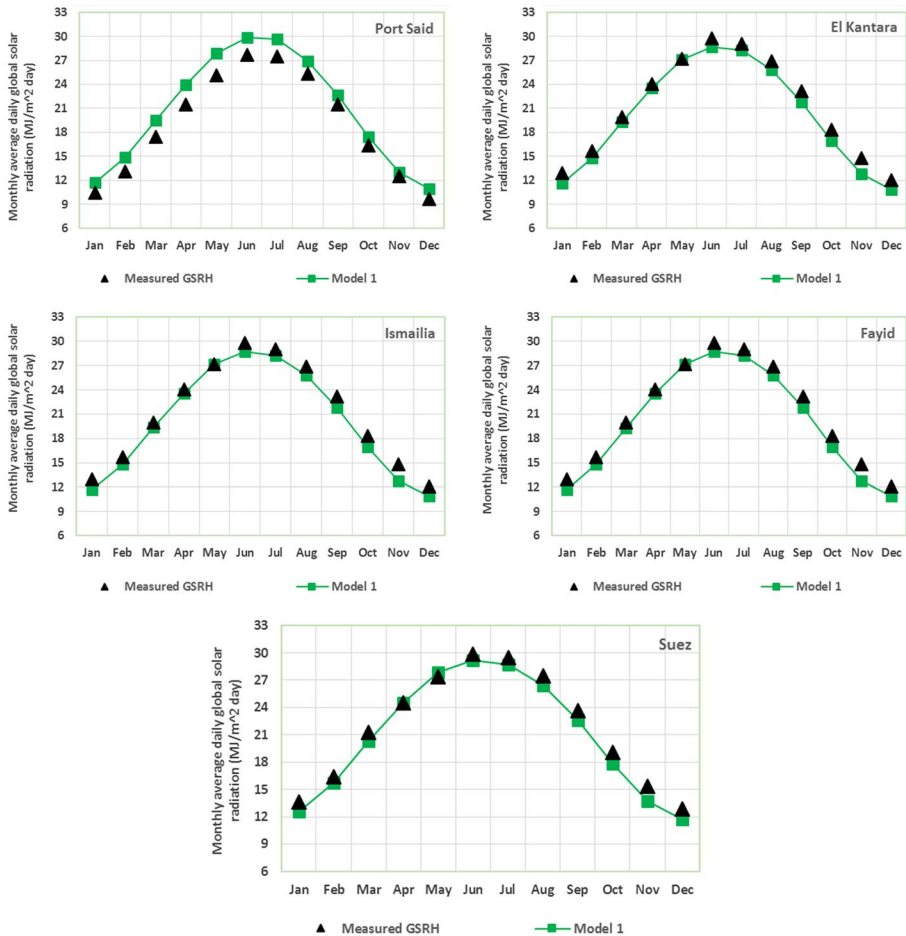Rows in bold refer to the most accurate model at each city

Fig. 3 Overall performance for the best models of local formulas using validation Data Set 1 (2018–2020)

94%. This marginal decline in models' performance at Suez City usually occurs in general models which are adjusted to cover larger areas or zones in contrast to local models that are calibrated on the local data (Hassan et al., 2016a, 2016b; Hassan et al., 2016a, 2016b; Li et al., 2014).

In general, it is also clear from Table 4 that the models' performance improved significantly at all locations except for Suez City, performance of the three models slightly decreased, as well as for Model 3 at Port Said City where performance dramatically collapsed. Additionally, Model 1 provides the best performance in comparison with the other models almost at all sites, especially at coastal ones. Moreover, its performance is approximately stable and reliable and does not experience a significant unfavorable change. In contrast, its prediction accuracy is improved in all cities except for Suez, where it has marginally declined. The R^2 values for Model 1 are between 0.95% and 0.98%, which refers to a good fitting between the model's estimation and measured values. Figure 5 displays the statistical error graph for the general models at the five selected sites using the validation
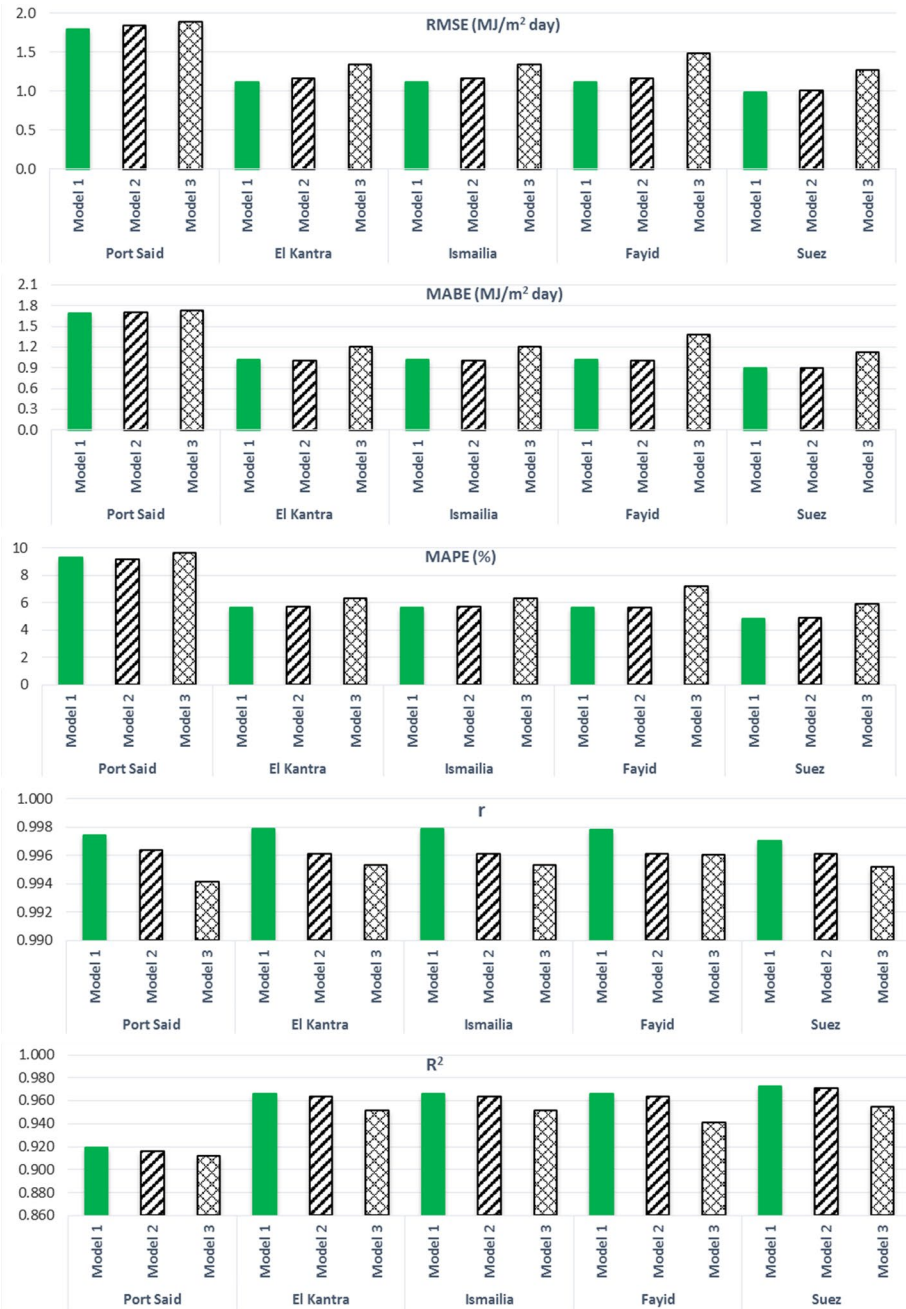
Fig. 4 Statistical errors graph for all models of the local formulas using validation Data Set 1 (2018–2020)

Data Set 1 (2018–2020). In addition, the forecast of the best general models compared against the observed data at the five selected cities is illustrated in Fig. 6.

Table 4 Statistical errors for the general formulas of the proposed models using validation Data Set 1 (2018–2020)

Type	Site	Model	MPE	MBE	RMSE	MAPE	MABE	r	R^2	Rank
Suez Canal's zone formulas	Port Said	Model 1	**6.7562**	**1.1688**	**1.3253**	**6.7562**	**1.1688**	**0.9954**	**0.9564**	**1**
		Model 2	6.9263	1.2219	1.4030	6.9263	1.2219	0.9943	0.9512	2
		Model 3	−12.4784	−2.5383	2.9285	12.4784	2.5383	0.9906	0.7872	3
	El Kantara	Model 1	−4.1501	−0.7002	0.8621	4.3253	0.7479	0.9982	0.9801	2
		Model 2	−4.1696	−0.6863	0.9165	4.4370	0.7498	0.9965	0.9776	3
		Model 3	**−0.6820**	**0.0126**	**0.7863**	**3.4319**	**0.6430**	**0.9949**	**0.9835**	**1**
	Ismailia	Model 1	−4.1501	−0.7002	0.8621	4.3253	0.7479	0.9982	0.9801	2
		Model 2	−4.1696	−0.6863	0.9165	4.4370	0.7498	0.9965	0.9776	3
		Model 3	**−0.6820**	**0.0126**	**0.7863**	**3.4319**	**0.6430**	**0.9949**	**0.9835**	**1**
	Fayid	Model 1	**−4.3326**	**−0.7279**	**0.8988**	**4.5357**	**0.7832**	**0.9981**	**0.9784**	**1**
		Model 2	−4.5113	−0.7366	0.9772	4.7493	0.7952	0.9964	0.9745	2
		Model 3	2.4329	0.6934	1.1204	4.4623	0.9701	0.9963	0.9665	3
	Suez	Model 1	**−6.4860**	**−1.2122**	**1.3194**	**6.5500**	**1.2297**	**0.9976**	**0.9507**	**1**
		Model 2	−6.5582	−1.2020	1.3467	6.5582	1.2020	0.9964	0.9486	2
		Model 3	−6.7539	−1.2994	1.3955	6.7539	1.2994	0.9967	0.9448	3

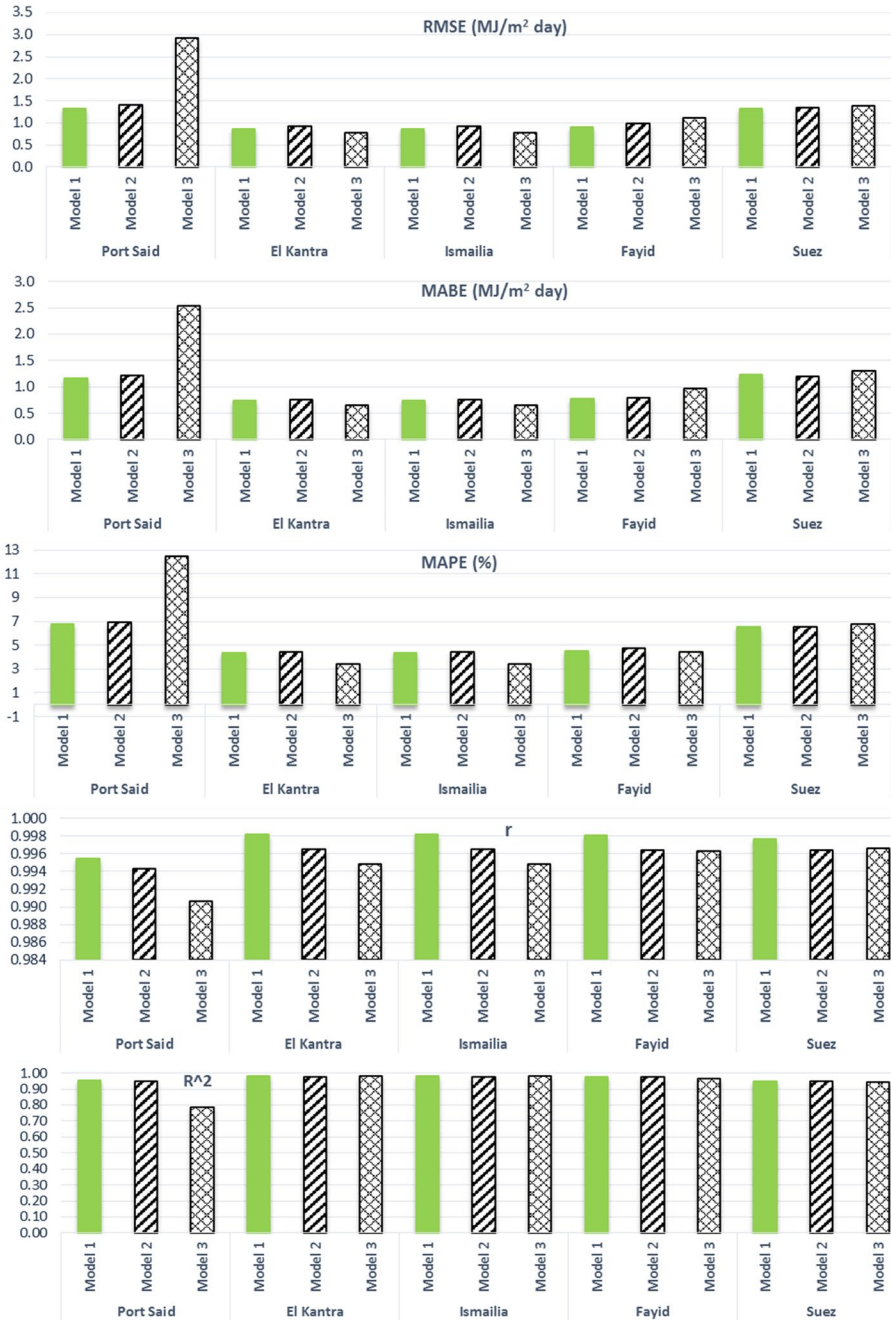Rows in bold refer to the most accurate model at each city

Fig. 5 Statistical errors graph for all models of the general formulas using validation Data Set 1 (2018–2020)

Regarding comparison, the performance for both the local and the general formulas is compared based on the result of the validation process using Data Set 1 (2018–2020). The best models from both formulas are selected and compared together as presented in Table 5. The most accurate model at each location is identified and indicated in bold as seen in the table. It is worth noting that while the general formula of Model 3 provided the best performance at two cities, El Kantara and Ismailia, with R^2 values > 98%, the formulas of Model 1, either the local one or the general one, showed the best performance at the remaining three cities, Port Said, Fayid, and Suez, and its R^2 values are greater than 95%,
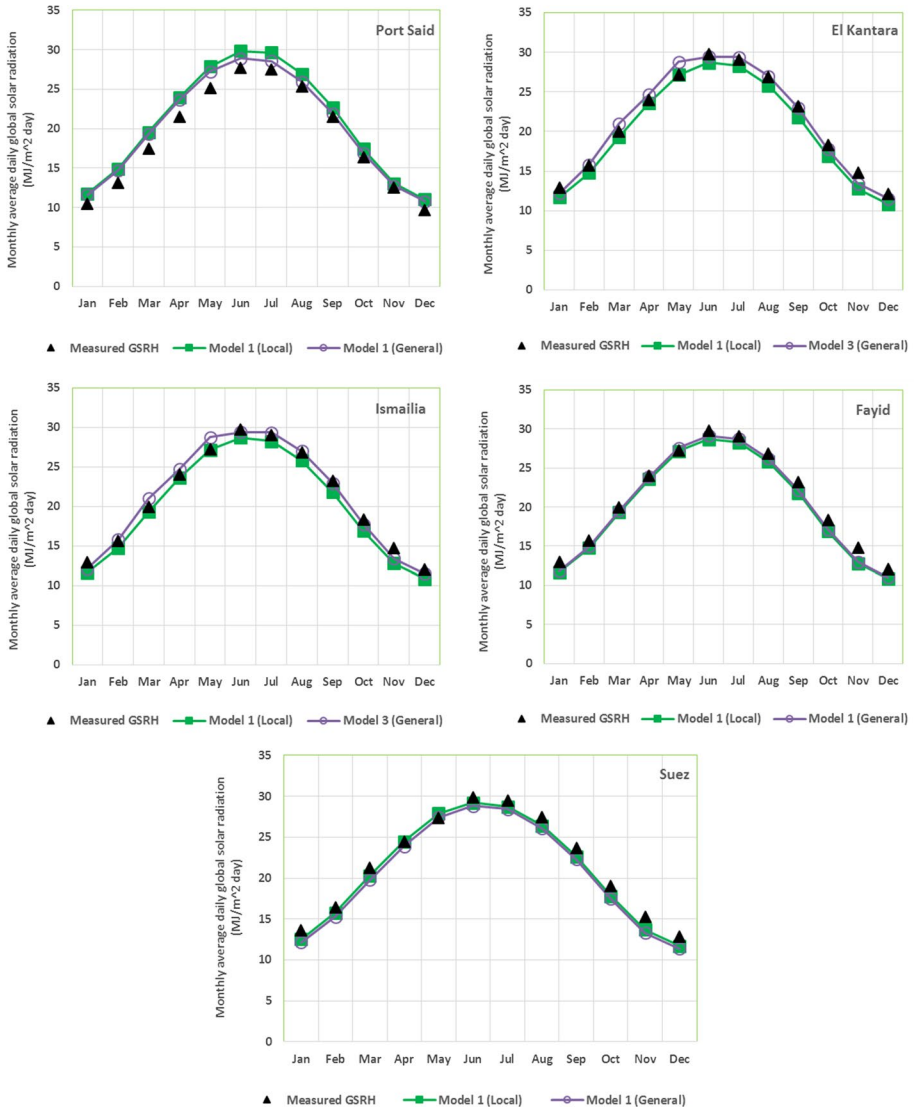
Fig. 6 Prediction of the best local models and the best general models using validation Data Set 1 (2018–2020)

Table 5 Performance comparison for the best models (local and general) using validation Data Set 1 (2018–2020)

Site	Model type	MPE	MBE	RMSE	MAPE	MABE	r	R^2	Rank
Port Said	Model 1-local	9.3092	1.6925	1.7998	9.3092	1.6925	0.9975	0.9196	2
	Model 1-Canal's zone	**6.7562**	**1.1688**	**1.3253**	**6.7562**	**1.1688**	**0.9954**	**0.9564**	**1**
El Kantara	Model 1-local	−5.6514	−1.0118	1.1207	5.6514	1.0118	0.9979	0.9664	**2**
	Model 3-Canal's zone	**−0.6820**	**0.0126**	**0.7863**	**3.4319**	**0.6430**	**0.9949**	**0.9835**	**1**
Ismailia	Model 1-local	−5.6512	−1.0118	1.1207	5.6512	1.0118	0.9979	0.9664	**2**
	Model 3-Canal's zone	**−0.6820**	**0.0126**	**0.7863**	**3.4319**	**0.6430**	**0.9949**	**0.9835**	**1**
Fayid	Model 1-local	−5.6506	−1.0151	1.1219	5.6506	1.0151	0.9979	0.9664	**2**
	Model 1-Canal's zone	**−4.3326**	**−0.7279**	**0.8988**	**4.5357**	**0.7832**	**0.9981**	**0.9784**	**1**
Suez	Model 1-local	**−4.4699**	**−0.8082**	**0.9823**	**4.8015**	**0.8989**	**0.9971**	**0.9727**	**1**
	Model 1-Canal's zone	−6.4860	−1.2122	1.3194	6.5500	1.2297	0.9976	0.9507	2

Rows in bold refer to the most accurate model at each city

97%, and 97%, respectively. Also, its local formulae at the remaining two sites, El Kantara and Ismailia, have excellent prediction with $R^2 > 96\%$, and its general formula at Port Said city improved the performance ($R^2 > 95\%$) with about 4% compared to its local formula.

For more clarification, the results of the performance comparison for the best local model and the best general model at the five locations are displayed in Fig. 6, and graphs of their statistical indicators are represented in Fig. 7. Furthermore, the relative percentage error (e) is calculated for these best models (local and general) through the twelve months and summarized in Table 6. Generally, what is clear from the table is that while local and general models performed lower at the coastal city, like Port Said, their values of relative error slightly exceeded the preferred range $\pm 10\%$ for some months, they showed perfect performance at the remaining four cities almost for all months except for some Winter's months, for example, the local formulas of Model 1 in November. The reason for this can be explained by different weather conditions particularly in Winter months such as clouds and rains (El-Metwally, 2004; Hassan et al., 2016a, 2016b; Hassan et al., 2016a, 2016b; Khorasanizadeh & Mohammadi, 2013a). Also, it is noted that the relative error values were reduced at Port Said (coastal city) in contrast with its values at Suez city. This may be interpolated by the averaging process for the measured data of the five selected cities leading to improved values of different parameters (T, T_{Max}, T_{Min}, RH, G, and G_0) at Port Said city, and vice versa at Suez city. On the contrary, the relative error values for the general formulas of Model 3 at El Kantara and Ismailia are in the range. For more clarification, the relative errors graph for the best local and general models is shown in Fig. 8.
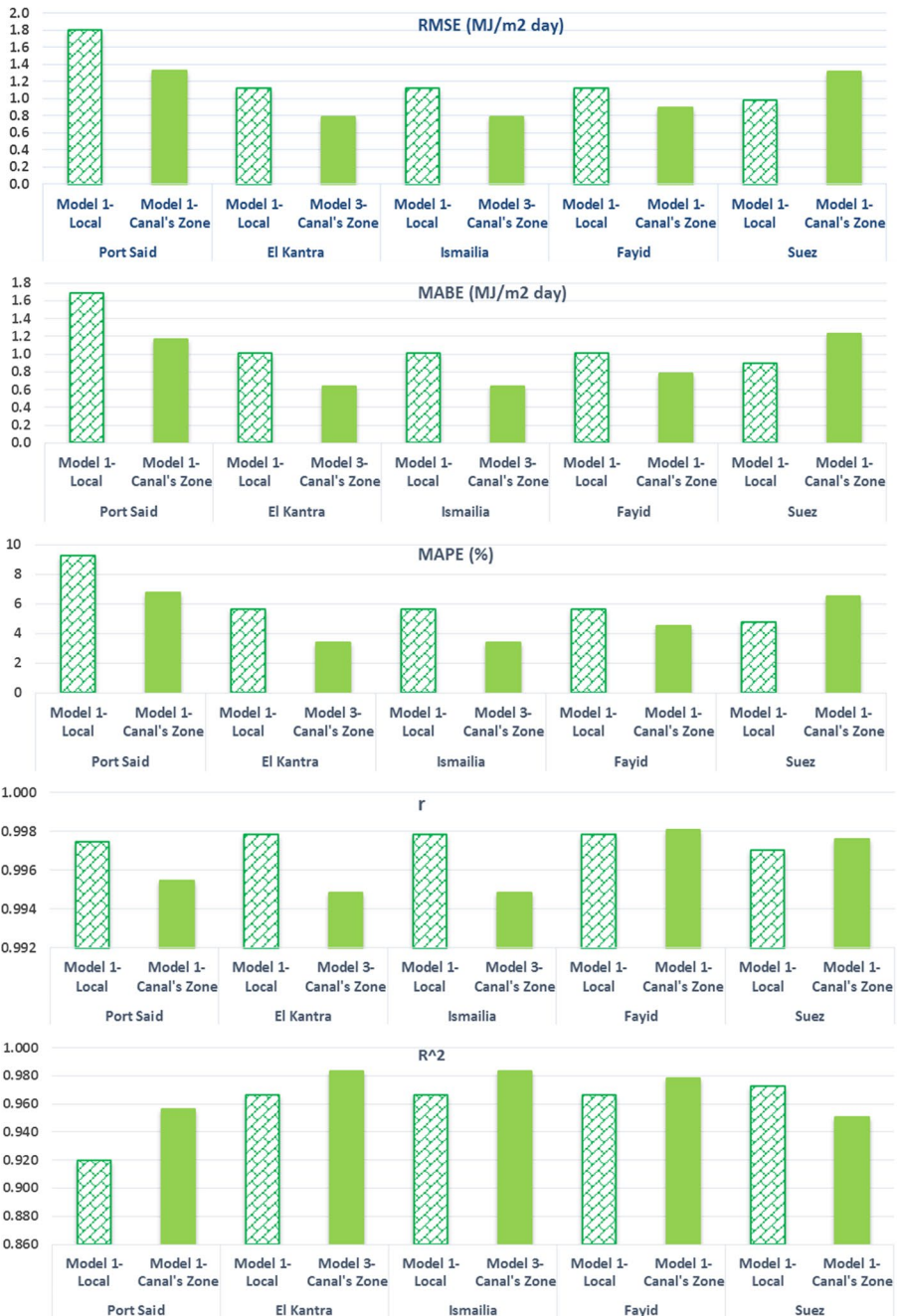
Fig. 7 Errors comparison for the best local and general models using validation Data Set 1 (2018–2020)

Table 6 Relative errors for the best local and general models using validation Data Set 1 (2018–2020)

| Type	Month	Port Said (Model 1)	El Kantara (Model 1)	Ismailia (Model 1)	Fayid (Model 1)	Suez (Model 1)
Local Formulas	Jan	12.2	−9.9	−9.9	−9.8	−8.2
	Feb	13.5	−5.8	−5.8	−5.8	−3.8
	Mar	11.7	−3.0	−3.0	−3.1	−4.4
	Apr	11.7	−1.8	−1.8	−1.9	0.0
	May	10.9	−0.2	−0.2	−0.2	2.0
	Jun	7.6	−3.6	−3.6	−3.7	−2.2
	Jul	7.7	−2.6	−2.6	−2.7	−2.5
	Aug	6.3	−4.0	−4.0	−4.0	−4.0
	Sep	5.4	−5.9	−5.9	−6.0	−4.4
	Oct	6.7	−7.6	−7.6	−7.6	−6.4
	Nov	4.4	−13.3	−13.3	−13.2	−10.6
	Dec	13.5	−10.0	−10.0	−9.8	−9.1
General Formulas		(Model 1)	(Model 3)	(Model 3)	(Model 1)	(Model 1)
	Jan	10.7	−6.0	−6.0	−8.7	−11.1
	Feb	12.1	0.8	0.8	−4.6	−6.7
	Mar	10.4	5.4	5.4	−2.0	−7.0
	Apr	10.2	2.8	2.8	−0.7	−2.4
	May	8.4	5.8	5.8	1.2	0.4
	Jun	4.3	−1.2	−1.2	−2.2	−3.5
	Jul	3.9	1.2	1.2	−1.1	−3.6
	Aug	2.4	0.5	0.5	−2.5	−5.1
	Sep	1.8	−1.0	−1.0	−4.5	−5.8
	Oct	3.6	−2.9	−2.9	−6.2	−8.3
	Nov	1.9	−9.1	−9.1	−12.0	−12.9
	Dec	11.3	−4.5	−4.5	−8.6	−11.8
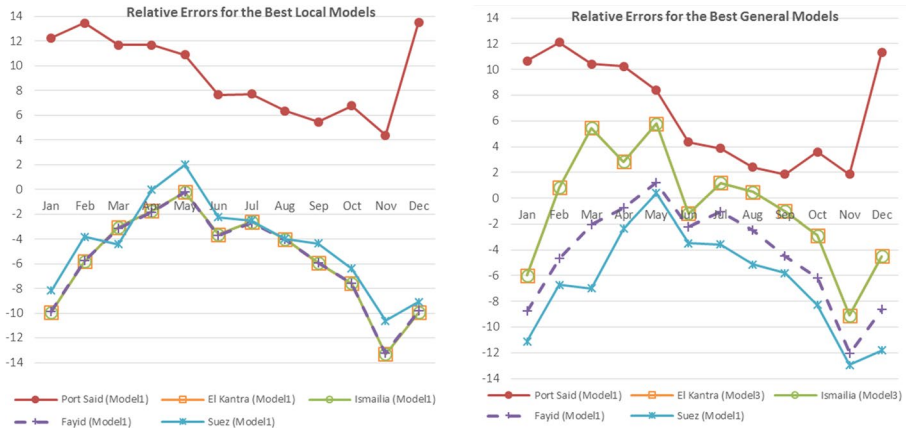
Fig. 8 Relative errors for the best models (local and general) using validation Data Set 1 (2018–2020)

3.2 Validation using Data Set 2 (2018)

Looking at the second data set, Validation Data Set 2 (2018), the same manner that has been carried out with Validation Data Set 1 is conducted with Data Set 2. Firstly, the measured data of all parameters for only one year, from January 1, 2018 to December 31, 2018, at each city are averaged to get the monthly average values. Then, models' predictions are compared against observed data, and all indicators are computed for both local and general formulas at each site as seen in Table 7. Also, the most accurate models are identified and pointed in bold as displayed in the table.

Focusing on the local formulas, the revealed results show that at the coastal city (Port Said), the three developed models showed low performance, and their MPE values slightly exceeded the range $\pm 10\%$, 10.5113, 10.3519, and 10.0890, respectively. Also, relative humidity, RH, which is an additional input parameter in Model 3 improves the model's performance by about 1% compared with the other models. This drop in models' performance returns to bad weather conditions at the coastal sites as mentioned before (El-Metwally, 2004; Khorasanizadeh & Mohammadi, 2013a). On the contrary, in the remaining four cities, while Model 3 provided moderate performance with R^2 values between 92 and 94%, Model 1 and Model 2 showed perfect performance with $R^2 > 96\%$ and their performance is almost identical. Generally, despite lower prediction at Port Said City, the models' prediction at the remaining cities was good especially Model 1 and Model 2, and their statistical indicators are in the acceptable ranges.

Regarding general formulas, it is very pertinent to note that at Port Said city, the performance of Model 3 is significantly deteriorated with R^2 equals 0.77% and MPE is larger than 12%, out of the range ($\pm 10\%$). This decrease can be explained by Model 3 mainly based on RH as one of its input parameters, which changed a lot when it was averaged for the five selected locations compared with its local values, as well as taking into consideration the difference within weather conditions at the coastal city (Port Said) and the other four cities. Model 1 gave the best prediction followed by Model 2 with $R^2 > 94\%$.

Table 7 Statistical errors for both formulas (local and general) of the proposed models using validation Data Set 2 (2018)

Type	Site	Model	MPE	MBE	RMSE	MAPE	MABE	r	R^2	Rank
Local formulas	Port Said	Model 1	10.5113	1.8491	2.0009	10.5113	1.8491	0.9949	0.9023	3
		Model 2	10.3519	1.8612	1.9893	10.3519	1.8612	0.9957	0.9034	2
		Model 3	**10.0890**	**1.7060**	**1.9097**	**10.0890**	**1.7060**	**0.9915**	**0.9110**	**1**
	El Kantara	Model 1	−5.3352	−0.9423	1.1776	6.0551	1.1312	0.9950	0.9623	2
		Model 2	**−5.4258**	**−0.9423**	**1.1675**	**6.0035**	**1.0939**	**0.9950**	**0.9629**	**1**
		Model 3	−6.7716	−1.3180	1.5294	6.9635	1.3684	0.9918	0.9364	3
	Ismailia	Model 1	−5.3350	−0.9423	1.1776	6.0549	1.1312	0.9950	0.9623	2
		Model 2	**−5.4258**	**−0.9423**	**1.1675**	**6.0035**	**1.0939**	**0.9950**	**0.9629**	**1**
		Model 3	−6.7716	−1.3180	1.5294	6.9635	1.3684	0.9918	0.9364	3
	Fayid	Model 1	−5.3578	−0.9508	1.1799	6.0634	1.1360	0.9950	0.9621	2
		Model 2	**−5.4382**	**−0.9472**	**1.1678**	**6.0051**	**1.0959**	**0.9951**	**0.9629**	**1**
		Model 3	−7.6445	−1.4999	1.6573	7.6445	1.4999	0.9933	0.9253	3
	Suez	**Model 1**	**−3.6626**	**−0.6399**	**1.0303**	**4.8051**	**0.9390**	**0.9927**	**0.9698**	**1**
		Model 2	−3.7344	−0.6319	1.0542	4.9124	0.9374	0.9919	0.9684	2
		Model 3	−6.0988	−1.1820	1.4173	6.2152	1.2105	0.9915	0.9429	3

Table 7 (continued)

Type	Site	Model	MPE	MBE	RMSE	MAPE	MABE	r	R^2	Rank
Suez Canal's zone formulas	Port Said	**Model 1**	**7.8213**	**1.3000**	**1.4965**	**7.8213**	**1.3000**	**0.9935**	**0.9454**	**1**
		Model 2	7.9415	1.3464	1.5184	7.9415	1.3464	0.9941	0.9437	2
		Model 3	−12.2228	−2.5589	3.0605	12.2228	2.5589	0.9847	0.7715	3
	El Kantara	Model 1	−3.8056	−0.6243	0.9682	4.7649	0.8760	0.9952	0.9745	3
		Model 2	−3.8695	−0.6186	0.9475	4.6841	0.8324	0.9952	0.9756	2
		Model 3	**−1.2140**	**−0.1170**	**0.9155**	**3.7969**	**0.7451**	**0.9914**	**0.9772**	**1**
	Ismailia	Model 1	−3.8056	−0.6243	0.9682	4.7649	0.8760	0.9952	0.9745	3
		Model 2	−3.8695	−0.6186	0.9475	4.6841	0.8324	0.9952	0.9756	2
		Model 3	**−1.2140**	**−0.1170**	**0.9155**	**3.7969**	**0.7451**	**0.9914**	**0.9772**	**1**
	Fayid	**Model 1**	**−4.0157**	**−0.6582**	**1.0025**	**4.9762**	**0.9102**	**0.9951**	**0.9727**	**1**
		Model 2	−4.2489	−0.6759	1.0074	5.0644	0.8899	0.9952	0.9724	2
		Model 3	1.9299	0.5550	1.0403	3.7556	0.8122	0.9944	0.9706	3
	Suez	**Model 1**	**−5.6500**	**−1.0350**	**1.2984**	**6.3978**	**1.2324**	**0.9932**	**0.9520**	**1**
		Model 2	−5.7608	−1.0332	1.3269	6.3898	1.1993	0.9922	0.9499	2
		Model 3	−6.7502	−1.3157	1.4698	6.7502	1.3157	0.9940	0.9385	3

Rows in bold refer to the most accurate model at each city

Alternatively, at the remaining locations, the three models displayed excellent estimation at El Kantara, Ismailia, and Fayid, where their performance is approximately alike, with R^2 values $> 97\%$. Moreover, Model 1 was the best one in Fayid and Suez cities. Overall, it can be observed that whereas some of the general formulas at some locations weakened or slightly decreased the performance, others, almost the majority of them, improved it. For example, the general formulas of Model 1 and Model 2 at Port Said city amended the performance significantly. Additionally, all models performed well except the general formula of Model 3 at Port Said city, and Model 1 has the most accurate prediction at three cities (Port Said, Faid, and Suez) and at the remaining two cities, its performance is approximately the same as the best model, Model 3. Of more interest, the statistical indicators for both local and general models at the five sites are graphed and demonstrated in Figs. 12 and 13 (in the Appendix).

Local and general formulas one-year comparison (2018), on the other hand, shows that firstly, Model 1 has the most stable and reliable performance among all models with good R^2 values. Also, if it is not ranked the first one, its performance will approximately coincide with the performance of the best. Performance of the general formulas generally improves models' performance with some exceptions, for instance, the general one of Model 3 at Port Said as well as the slight decrease in models' performance at Suez compared with its local one. Also, the general formulae of Model 3 is not recommended for use at the coastal sites. The overall performance comparison for both the best models, both local and general, at all cities and their statistical indicators are shown in Fig. 9, Fig. 10, respectively. Furthermore, the relative error values, (*e*), are calculated for these best models (local and general) through the year months and summarized in Table 8. It is worth mentioning that while relative error values for some months at the coastal city (Port Said) for both local and general formulas overstep the preferred range, the majority of its values at the other cities are in the range except for some winter's months own to unwell weather conditions as mentioned previously, where its values rarely and slightly exceed the scope. Also, the relative error values of general formulas are slightly better than local ones. For more illumination, the relative errors for the best local and general models are shown in Fig. 14 (in the Appendix).

3.3 Validation data sets comparison

Of more interest, the two validation data sets are compared together, Validation Data Set 1 (Three years data average: 2018–2020) and Validation Data Set 2 (One-year data: 2018). The obtained results from this comparison can provide several valuable information; for example, knowing the effect of changing the length of the validation data set on models' performance and accuracy, as well as determining the most accurate, stable, and reliable model at various sites that have different weather conditions. Therefore, the received results of statistical indicators from the two validation data sets, as previously shown in Table 3, 4, and 7, are compared together, as well as the revealed results of relative error in Table 6 and 8.

Looking at the local formulas of Validation Data Set 1 and 2, it can be marked that at the coastal city (Port Said city), while the performance of Model 1 and 2 marginally
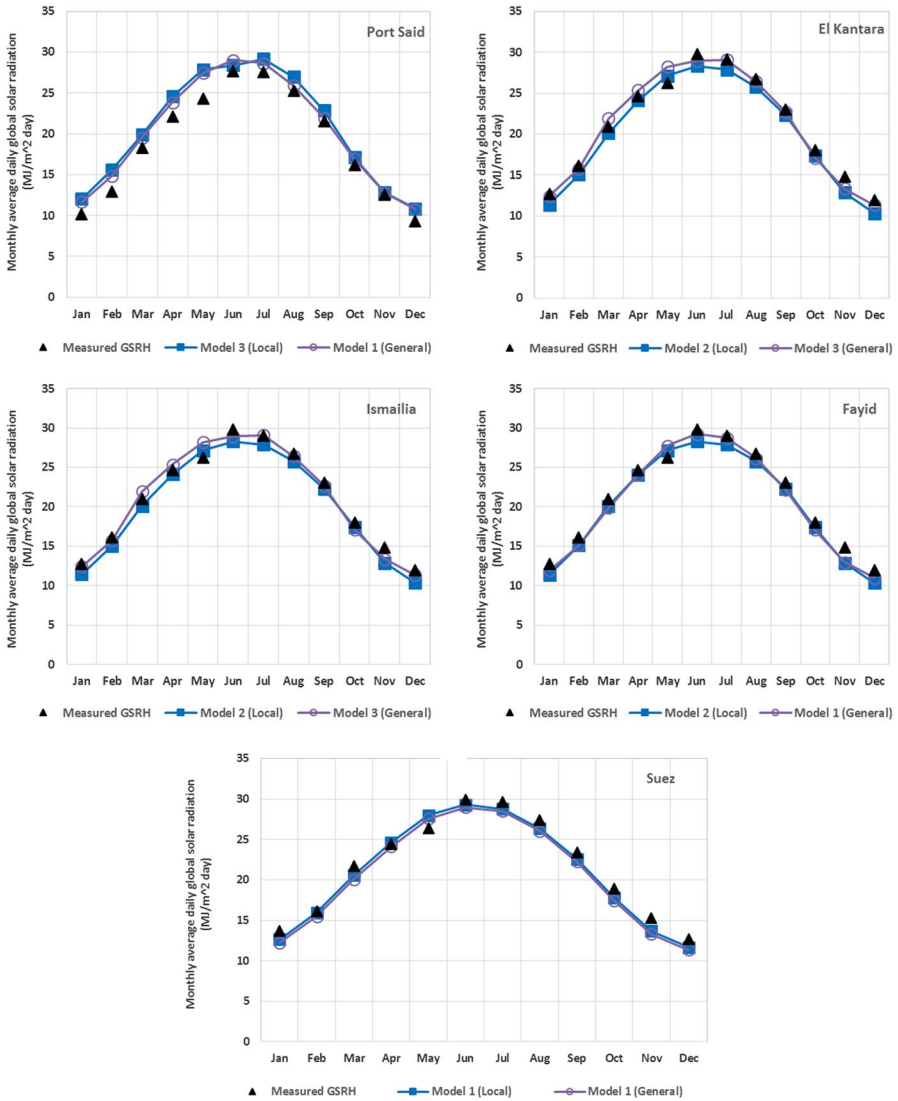
Fig. 9 Prediction of the best local models and the best general models using validation Data Set 2 (2018)

declined by about 1% (from 91 to 90% for Validation Data Set 1 and Validation Data Set 1, respectively), the performance of Model 3 is almost stable. Additionally, the MPE values for Validation Data Set 2 (2018) are just over 10%. For Kantara, Ismailia, and Fayid, although the performance of Model 3 decreased by about 2%, the performance of Model 1
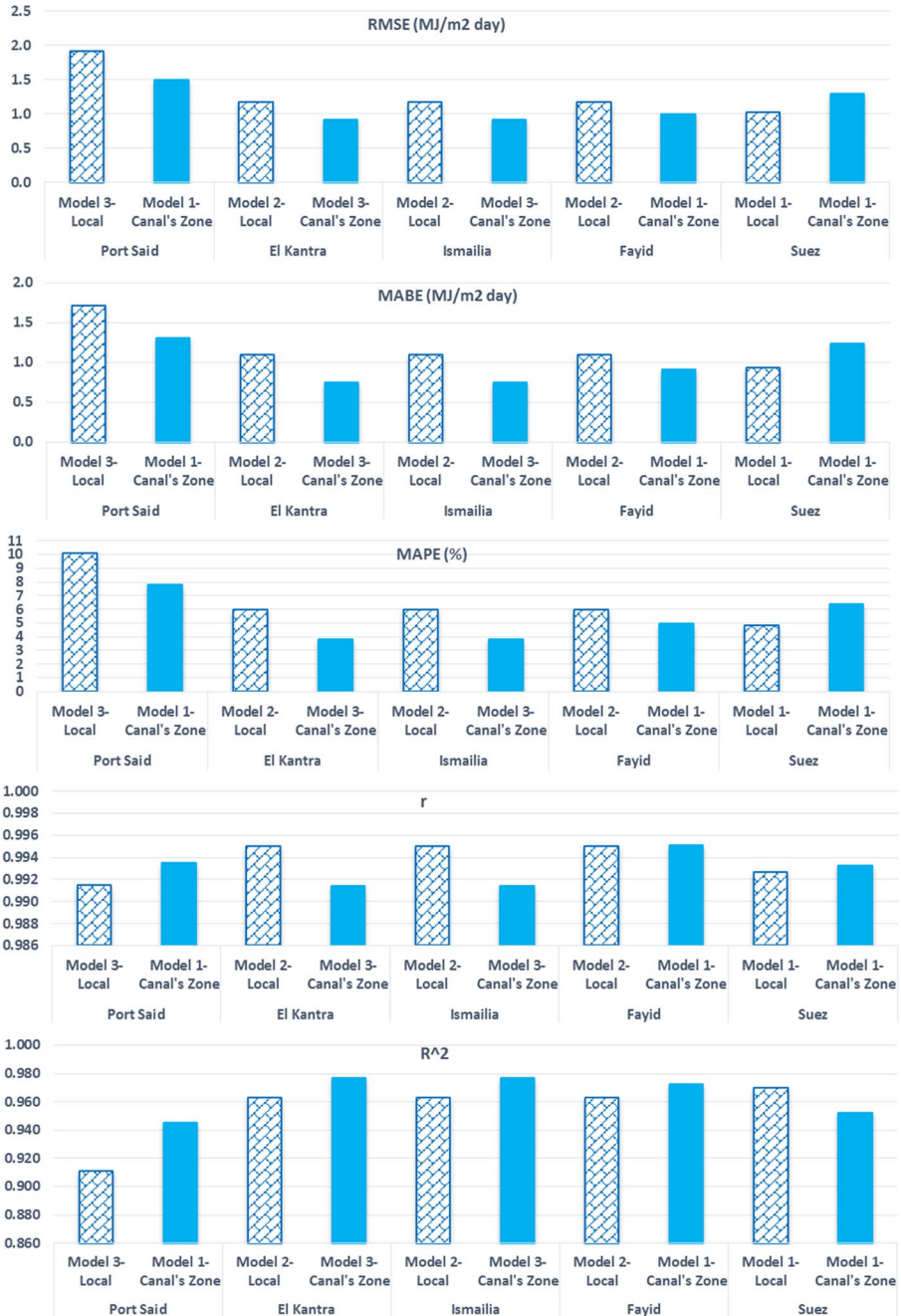
Fig. 10 Errors comparison for the best local and general models using validation Data Set 2 (2018)

Table 8 Relative errors for the best local and general models using validation Data Set 2 (2018)

Type	Month	Port Said (Model 3)	El Kantara (Model 2)	Ismailia (Model 2)	Fayid (Model 2)	Suez (Model 1)
Local formulas	Jan	19.1	−10.3	−10.3	−10.4	−7.6
	Feb	21.5	−6.5	−6.5	−6.2	−1.4
	Mar	8.6	−4.0	−4.0	−4.0	−5.2
	Apr	11.3	−2.0	−2.0	−2.2	0.8
	May	14.5	3.5	3.5	3.4	6.1
	Jun	2.8	−4.9	−4.9	−5.0	−2.2
	Jul	5.9	−4.0	−4.0	−4.0	−2.8
	Aug	6.7	−3.7	−3.7	−3.7	−3.9
	Sep	6.2	−3.3	−3.3	−3.3	−3.6
	Oct	5.8	−3.5	−3.5	−3.6	−5.9
	Nov	2.3	−13.2	−13.2	−13.2	−10.5
	Dec	16.3	−13.2	−13.2	−13.2	−7.9
Suez Canal's zone formulas		(Model 1)	(Model 3)	(Model 3)	(Model 1)	(Model 1)
	Jan	15.4	−2.2	−2.2	−6.1	−10.6
	Feb	15.3	−2.0	−2.0	−6.2	−4.1
	Mar	6.9	4.7	4.7	−5.1	−7.5
	Apr	7.9	3.1	3.1	−2.1	−1.4
	May	12.7	7.5	7.5	5.8	4.5
	Jun	4.9	−2.9	−2.9	−1.8	−3.4
	Jul	4.0	0.1	0.1	−1.0	−3.8
	Aug	2.6	−1.3	−1.3	−2.1	−5.1
	Sep	1.9	−1.5	−1.5	−3.9	−5.1
	Oct	4.8	−5.2	−5.2	−5.3	−7.8
	Nov	2.0	−10.1	−10.1	−12.6	−12.8
	Dec	15.6	−4.8	−4.8	−7.7	−10.7

and 2 is approximately the same, with $R^2 > 96\%$. As for Suez City, all models' performance falls somewhat with about 1%, which is normal at the coastal places.

On the other side, the general formulas of the two validation data sets, for Port Said city, Model 1 and 2 showed the best performance, which marginally slipped by about 1% from 95 to 94%. Conversely, the general formulas of Model 3, for both validation data sets, are dramatically collapsed with R^2 values around 77%. At Kantara and Ismailia, whereas the performance of Model 2 was almost unchanged with R^2 equals 97%, the performance of Model 1 and 3 had very little change with about 1%, from 98 to 97%. At Fayid City, the performance of Models 1 and 2 did not change (R^2 is 97%) while the performance of Model 3 increased from 96 to 97%. For Suez City, Model 1 and Model 2 almost have the same performance, where R^2 is 95% and 94% for the two validation data sets. Also, the performance of Model 3 slightly falls from 94 to 93%, about 1%.

Generally, according to the obtained results of the two validation data sets comparison (Data Set 1: 2018–2020, and Data Set 2: 2018) for both the local and general
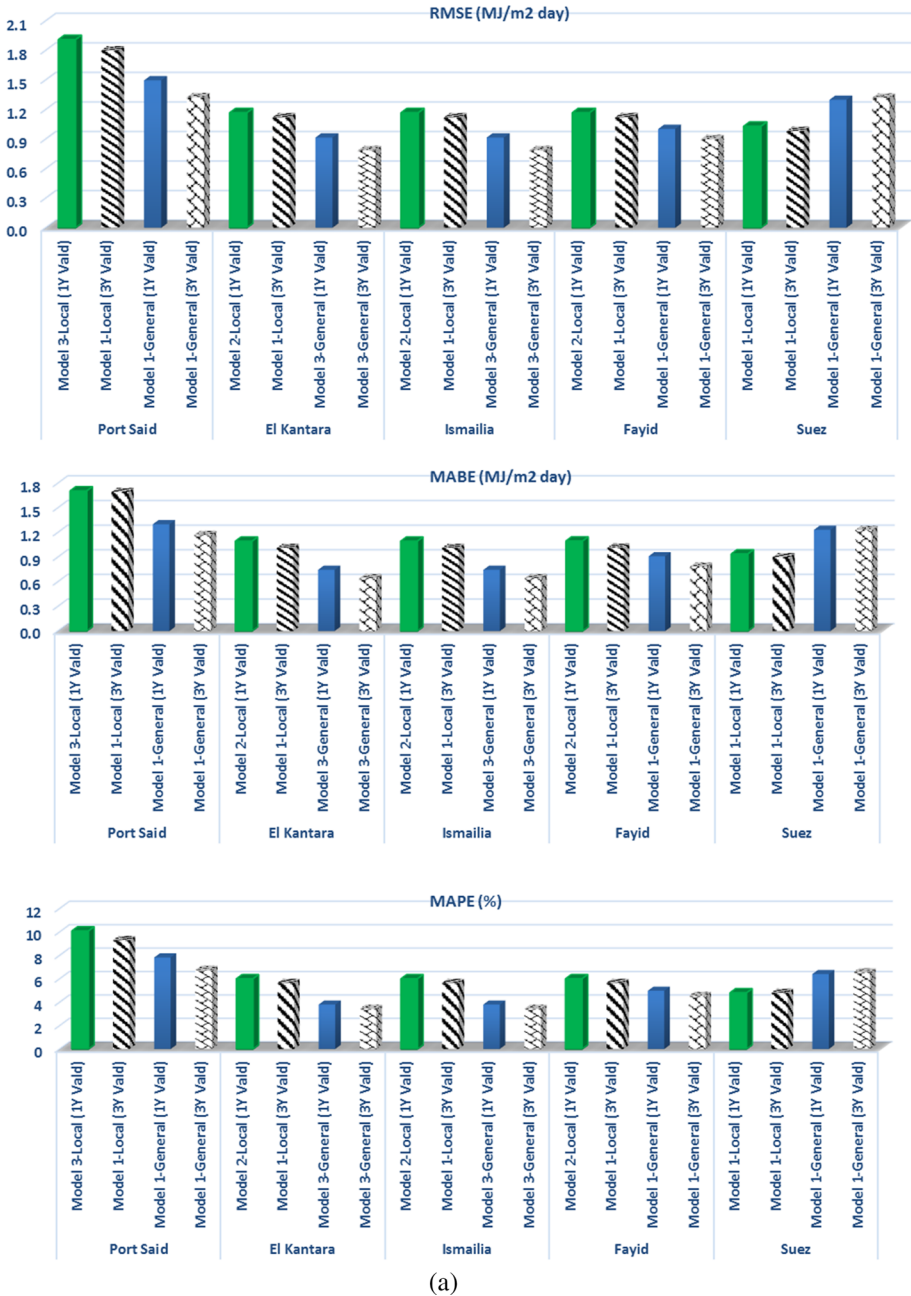
Table 9 Statistical error comparison for all best models (locals and the generals) using the Validation Data Set 1 (2018–2020) and Validation Data Set 2 (2018)

Site	Model type	MPE	MBE	RMSE	MAPE	MABE	r	R^2
Port Said	Model 3-Local (1Y Vald)	10.0890	1.7060	1.9097	10.0890	1.7060	0.9915	0.9110
	Model 1-Local (3Y Vald)	9.3092	1.6925	1.7998	9.309	1.6925	0.9975	0.9196
	Model 1-General (1Y Vald)	7.8213	1.3000	1.4965	7.8213	1.3000	0.9935	0.9454
	Model 1-General (3Y Vald)	6.7562	1.1688	1.3253	6.7562	1.1688	0.9954	0.9564
El Kantara	Model 2-Local (1Y Vald)	−5.4258	−0.9423	1.1675	6.0035	1.0939	0.9950	0.9629
	Model 1-Local (3Y Vald)	−5.6514	−1.0118	1.1207	5.6514	1.0118	0.9979	0.9664
	Model 3-General (1Y Vald)	−1.2140	−0.1170	0.9155	3.7969	0.7451	0.9914	0.9772
	Model 3-General (3Y Vald)	−0.6820	0.0126	0.7863	3.4319	0.6430	0.9949	0.9835
Ismailia	Model 2-Local (1Y Vald)	−5.4258	−0.9423	1.1675	6.0035	1.0939	0.9950	0.9629
	Model 1-Local (3Y Vald)	−5.6512	−1.0118	1.1207	5.6512	1.0118	0.9979	0.9664
	Model 3-General (1Y Vald)	−1.2140	−0.1170	0.9155	3.7969	0.7451	0.9914	0.9772
	Model 3-General (3Y Vald)	−0.6820	0.0126	0.7863	3.4319	0.6430	0.9949	0.9835
Fayid	Model 2-Local (1Y Vald)	−5.4382	−0.9472	1.1678	6.0051	1.0959	0.9951	0.9629
	Model 1-Local (3Y Vald)	−5.6506	−1.0151	1.1219	5.6506	1.0151	0.9979	0.9664
	Model 1-General (1Y Vald)	−4.0157	−0.6582	1.0025	4.9762	0.9102	0.9951	0.9727
	Model 1-General (3Y Vald)	−4.3326	−0.7279	0.8988	4.5357	0.7832	0.9981	0.9784
Suez	Model 1-Local (1Y Vald)	−3.6626	−0.6399	1.0303	4.8051	0.9390	0.9927	0.9698
	Model 1-Local (3Y Vald)	−4.4699	−0.8082	0.9823	4.8015	0.8989	0.9971	0.9727
	Model 1-General (1Y Vald)	−5.6500	−1.0350	1.2984	6.3978	1.2324	0.9932	0.9520
	Model 1-General (3Y Vald)	−6.4860	−1.2122	1.3194	6.5500	1.2297	0.9976	0.9507

Table 10 Relative error comparison for all best models (locals and the generals) using the validation Data Set 1 (2018–2020) and validation Data Set 2 (2018)

Site	Model type	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Port Said	Model 3-local (1Y Vald)	19.1	21.5	8.6	11.3	14.5	2.8	5.9	6.7	6.2	5.8	2.3	16.3
	Model 1-Local (3Y Vald)	12.2	13.5	11.7	11.7	10.9	7.6	7.7	6.3	5.4	6.7	4.4	13.5
	Model 1-General (1Y Vald)	15.4	15.3	6.9	7.9	12.7	4.9	4.0	2.6	1.9	4.8	2.0	15.6
	Model 1-General (3Y Vald)	10.7	12.1	10.4	10.2	8.4	4.3	3.9	2.4	1.8	3.6	1.9	11.3
El Kantara	Model 2-Local (1Y Vald)	−10.3	−6.5	−4.0	−2.0	3.5	−4.9	−4.0	−3.7	−3.3	−3.5	−13.2	−13.2
	Model 1-Local (3Y Vald)	−9.9	−5.8	−3.0	−1.8	−0.2	−3.6	−2.6	−4.0	−5.9	−7.6	−13.3	−10.0
	Model 3-General (1Y Vald)	−2.2	−2.0	4.7	3.1	7.5	−2.9	0.1	−1.3	−1.5	−5.2	−10.1	−4.8
	Model 3-General (3Y Vald)	−6.0	0.8	5.4	2.8	5.8	−1.2	1.2	0.5	−1.0	−2.9	−9.1	−4.5
Ismailia	Model 2-Local (1Y Vald)	−10.3	−6.5	−4.0	−2.0	3.5	−4.9	−4.0	−3.7	−3.3	−3.5	−13.2	−13.2
	Model 1-Local (3Y Vald)	−9.9	−5.8	−3.0	−1.8	−0.2	−3.6	−2.6	−4.0	−5.9	−7.6	−13.3	−10.0
	Model 3-General (1Y Vald)	−2.2	−2.0	4.7	3.1	7.5	−2.9	0.1	−1.3	−1.5	−5.2	−10.1	−4.8
	Model 3-General (3Y Vald)	−6.0	0.8	5.4	2.8	5.8	−1.2	1.2	0.5	−1.0	−2.9	−9.1	−4.5
Fayid	Model 2-Local (1Y Vald)	−10.4	−6.2	−4.0	−2.2	3.4	−5.0	−4.0	−3.7	−3.3	−3.6	−13.2	−13.2
	Model 1-Local (3Y Vald)	−9.8	−5.8	−3.1	−1.9	−0.2	−3.7	−2.7	−4.0	−6.0	−7.6	−13.2	−9.8
	Model 1-General (1Y Vald)	−6.1	−6.2	−5.1	−2.1	5.8	−1.8	−1.0	−2.1	−3.9	−5.3	−12.6	−7.7
	Model 1-General (3Y Vald)	−8.7	−4.6	−2.0	−0.7	1.2	−2.2	−1.1	−2.5	−4.5	−6.2	−12.0	−8.6
Suez	Model 1-Local (1Y Vald)	−7.6	−1.4	−5.2	0.8	6.1	−2.2	−2.8	−3.9	−3.6	−5.9	−10.5	−7.9
	Model 1-Local (3Y Vald)	−8.2	−3.8	−4.4	0.0	2.0	−2.2	−2.5	−4.0	−4.4	−6.4	−10.6	−9.1
	Model 1-General (1Y Vald)	−10.6	−4.1	−7.5	−1.4	4.5	−3.4	−3.8	−5.1	−5.1	−7.8	−12.8	−10.7
	Model 1-General (3Y Vald)	−11.1	−6.7	−7.0	−2.4	0.4	−3.5	−3.6	−5.1	−5.8	−8.3	−12.9	−11.8

Fig. 11 **a** Statistical errors' comparison for all best models, locals and generals, using both validation data sets, Data Set 1 (2018–2020) and Data Set 2 (2018). **b** Statistical errors' comparison for all best models, locals and generals, using both validation data sets, Data Set 1 (2018–2020) and Data Set 2 (2018)
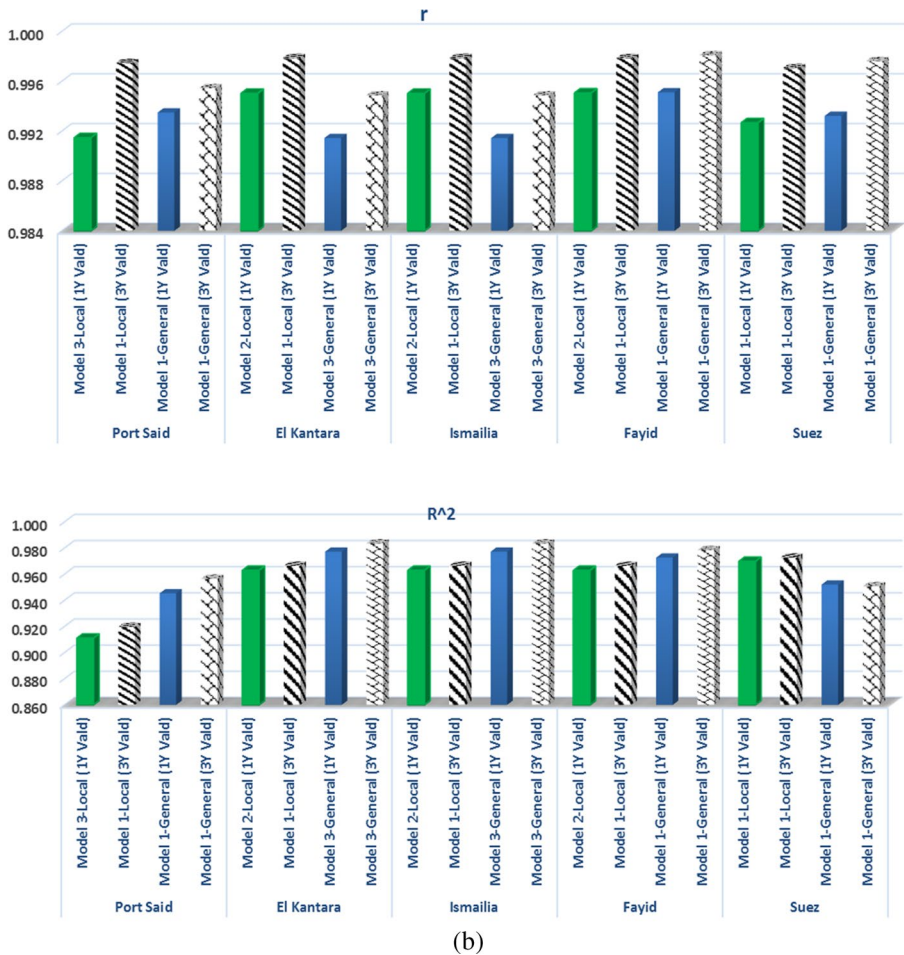
Fig. 11 (continued)

models, it is worthwhile to note that firstly, the length of the validation data set does not have a big effect on models' performance. Actually, the change in models' accuracy is very slight, from about 1% to 2%. Furthermore, Model 1 can be considered the most stable, accurate, and reliable model at different sites, where its performance is located between 95 and 98% except for Port Said City, which is from 91 to 95%. Similarly, the relative error value is compared for the two validation data sets, and it is noticed that while it is increased for some months, especially winter ones, at some locations like Port Said city, it is also decreased at other places such as Suez city.

Of more interest, the best locals' and generals' formulas for the Validation Data Set 1 (2018–2020) and the Validation Data Set 2 (2018) are compared together. Table 9 shows the statistical indicators comparison, while Table 10 introduces the relative

errors comparison. The results show that all statistical errors are within the range of ± 10%, and the performance of the best models is very close to each other. The formulae of Model 1 is approximately the most accurate one among all models, as noted in Table 9, as well as it is the most repeated one as the best formulae, with perfect R^2 values. Model 1 also is the most stable model and has the most accurate estimation at the majority of locations, even at coastal sites with a slight decline in its performance. Moreover, the length of the validation data set has a slight effect on the models' performance and accuracy. For more illustration, the statistical errors for all best models, locals and generals, using both validation data sets, Data Set 1 (2018–2020) and Data Set 2 (2018), are clarified in Fig. 11a and b. Based on the revealed results from the comparison of the two validation data sets together, as well as the locals' and the generals' formulas, Model 1 (Hassan Model (Hassan et al., 2016a, 2016b), Eq. 1) can be considered the most recommended one to estimate GSR on a horizontal plane at the study's locations (five selected cities and the Suez Canal's Zone).

Based on the predicted solar energy resources for the five cities located in the Suez Canal Economic Zone (SCZONE), the results show good average values of solar radiation around the year. These results are consistent with the fact that SCZONE is located within the global solar belt (between the latitudes of ± 35°) with high solar potential which has still not been effectively exploited. The proposed mathematical model that considered a simple accurate quantitative tool to prove the feasibility of solar energy utilization in this newly developed economic region. The results of this accurate tool support the suggested polices for green and sustainable development in this new economic region. These policies attract a lot of green investments which aim to produce different products with low carbon foot print to satisfy the future and modern international requirements of clean production and environmental-friendly products consistent with global Sustainable Development Goals (SDGs).

In general, the issue of providing preliminary data for global solar radiation in an accurate manner has become easier, as the results indicate that the best models in this study performed well in the GSR forecast. These best models can anticipate GSR in the analyzed areas with high accuracy at any moment, and they can do so more quickly and with more applicability. Their R^2 ranges from 95 to 98% at all sites except for coastal one, which is from 91 to 95%. As a result, they are the ones that are advised to forecast GSR at the investigated areas in particular or throughout the Suez Canal Zone. Additionally, it is suitable for horizontal GSR forecasting at any location, and its empirical coefficients have to be calibrated using local data when it is used in other sites because it is site-dependent. Moreover, by using this developed model, Model 1, forecasting solar radiation may be readily achieved with adequate dependability at a variety of locations when no device for detecting solar radiation is available. Furthermore, the temperature parameter is being captured simply and constantly for other purposes, which is seen to be the key advantage of this strategy in the current study. For more clarification, the graphical presentations of the study's findings are presented in Fig. 15 (in the Appendix). In other words, the developed mathematical models in this study are considered an accurate and simple tool based on the available temperature data to evaluate the available solar energy resources for different locations. This can help the policymakers to identify different opportunities for solar energy utilization

as a part of the green sustainable plan reducing the carbon footprint toward the low or zero-carbon policies for a certain region such as the Suez Canal Economic Zone.

4 Conclusion and future work

Solar energy is believed to be the most promising renewable energy source globally, and as such, accurate data on solar radiation is essential in evaluating solar resources for various applications. These data are considered the first step in achieving energy sustainability goals and satisfying the world's energy demand. However, the unavailability of solar radiation measurements in many regions of the world has necessitated the development of several global solar radiation (GSR) models to predict solar radiation in these locations. In this regard, the present work evaluated the proficiency of several GSR models at five new locations and determined the most suitable one for GSR prediction. The study has further developed solar radiation models for these new locations, as well as general ones for the entire region, which does not have any GSR models regardless of the existence of many proposed solar energy projects in this zone. These sites are Port Said, El Kantara, Ismailia, Fayid, and Suez. Additionally, the study investigates the effect of changing the length of the validation dataset on models' performance and accuracy, as well as assesses the introduced models' generalization capability. The study revealed that Model 1 (Hassan et al. Model), Eq. 1, outperformed other models in terms of stability, accuracy, and reliability at different locations. Its performance, measured by R^2 value, was found to be between 95 and 98%, except for the coastal city (Port Said), where it ranged between 91 and 95%. The study also found that the length of the validation data set had a negligible effect on the models' performance, which ranged from about 1% to 2%. Furthermore, Model 1 showed consistent performance even after a prolonged period, and its accuracy was still highly accurate when compared to its 2016 results. Therefore, Model 1 is an excellent GSR estimation model on a flat plane with high precision, and it can be combined with other long- or short-term weather forecast methods to achieve high applicability. In particular, accurate temperature predictions from these weather forecast methods can be used as inputs to Model 1 for precise GSR forecast. Consequently, Model 1 is the recommended solar radiation model for the studied locations and other sites globally, with an accepted error limit and high precision.

However, because of its slightly low performance at the coastal location, it is planned to investigate and study integrating effective weather parameters such as relative humidity into Model 1 (Eq. 1) and check its effects on the model's prediction and accuracy. Additionally, the study aims to compare the performance of the best empirical model obtained from this study with other techniques such as machine learning methods, including artificial neural networks, at the same studied locations, especially the coastal ones.

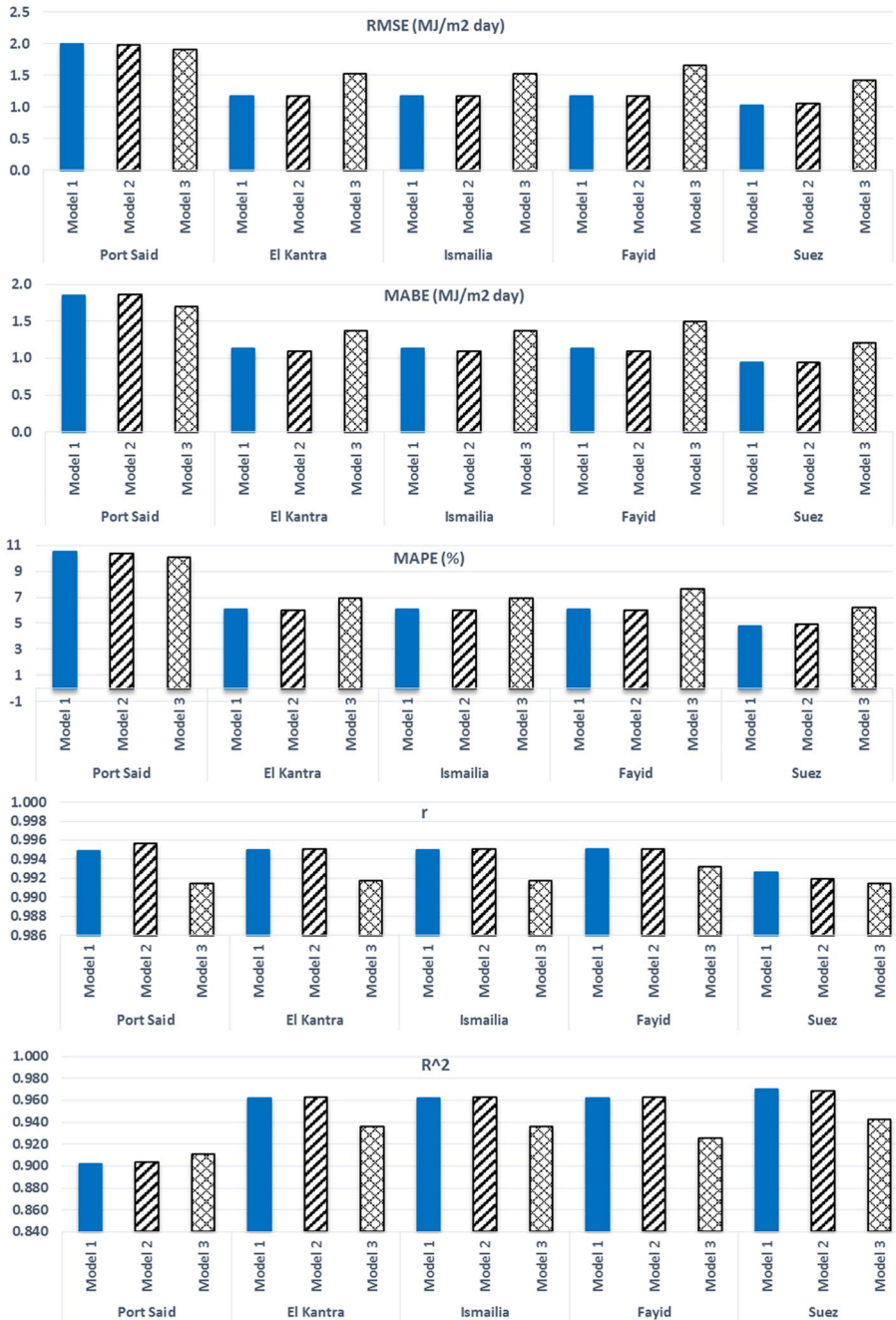## Appendix

See Figs. 12, 13, 14 and 15.

Fig. 12 Statistical errors graph for all models of the local formulas using validation Data Set 2 (2018)
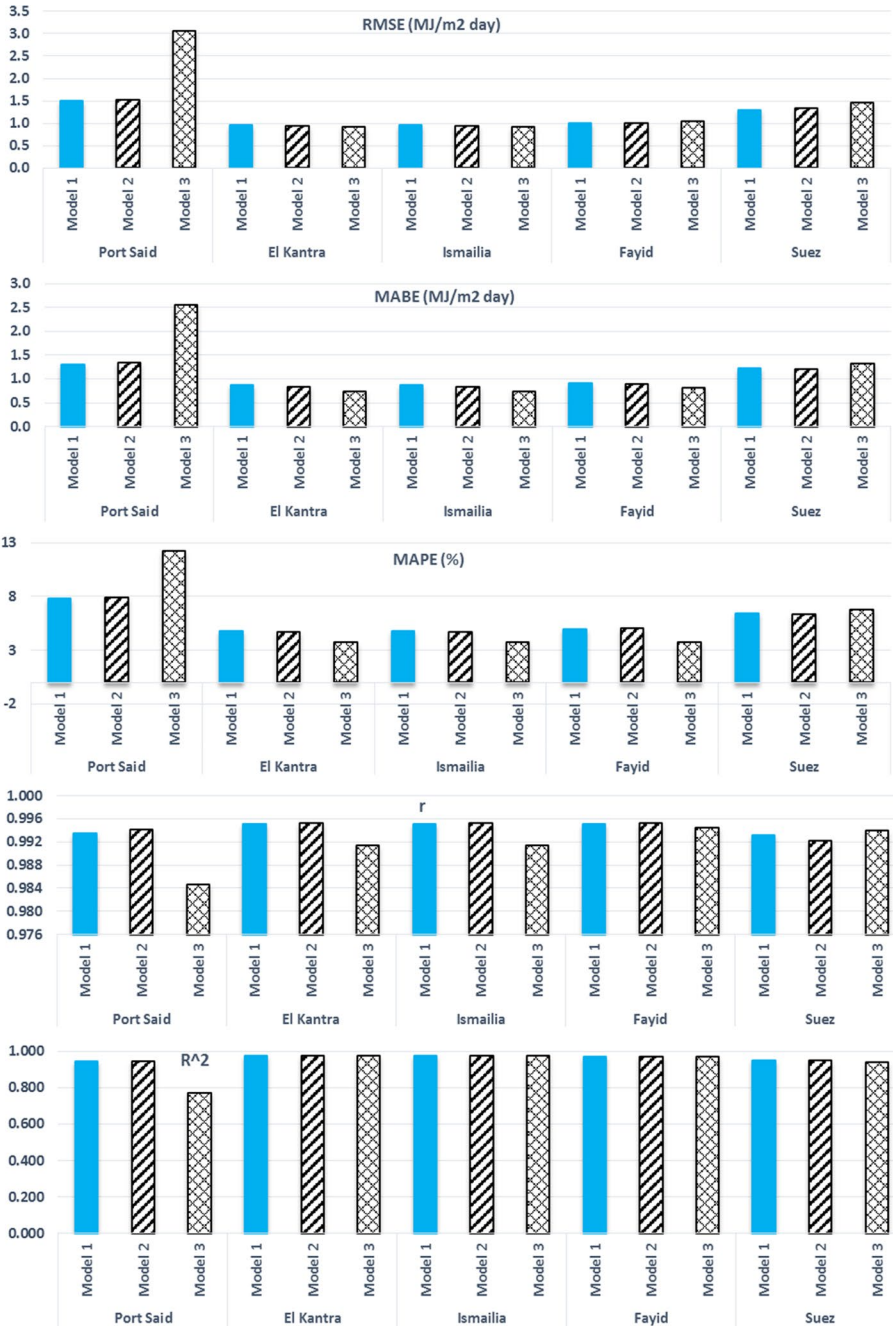
Fig. 13 Statistical errors graph for all models of the general formulas using validation Data Set 2 (2018)
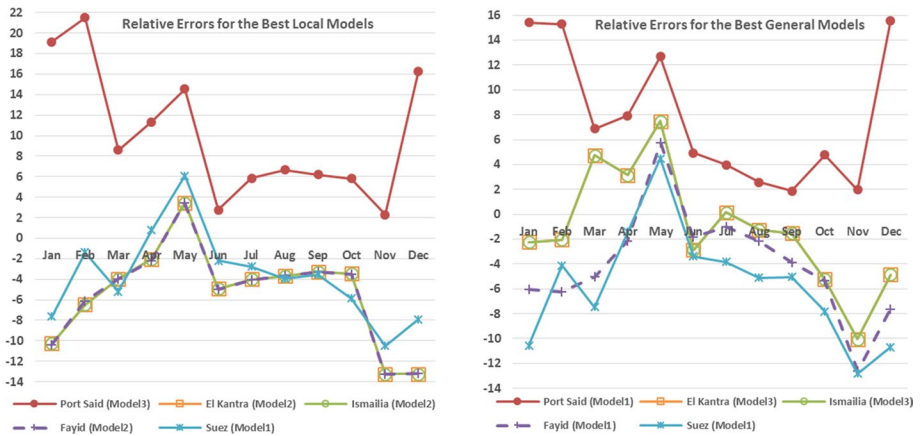
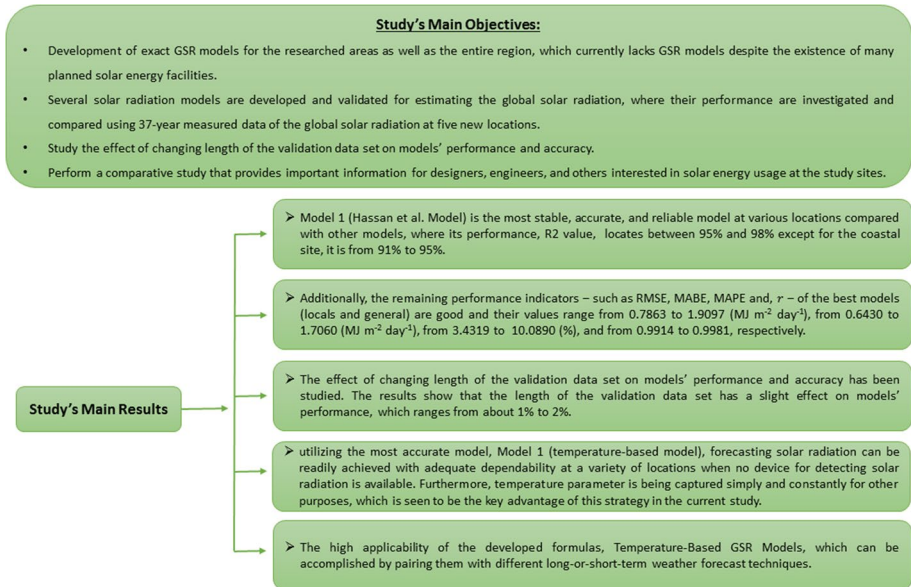**Fig. 14** Relative errors for the best local and general models using validation Data Set 2 (2018)

Fig. 15 Graphical presentations of the study's findings

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abdo, T., & El-Shimy, M. (2013). Estimating the global solar radiation for solar energy projects–Egypt case study. *International Journal of Sustainable Energy, 32*(6), 682–712. https://doi.org/10.1080/14786451.2013.822872

Agarwal, V., Malhotra, S., Dagar, V., & Pavithra, M. R. (2023). Coping with public-private partnership issues: A path forward to sustainable agriculture. *Socio-Economic Planning Sciences, 89*, 101703. https://doi.org/10.1016/j.seps.2023.101703

Ajayi, O. O., Ohijeagbon, O. D., Nwadialo, C. E., & Olasope, O. (2014). New model to estimate daily global solar radiation over Nigeria. *Sustainable Energy Technologies and Assessments, 5*, 28–36. https://doi.org/10.1016/j.seta.2013.11.001

Ali, M. A., Hassan, G. E., & Youssef, M. E. (2016). Assessment the performance of artificial neural networks in estimating global solar radiation. In *International conference on new trends for sustainable energy-ICNTSE* (pp. 148–150).

Ali, Z., Qingmei, B., Hafiz, T., & Kamran, W. (2022). A multi-perspective assessment approach of renewable energy production: Policy perspective analysis. *Environment, Development and Sustainability, 24*(2), 2164–2192. https://doi.org/10.1007/s10668-021-01524-8

Allen, R. G. (1997). Self-calibrating method for estimating solar radiation from air temperature. *Journal of Hydrologic Engineering, 2*(2), 56–67. https://doi.org/10.1061/(ASCE)1084-0699(1997)2:2(56)

Almorox, J., Benito, M., & Hontoria, C. (2005). Estimation of monthly Angström–Prescott equation coefficients from measured daily data in Toledo Spain. *Renewable Energy, 30*(6), 931–936. https://doi.org/10.1016/j.renene.2004.08.002

Almorox, J., Hontoria, C., & Benito, M. (2011). Models for obtaining daily global solar radiation with measured air temperature data in Madrid (Spain). *Applied Energy, 88*(5), 1703–1709. https://doi.org/10.1016/j.apenergy.2010.11.003

Al-Mostafa, Z. A., Maghrabi, A. H., & Al-Shehri, S. M. (2014). Sunshine-based global radiation models: A review and case study. *Energy Conversion and Management, 84*, 209–216. https://doi.org/10.1016/j.enconman.2014.04.021

Alvarado, R., Tillaguango, B., Cuesta, L., Pinzon, S., Alvarado-Lopez, M. R., Işık, C., & Dagar, V. (2022). Biocapacity convergence clubs in Latin America: An analysis of their determining factors using quantile regressions. *Environmental Science and Pollution Research, 29*(44), 66605–66621. https://doi.org/10.1007/s11356-022-20567-6

Angstrom, A. (1924). Solar and terrestrial radiation. *Quarterly Journal of the Royal Meteorological Society, 50*, 121–125.

Annandale, J., Jovanovic, N., Benadé, N., & Allen, R. (2002). Software for missing data error analysis of Penman–Monteith reference evapotranspiration. *Irrigation Science, 21*(2), 57–67. https://doi.org/10.1007/s002710100047

Atkins, N. (n.d.). Daily temperature variations–diurnal variation of solar radiation. Yndon State Collage Survey of Meterology.

Ayodele, T. R., & Ogunjuyigbe, A. S. O. (2015). Prediction of monthly average global solar radiation based on statistical distribution of clearness index. *Energy, 90*, 1733–1742. https://doi.org/10.1016/j.energy.2015.06.137

Babatunde, O. M., Munda, J. L., Hamam, Y., & Monyei, C. G. (2023). A critical overview of the (Im) practicability of solar radiation forecasting models. *e-Prime–advances in Electrical Engineering, Electronics and Energy*. https://doi.org/10.1016/j.prime.2023.100213

Barbaro, S., Coppolino, S., Leone, C., & Sinagra, E. (1978). Global solar radiation in Italy. *Solar Energy, 20*(5), 431–435.

Besharat, F., Dehghan, A. A., & Faghih, A. R. (2013). Empirical models for estimating global solar radiation: A review and case study. *Renewable and Sustainable Energy Reviews, 21*, 798–821. https://doi.org/10.1016/j.rser.2012.12.043

Bounoua, Z., Ouazzani Chahidi, L., & Mechaqrane, A. (2021). Estimation of daily global solar radiation using empirical and machine-learning methods: A case study of five Moroccan locations. *Sustainable Materials and Technologies, 28*, e00261. https://doi.org/10.1016/j.susmat.2021.e00261

Camacho, E., Berenguel, M., Rubio, F., & Martínez, D. (2012). Chapter 1-solar energy fundamentals. *Control of solar energy systems* (pp. 1–23). Springer.

Chanchangi, Y. N., Adu, F., Ghosh, A., Sundaram, S., & Mallick, T. K. (2022). Nigeria's energy review: Focusing on solar energy potential and penetration. *Environment, Development and Sustainability.* https://doi.org/10.1007/s10668-022-02308-4

Chand, V., Arunendra, S., & Vilas, K. T. (2021). Experimental investigations on the seasonal performance variations of directly coupled solar photovoltaic water pumping system using centrifugal pump. *Environment, Development and Sustainability, 23*(6), 8288–8306. https://doi.org/10.1007/s10668-020-00965-x

Chaudhary, R., & Yadav, A. (2021). Experimental investigation of solar cooking system based on evacuated tube solar collector for the preparation of concentrated sugarcane juice used in Jaggery making. *Environment, Development and Sustainability, 23*(1), 647–663. https://doi.org/10.1007/s10668-020-00601-8

Chen, P., & Dagestani, A. A. (2023a). Greenwashing behavior and firm value–from the perspective of board characteristics. *Corporate Social Responsibility and Environmental Management, 30*(5), 2330–2343. https://doi.org/10.1002/csr.2488

Chen, P., & Dagestani, A. A. (2023b). Urban planning policy and clean energy development Harmony–evidence from smart city pilot policy in China. *Renewable Energy, 210*, 251–257. https://doi.org/10.1016/j.renene.2023.04.063

Chen, P., Dagestani, A. A., & Kim, S. K. (2023). Corporate social responsibility and green exploratory innovation - the moderating role of three environmental regulations. *Technology Analysis and Strategic Management.* https://doi.org/10.1080/09537325.2023.2196585

Cheng, S., Fan, Q., & Dagestani, A. A. (2023). Opening the black box between strategic vision on digitalization and SMEs digital transformation: the mediating role of resource orchestration. *Kybernetes.* https://doi.org/10.1108/K-01-2023-0073

Cihan, P., Ozel, H., & Ozcan, H. K. (2021). Modeling of atmospheric particulate matters via artificial intelligence methods. *Environmental Monitoring and Assessment, 193*(5), 1–15. https://doi.org/10.1007/s10661-021-09091-1

Dagar, V., & Malik, S. (2023). Nexus between macroeconomic uncertainty, oil prices, and exports: Evidence from quantile-on-quantile regression approach. *Environmental Science and Pollution Research, 30*(16), 48363–48374. https://doi.org/10.1007/s11356-023-25574-9

Dagestani, A. A. (2022). An analysis of the impacts of COVID-19 and Freight cost on trade of the economic belt and the maritime silk road. *International Journal of Industrial Engineering and Production Research, 33*(3), 1–16. https://doi.org/10.22068/ijiepr.33.3.7

Dagestani, A. A., & Qing, L. (2022). The impact of environmental information disclosure on Chinese firms' environmental and economic performance in the 21st century: A systematic review. *IEEE Engineering Management Review, 50*(4), 203–214. https://doi.org/10.1109/EMR.2022.3210465

Dagestani, A. A., Qing, L., & Abou Houran, M. (2022). What remains unsolved in sub-african environmental exposure information disclosure: A review. *Journal of Risk and Financial Management, 15*(10), 487. https://doi.org/10.3390/jrfm15100487

Dagestani, A. A., Shang, Y., Schneider, N., Cifuentes-Faura, J., & Zhao, X. (2023). Porter in China: A quasi-experimental view of market-based environmental regulation effects on firm performance. *Energy Economics, 126*, 106966. https://doi.org/10.1016/j.eneco.2023.106966

Dagher, L. (2012). Natural gas demand at the utility level: An application of dynamic elasticities. *Energy Economics, 34*(4), 961–969. https://doi.org/10.1016/j.eneco.2011.05.010

Dagher, L., & El Hariri, S. (2013). The impact of global oil price shocks on the Lebanese stock market. *Energy, 63*(2013), 366–374. https://doi.org/10.1016/j.energy.2013.10.012

Dagher, L., & Ruble, I. (2010). Challenges for CO_2 mitigation in the Lebanese electric-power sector. *Energy Policy, 38*(2), 912–918. https://doi.org/10.1016/j.enpol.2009.10.042

Dagher, L., & Ruble, I. (2011). Modeling Lebanon's electricity sector: Alternative scenarios and their implications. *Energy, 36*(7), 4315–4326. https://doi.org/10.1016/j.energy.2011.04.010

Dagher, L., & Yacoubian, T. (2012). The causal relationship between energy consumption and economic growth in Lebanon. *Energy Policy, 50*(2012), 795–801. https://doi.org/10.1016/j.enpol.2012.08.034

Despotovic, M., Nedic, V., Despotovic, D., & Cvetanovic, S. (2015). Review and statistical analysis of different global solar radiation sunshine models. *Renewable and Sustainable Energy Reviews, 52*, 1869–1880. https://doi.org/10.1016/j.rser.2015.08.035

Dincer, I., Dilmac, S., Ture, I. E., & Edin, M. (1996). A simple technique for estimating solar radiation parameters and its application for Gebze. *Energy Conversion and Management, 37*(2), 183–198. https://doi.org/10.1016/0196-8904(95)00168-D

Doorga, J. R. S., Rughooputh, S. D. D. V., & Boojhawon, R. (2019). Modelling the global solar radiation climate of Mauritius using regression techniques. *Renewable Energy, 131*, 861–878. https://doi.org/10.1016/j.renene.2018.07.107

Eesley, C., Li, J. B., & Yang, D. (2016). Does institutional change in universities influence high-tech entrepreneurship? Evidence from China's project. *Oranization Sciene, 27*(2), 446–461. https://doi.org/10.1287/orsc.2015.1038

El Mghouchi, Y., El Bouardi, A., Sadouk, A., Fellak, I., & Ajzoul, T. (2016). Comparison of three solar radiation models and their validation under all sky conditions–case study: Tetuan city in northern of Morocco. *Renewable and Sustainable Energy Reviews, 58*, 1432–1444. https://doi.org/10.1016/j.rser.2015.12.354

El-Metwally, M. (2004). Simple new methods to estimate global solar radiation based on meteorological data in Egypt. *Atmospheric Research, 69*(3–4), 217–239. https://doi.org/10.1016/j.atmosres.2003.09.002

El-Metwally, M. (2005). Sunshine and global solar radiation estimation at different sites in Egypt. *Journal of Atmospheric and Solar-Terrestrial Physics, 67*(14), 1331–1342. https://doi.org/10.1016/j.jastp.2005.04.004

Eroğlu, H. (2022). Development of a novel solar energy need index for identifying priority investment regions : A case study and current status in Turkey. *Environment, Development and Sustainability, 24*(6), 8840–8855. https://doi.org/10.1007/s10668-021-01812-3

Ertekin, C., & Yaldiz, O. (1999). Estimation of monthly average daily global radiation on horizontal surface for Antalya (Turkey). *Renewable Energy, 17*(1), 95–102. https://doi.org/10.1016/S0960-1481(98)00109-8

Fadare, D. A. (2009). Modelling of solar energy potential in Nigeria using an artificial neural network model. *Applied Energy, 86*(9), 1410–1422. https://doi.org/10.1016/j.apenergy.2008.12.005

Feng, Y., Gong, D., Zhang, Q., Jiang, S., Zhao, L., & Cui, N. (2019). Evaluation of temperature-based machine learning and empirical models for predicting daily global solar radiation. *Energy Conversion and Management, 198*, 111780. https://doi.org/10.1016/j.enconman.2019.111780

Guermoui, M., Melgani, F., Gairaa, K., & Mekhalfi, M. L. (2020). A comprehensive review of hybrid models for solar radiation forecasting. *Journal of Cleaner Production, 258*, 120357. https://doi.org/10.1016/j.jclepro.2020.120357

Guru, S., Verma, S., Baheti, P., & Dagar, V. (2023). Assessing the feasibility of hyperlocal delivery model as an effective distribution channel. *Management Decision, 61*(6), 1634–1655. https://doi.org/10.1108/MD-03-2022-0407

Hai, T., Sharafati, A., Mohammed, A., Salih, S. Q., Deo, R. C., Al-Ansari, N., & Yaseen, Z. M. (2020). Global solar radiation estimation and climatic variability analysis using extreme learning machine based predictive model. *IEEE Access, 8*, 12026–12042. https://doi.org/10.1109/ACCESS.2020.2965303

Hargreaves, G. H., & Samani, Z. A. (1982). Estimating potential evapotranspiration. *Journal of the Irrigation and Drainage Division, 108*(3), 225–230.

Hassan, G. E., Youssef, M. E., Ali, M. A., Mohamed, Z. E., & Hanafy, A. A. (2018). Evaluation of different sunshine-based models for predicting global solar radiation–case study: New Borg El-Arab City Egypt. *Thermal Science, 22*(2), 979–992. https://doi.org/10.2298/TSCI160803085H

Hassan, G. E., Youssef, M. E., Ali, M. A., Mohamed, Z. E., & Shehata, A. I. (2016a). Performance assessment of different day-of-the-year-based models for estimating global solar radiation–case study: Egypt. *Journal of Atmospheric and Solar-Terrestrial Physics, 149*, 69–80. https://doi.org/10.1016/j.jastp.2016.09.011

Hassan, G. E., Youssef, M. E., Mohamed, Z. E., Ali, M. A., & Hanafy, A. A. (2016b). New temperature-based models for predicting global solar radiation. *Applied Energy, 179*, 437–450. https://doi.org/10.1016/j.apenergy.2016.07.006

He, C., Liu, J., Xu, F., Zhang, T., Chen, S., Sun, Z., et al. (2020). Improving solar radiation estimation in China based on regional optimal combination of meteorological factors with machine learning methods. *Energy Conversion and Management, 220*, 113111. https://doi.org/10.1016/j.enconman.2020.113111

I-cias. (n.d.). Suez canal map. https://www.cruiserswiki.org/wiki/Suez_Canal

Jiang, Y. (2009). Estimation of monthly mean daily diffuse radiation in China. *Applied Energy, 86*(9), 1458–1464. https://doi.org/10.1016/j.apenergy.2009.01.002

Karakoti, I., Das, P. K., & Singh, S. K. (2012). Predicting monthly mean daily diffuse radiation for India. *Applied Energy, 91*(1), 412–425. https://doi.org/10.1016/j.apenergy.2011.10.012

Karaman, Ö. A., Tanyıldızı Ağır, T., & Arsel, İ. (2021). Estimation of solar radiation using modern methods. *Alexandria Engineering Journal, 60*(2), 2447–2455. https://doi.org/10.1016/j.aej.2020.12.048

Khalil, S. A., & Shaffie, A. M. (2013a). Performance of statistical comparison models of solar energy on horizontal and inclined surface. *International Journal of Energy and Power, 2*(1), 8–25.

Khalil, S. A., & Shaffie, A. M. (2013b). A comparative study of total, direct and diffuse solar irradiance by using different models on horizontal and inclined surfaces for Cairo Egypt. *Renewable and Sustainable Energy Reviews, 27*, 853–863. https://doi.org/10.1016/j.rser.2013.06.038

Khorasanizadeh, H., & Mohammadi, K. (2013a). Prediction of daily global solar radiation by day of the year in four cities located in the sunny regions of Iran. *Energy Conversion and Management, 76*, 385–392. https://doi.org/10.1016/j.enconman.2013.07.073

Khorasanizadeh, H., & Mohammadi, K. (2013b). Introducing the best model for predicting the monthly mean global solar radiation over six major cities of Iran. *Energy, 51*, 257–266. https://doi.org/10.1016/j.energy.2012.11.007

Kumar, P., Gupta, S., & Dagar, V. (2023). Sustainable energy development through non-residential rooftop solar photovoltaic adoption: Empirical evidence from India. *Sustainable Development*. https://doi.org/10.1002/sd.2644

Kumar, R., Aggarwal, R. K., & Sharma, J. D. (2015). Comparison of regression and artificial neural network models for estimation of global solar radiations. *Renewable and Sustainable Energy Reviews, 52*, 1294–1299. https://doi.org/10.1016/j.rser.2015.08.021

Li, H., Cao, F., Wang, X., & Ma, W. (2014). A temperature-based model for estimating monthly average daily global solar radiation in China. *The Scientific World Journal, 2014*, 128754.

Li, H., Ma, W., Lian, Y., & Wang, X. (2010). Estimating daily global solar radiation by day of year in China. *Applied Energy, 87*(10), 3011–3017. https://doi.org/10.1016/j.apenergy.2010.03.028

Makade, R. G., Chakrabarti, S., & Jamil, B. (2021). Development of global solar radiation models: A comprehensive review and statistical analysis for Indian regions. *Journal of Cleaner Production, 293*, 126208. https://doi.org/10.1016/j.jclepro.2021.126208

Mawire, A., Simelane, S. M., & Abedigamba, P. O. (2021). Energetic and exergetic performance comparison of three solar cookers for developing countries. *Environment, Development and Sustainability, 23*(10), 14528–14555. https://doi.org/10.1007/s10668-021-01255-w

Mecibah, M. S., Boukelia, T. E., Tahtah, R., & Gairaa, K. (2014). Introducing the best model for estimation the monthly mean daily global solar radiation on a horizontal surface (Case study: Algeria). *Renewable and Sustainable Energy Reviews, 36*, 194–202. https://doi.org/10.1016/j.rser.2014.04.054

NASA Surface Meteorology and Solar Energy. (n.d.). These data were obtained from the NASA langley research center (LaRC) POWER project funded through the NASA earth science/applied science program. https://power.larc.nasa.gov/data-access-viewer/

NASA. (n.d.). Accuracy of NASA surface meteorology and solar energy data. https://www.pvsyst.com/help/meteo_source_nasa.htm and https://power.larc.nasa.gov/docs/methodology/ and https://power.larc.nasa.gov/

Nwokolo, S. C., & Ogbulezie, J. C. (2018). A qualitative review of empirical models for estimating diffuse solar radiation from experimental data in Africa. *Renewable and Sustainable Energy Reviews, 92*(2017), 353–393. https://doi.org/10.1016/j.rser.2018.04.118

Piwowar-Sulej, K., Malik, S., Shobande, O. A., Singh, S., & Dagar, V. (2023). A contribution to sustainable human resource development in the era of the COVID-19 pandemic. *Journal of Business Ethics*. https://doi.org/10.1007/s10551-023-05456-3

Prescott, J. A. (1940). Evaporation from water surface in relation to solar radiation. *Transactions of the Royal Society of Australia, 46*, 114–118.

Quej, V. H., Almorox, J., Ibrakhimov, M., & Saito, L. (2016). Empirical models for estimating daily global solar radiation in Yucatán Peninsula, Mexico. *Energy Conversion and Management, 110*, 448–456. https://doi.org/10.1016/j.enconman.2015.12.050

Rao, A., Dagar, V., Sohag, K., Dagher, L., & Tanin, T. I. (2023). Good for the planet, good for the wallet: The ESG impact on financial performance in India. *Finance Research Letters, 56*, 104093. https://doi.org/10.1016/j.frl.2023.104093

Robaa, S. M. (2009). Validation of the existing models for estimating global solar radiation over Egypt. *Energy Conversion and Management, 50*(1), 184–193. https://doi.org/10.1016/j.enconman.2008.07.005

Saleh, A., Mohammad, A., Naim, F., Mohd, T., & Adzman, R. (2021). Impacts of albedo and atmospheric conditions on the efficiency of solar energy: A case study in temperate climate of Choman, Iraq. *Environment, Development and Sustainability, 23*(1), 989–1018. https://doi.org/10.1007/s10668-019-00568-1

Samuel Chukwujindu, N. (2017). A comprehensive review of empirical models for estimating global solar radiation in Africa. *Renewable and Sustainable Energy Reviews, 78*, 955–995. https://doi.org/10.1016/j.rser.2017.04.101

Shen, B., Yang, X., Xu, Y., Ge, W., Liu, G., Su, X., et al. (2023). Can carbon emission trading pilot policy drive industrial structure low-carbon restructuring: New evidence from China. *Environmental Science and Pollution Research, 30*(14), 41553–41569. https://doi.org/10.1007/s11356-023-25169-4

Su, D., Xu, S., & Tong, Z. (2023). Credit availability and corporate risk-taking: Evidence from China's green credit policy. *Post-Communist Economies, 35*(3), 236–270. https://doi.org/10.1080/14631377.2023.2169516

Tao, H., Ewees, A. A., Al-Sulttani, A. O., Beyaztas, U., Hameed, M. M., Salih, S. Q., et al. (2021). Global solar radiation prediction over North Dakota using air temperature: Development of novel hybrid intelligence model. *Energy Reports, 7*, 136–157. https://doi.org/10.1016/j.egyr.2020.11.033

Tao, M., Dagestani, A. A., Goh, L. T., Zheng, Y., & Le, W. (2023). Do China's anti-corruption efforts improve corporate productivity? A difference-in-difference exploration of Chinese listed enterprises. *Socio-Economic Planning Sciences, 87*, 101594. https://doi.org/10.1016/j.seps.2023.101594

Tlili, I. (2015). Renewable energy in Saudi Arabia: current status and future potentials. *Environment, Development and Sustainability, 17*, 859–886. https://doi.org/10.1007/s10668-014-9579-9

Türk Togrul, I., & Onat, E. (1999). Study for estimating solar radiation in Elazig using geographical and meteorological data. *Energy Conversion and Management, 40*(14), 1577–1584. https://doi.org/10.1016/S0196-8904(99)00035-7

Yıldırım, H. B., Çelik, Ö., Teke, A., & Barutçu, B. (2018). Estimating daily global solar radiation with graphical user interface in eastern mediterranean region of Turkey. *Renewable and Sustainable Energy Reviews, 82*(2017), 1528–1537. https://doi.org/10.1016/j.rser.2017.06.030

Yorukoglu, M., & Celik, A. N. (2006). A critical review on the estimation of daily global solar radiation from sunshine duration. *Energy Conversion and Management, 47*(15–16), 2441–2450. https://doi.org/10.1016/j.enconman.2005.11.002

You, G., Gan, S., Guo, H., & Dagestani, A. A. (2022). Public opinion spread and guidance strategy under COVID-19: A SIS model analysis. *Axioms, 11*(6), 296. https://doi.org/10.3390/axioms11060296

Youssef, M. E., Hassan, G. E., Mohamed, Z. E., & Ali, M. A. (2016). Investigating the performance of different models in estimating global solar radiation. *Advances in Natural and Applied Sciences, 10*(4), 379–389.

Zaaoumi, A., Bah, A., Ciocan, M., Sebastian, P., Balan, M. C., Mechaqrane, A., & Alaoui, M. (2021). Estimation of the energy production of a parabolic trough solar thermal power plant using analytical and artificial neural networks models. *Renewable Energy, 170*, 620–638. https://doi.org/10.1016/j.renene.2021.01.129

Zhang, S., Shinwari, R., Zhao, S., & Dagestani, A. A. (2023). Energy transition, geopolitical risk, and natural resources extraction: A novel perspective of energy transition and resources extraction. *Resources Policy, 83*, 103608. https://doi.org/10.1016/j.resourpol.2023.103608

Zhao, S., Tian, W., & Dagestani, A. A. (2022). How do R&D factors affect total factor productivity: based on stochastic frontier analysis method. *Economic Analysis Letters, 1*(2), 28–34. https://doi.org/10.58567/eal01020005

Zhou, H., Awosusi, A. A., Dagar, V., Zhu, G., & Abbas, S. (2023). Unleashing the asymmetric effect of natural resources abundance on carbon emissions in regional comprehensive economic partnership: What role do economic globalization and disaggregating energy play? *Resources Policy, 85*, 103914. https://doi.org/10.1016/j.resourpol.2023.103914

Authors and Affiliations

Mohamed A. Ali[1,2,3] [ORCID] **· Ashraf Elsayed[2,4] · Islam Elkabani[2,4] · M. Elsayed Youssef[1] · Gasser E. Hassan[1,3]**

✉ Mohamed A. Ali
  mohammed.abdulfattahali@gmail.com; mabdelfatah@srtacity.sci.eg

1   Computer Based Engineering Applications Department, Informatics Research Institute, City of Scientific Research and Technological Applications (SRTA-City), New Borg El-Arab City, Alexandria 21934, Egypt

2   Department of Mathematics and Computer Science, Faculty of Science, Alexandria University, Alexandria, Egypt

3   Faculty of Industrial and Energy Technology, Borg Al Arab Technological University, Alexandria, Egypt

4   Faculty of Computer Science and Engineering, Al Alamein International University, El Alamein, Egypt