ORIGINAL ARTICLE

Open Access

Check for updates

Deep learning-based estimation of PV power plant potential under climate change: a case study of El Akarit, Tunisia

Afef Ben Othman[1,2*], Ayoub Ouni[2,3] and Mongi Besbes[2,4] (ID)

Abstract

Background: Several climatologists and experts in the renewable energy field agree that GHI and DNI calculation models must be revised because of the increasingly unpredictable and powerful climatic disturbances. The construction of analytical mathematical models for the prediction of these disturbances is almost impossible because the physical phenomena relating to the climate are often complex. We raise the question over the current and future PV system's sustainable energy production and whether climate disturbances will be affecting this sustainability and resulting in supply decline.

Methods: In this paper, we tried to use deep learning as a tool to predict the evolution of the future production of any geographic site. This approach can allow for improvements in decision-making concerning the implantation of solar PV or CSP plants. To reach this aim, we have deployed the databases of NASA and the Tunisian National Institute of Meteorology relating to the climatic parameters of the case study region of El Akarit, Gabes, Tunisia. In spite of the colossal amount of processed data that dates back to 1985, the use of deep learning algorithms allowed for the validation of the previously made estimates of the energy potential in the studied region.

Results: The calculation results suggested an increase in production as it was confirmed by the 2019 measures. The findings obtained from the case study region were reliable and seemed to be very promising. The results obtained using deep learning algorithms were similar to those produced by conventional calculation methods. However, while conventional approaches based on measurements obtained using hardware solutions (ground sensors) are expensive and very difficult to implement, the suggested new approach is cheaper and more convenient.

Conclusions: In the existence of a protracted controversy over the hypothetical effects of climate change, making advances in artificial intelligence and using new deep learning algorithms are critical procedures to strengthening conventional assessment tools of the production sites of photovoltaic energy and CSP plants.

Keywords: GHI estimation, Deep learning, PV production prediction, ANN modeling and greenhouse effect

* Correspondence: afef@benothman.de
[1]National School of Engineers of Carthage, University of Carthage, Tunis, Tunisia
[2]Laboratory of Robotics, Informatics and Complex Systems, University Tunis El Manar, Tunis, Tunisia
Full list of author information is available at the end of the article

Background
Introduction
For over a century, energy deficiency has become one of the major universal challenges. The global industry and commerce demand for energy has been urgently increasing. Meeting this demand with conventional resources as fossil fuels is no longer an option because of diminishing resources and global economic growth. Therefore, renewable energy is now an alternative means to sustain energy supply. Indeed, unlike fossil fuels, renewable energy resources are safe, unlimited, and easily exploitable.

Reliance on solar energy as one of the most significant renewable resources is mandatory to cope with this rapidly growing demand. Solar energy efficiency depends highly on wise decisions of where to locate a PV plant that puts solar radiation availability into consideration. This availability is corroborated by the estimation of the highest Global Horizontal Irradiation (GHI), i.e., the total amount of direct and diffuse solar radiation at a given location. Therefore, to choose an optimal location for a photovoltaic or CSP power plant, the following parameters must be considered [1]:

- Solar resource.

In this stage, as a study case area, we must indicate the area with the highest DNI (Direct Normal Irradiation) and GHI incidence or the area whose features meet the formal tender requirements.

- Transmission lines and electrical substations.

The proximity of electrical transmission lines and substations to the chosen area will reduce additional costs of connecting the station to the electricity transmission and distribution network.

- Adequate spatial location.

A third basic requirement is a geographically suitable area for the site. Vast, relatively flat but not floodplain wetlands, and less densely populated areas can be the most appropriate option. For example, the land area required by a 50-MW solar plant with a 6-h storage capacity is about 300 ha. Parabolic troughs and power tower plants require a piece of land with a slope of less than 2% (50 meters altitude difference).

- Other factors.

Many environmental, societal, economic, and political issues such as landownership, water resources, benefits and costs, and policy making should be taken account of before a photovoltaic or CSP power plant project implementation.

Unfortunately, although the aforementioned conditions can be met, it is still globally difficult to obtain radiation measurement data due to equipment high costs. In Tunisia, reliable radiation data are always collected over short periods of time and are only available at few meteorological stations. This can create a severe limitation for solar model application. Moreover, most of the classic mathematical models were developed to operate only under clear sky conditions. Therefore, this makes of them an inaccurate tool to predict global radiation data [2].

The increase in greenhouse gasses has already resulted in significant climate changes. By absorbing infra-red radiation, these gasses control how natural energy flows through the climate system [3]. As a result of anthropogenic emissions, the climate has started to adapt to a "thicker layer" of greenhouse gasses to maintain the balance between energy from the sun and that which is sent back into space.

Observations showed that global temperatures rose by about 0.6 °C during the twentieth century. Unfortunately, it is becoming increasingly evident that most of the heating of Earth's climate system observed over the past 50 years is due to human activities. Climate models expect that the planet's temperature to rise up from 1.4 °C to 5.8 °C by 2100. This will be a significant atmospheric warming, if compared to the temperature degrees experienced in the last 10,000 years. Climate models are based on a wide range of assumptions that take into account factors such as population growth and technological change that may result in future gas emissions. Nevertheless, those future climate projections have not reflected any measures to reduce gas emissions that threaten the climate. As a result, many uncertainties regarding the magnitude and the impact of climate change still persist, particularly at regional levels. Due to the delaying effect of the oceans, surface temperatures do not respond immediately to greenhouse gas emissions. As a result, climate change might continue to exist for centuries even after atmospheric concentrations stabilize.

Climate change is likely to have significant impacts on the planet's environment. In general, the faster climate change progresses, the greater the risk of damage will be. An average sea level rise of 9 to 88 cm is expected by 2100 [4], which would result in the flooding of low-lying areas and other major damages [5, 6]. Other effects would include increased global precipitation and changes in the severity or the frequency of extreme events. Climate zones could move vertically (instead of being belt-shaped and circular around the Poles) towards the Earth's poles, disrupting forests, deserts, grasslands, and other unmanaged ecosystems. This would lead to the decline or disintegration of some of these ecosystems and the extinction of certain species. In spite of these assertions, current knowledge on the assessment

and formal representation of the impacts of climate change remains extremely modest. This results in great uncertainty about the results of modeled approaches based on the evaluation of potential economic and social costs.

Scientific research suggests four sources of uncertainty related to the representation of impacts:

- Assessment scarcity and lack of data on significant, long-term global warming in the presence of serious climate disasters and triggered environmental crises in developing regions;
- The incompleteness of assessments along with the persistent present uncertainties in developing countries;
- Developing country's geographic aggregation, with a masking effect concealing large heterogeneities, including potential very localized ruptures with serious socio-economic consequences;
- The difficulties of monetary valuation, the extrapolation of the response functions is therefore based on general hypotheses concerning their shape and amplitude;

Recurrently, the literature postulates simple functional forms of the power or a polynomial type of order 2, which rule out the possibility of non-linearity. However, apart from questions of technical feasibility, nothing seems to justify their use.

Accordingly, in the absence of feasible solutions to cope with global warming and while waiting for these currently inaccessible functions, our work suggests to build integrated models by means of relatively simple functions. The efficiency of this approach can only be translated in the functions' ability to satisfy the analytical and numerical requirement of global warming's impacts and respect the previously acquired data in this field. Therefore, we suggest that artificial neural networks (ANN) and deep learning (DL), which can use many meteorological and geographical parameters as input, can be a more efficient method to predict GHI using [7]. In this study, we aim to reassess global solar radiation incidence using deep learning software and re-investigating big data record parameters collected by NASA's meteorological data over 34 years (from 1985 to 2018 and available at the Power Data Access Viewer Application). The case study area is El-Akarit, Gabes, which lies (34° 10′ 80″ North, 9° 97′ 28″ East) in the Middle East of Tunisia.

This paper is divided into three parts in addition to the introduction and the conclusion. The first part details the conditions and criteria for choosing a PV production site. We emphasize the digital evaluation methods, the data collection, and the expected results. We discuss the methods' scope and limitations.

We devote the second part to explain ANN's principles of functioning and terms and conditions of use as well as the possible horizons of the generated solutions.

In the third part, we propose a case study to illustrate ANN's potential and provide a comprehensive discussion of the future possible scenarios of climate change and global warming.
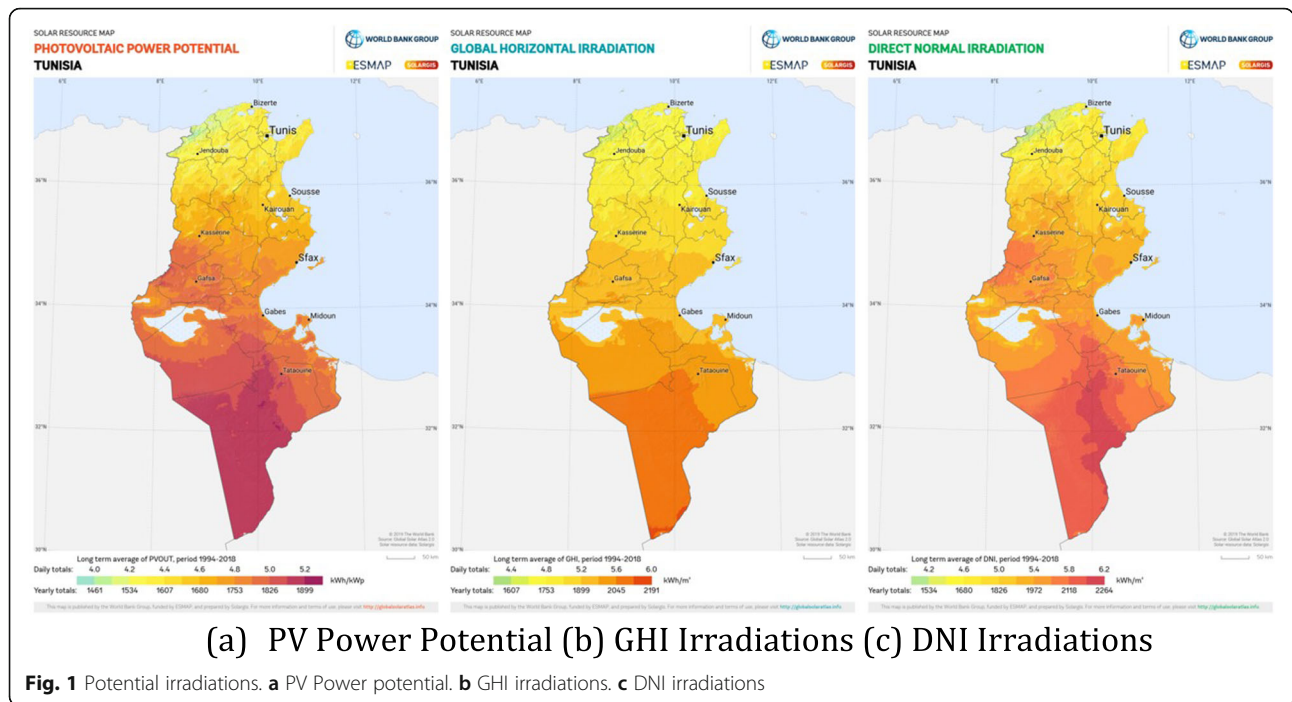
Site assessment methodology

Choosing the appropriate solar plant site is a priority for interested companies. Many criteria should be met in order to determine the optimal solar power plant location. Some of these criteria that might allow optimal exploitation of solar resources are the availability of high solar radiation and the possibility of predicting its spatial and temporal distribution [8, 9]. In this research, we chose El Akarit, Gabes, as a case study area to conduct our estimation of PV power potential (see Fig. 1). The area is located in the middle east of Tunisia which, according to *Solar GIS* maps [10], receives a huge amount of solar radiation [11].

Another reason behind our choice was that the same area was chosen by the National Company of Electricity and Gas (*STEG*) to install the first CSP-type solar power plant in Tunisia. Preliminary studies yielded evidence that the site had enormous energy potential. As a result, we predicted that any future installation of power plants in the same site might be highly profitable (see Fig. 2 and Table 1). Moreover, the site has additional advantages such as flatness, a rocky nature that does not produce high amounts of dust, and its proximity to power lines. Several published studies have given strong evidence of the appropriateness of determining the power plant location based on calculations generated using theoretical numerical models from the laws of physics [12, 13]. However, previous works provided 1-year measurements instead of providing robust measurements based on empirical studies. This limitation could not allow the validation of a certain site choice which normally requires more than 11 years of confirmed measurements with an energy production rate higher than $1.1 \, kW/m^2$.

Climatic condition integration in numerical models designed to predict a site's production stands as a significant constraint and is likely to complicate the calculations. This limitation generates an additional problem with this approach since it does not consider the climatic conditions which might affect the state of the sky (clear, clouds...) [13–15].

To cope with this limitation, we combine solar radiation measurements collected by satellites and the different climatic parameters (wind, temperature, humidity...) with those of the results generated by the models and the corrected data made by the available ground measurements [16–18].

(a) PV Power Potential (b) GHI Irradiations (c) DNI Irradiations

Fig. 1 Potential irradiations. **a** PV Power potential. **b** GHI irradiations. **c** DNI irradiations

We realized that if our work aims to provide accurate and thorough data, it should not rely just on satellite images because they will not be sufficient since both snow cover and altitude errors caused by 2D images might distort the measurements. Moreover, research has highly emphasized the importance of the geographical location of the site as a basic criterion for efficient energy production. Nevertheless, the production is inevitably dependent on the cloud cover which is determined by global climatic conditions (wind, humidity, snowmelt, Gulf Stream...) [19]. Therefore, our objective will be to analyze the effect

Fig. 2 Sunshine over the area of El AKARIT
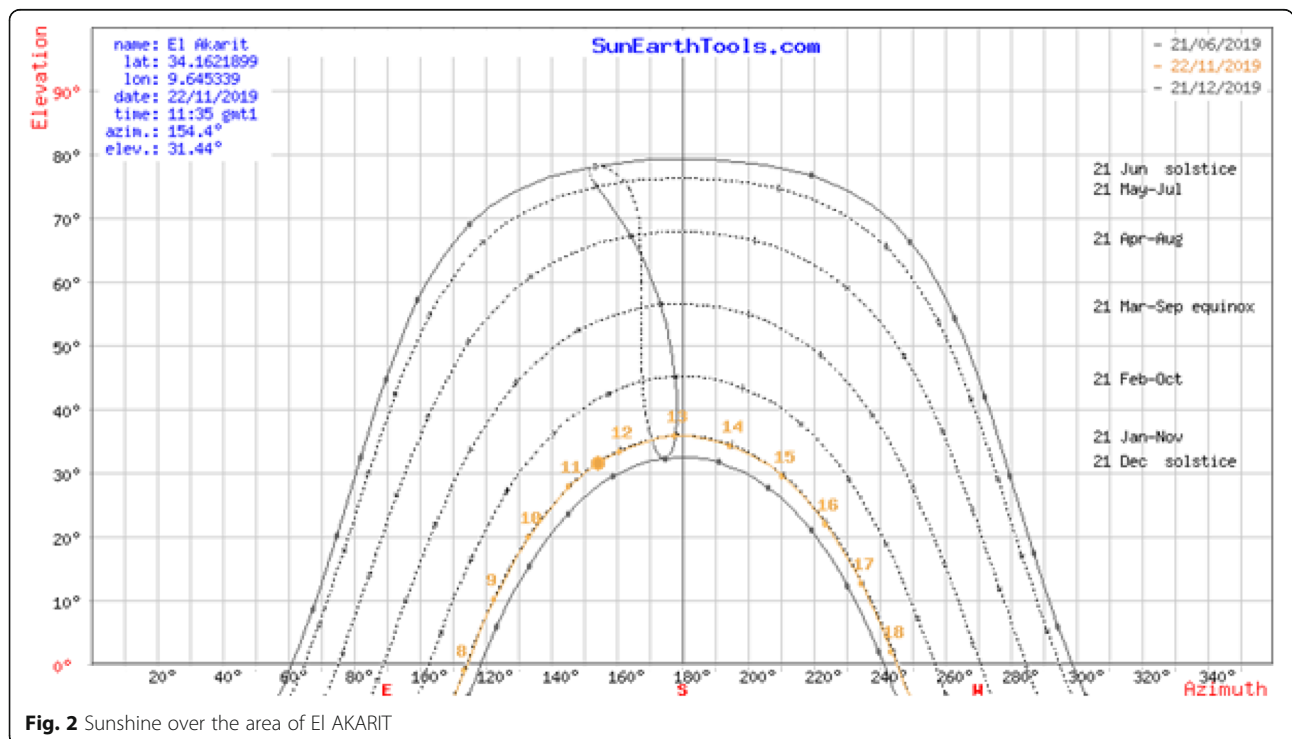
Table 1 Values measured by NASA satellites

Year	GHI (kWh/m²)	T_{Max} (℃)	T_{Min} (℃)	T_{Ave} (℃)	Wind (m/s)	Sky $_{Clear}$ (%)	Sky$_{Partly\ Cloudy}$ %)	Sky $_{Cloudy}$ (%)
1985	1785.46	34.18	4.69	19.54	4.13	31.25	61.93	6.82
1986	1778.50	32.42	6.82	19.56	4.14	30.97	62.50	6.53
1987	1700.50	34.46	4.65	19.69	4.09	23.53	68.63	7.84
...								
...								
2015	1950.51	34.68	5.33	19.92	3.98	60.66	36.64	2.70
2016	1922.98	34.67	7.59	21.04	4.16	52.92	43.38	3.69
2017	1943.06	35.44	5.60	19.64	4.24	60.42	36.86	2.72
2018	1898.71	36.23	7.94	20.18	4.07	52.60	43.12	4.28

of climate change and global warming on PV production in the study area. We will realize this through studying the effects of the variations in both cloud cover and temperature [20]. Considering KT a clear-sky index, we assume that the sky clarity can be classified as follows:

• $KT \in[0\ 0,35]$➡ Cloudy;• $KT \in[0,36\ 0,65]$ ➡ partly Cloudy;• $KT \in[0,66\ 1]$ ➡ Clear.

Data visualization (Fig. 3) reveals that GHI had been determined by climate change over the observed years. Findings show that the clear sky had been the dominant climatic feature in the observed 10-year period (2008–2018). This validates the assumption that there was increase in solar radiation values in the same period.

Table 1 shows that the values of climatic parameters taken in 1980s were different from those recently received.

1. Maximum temperatures increased by about 1.75 ℃.
2. Minimum temperatures increased by about 1.65 ℃.
3. Average temperatures increased by 0.69 ℃.
4. Total sunny days increased by 98 days year-round.
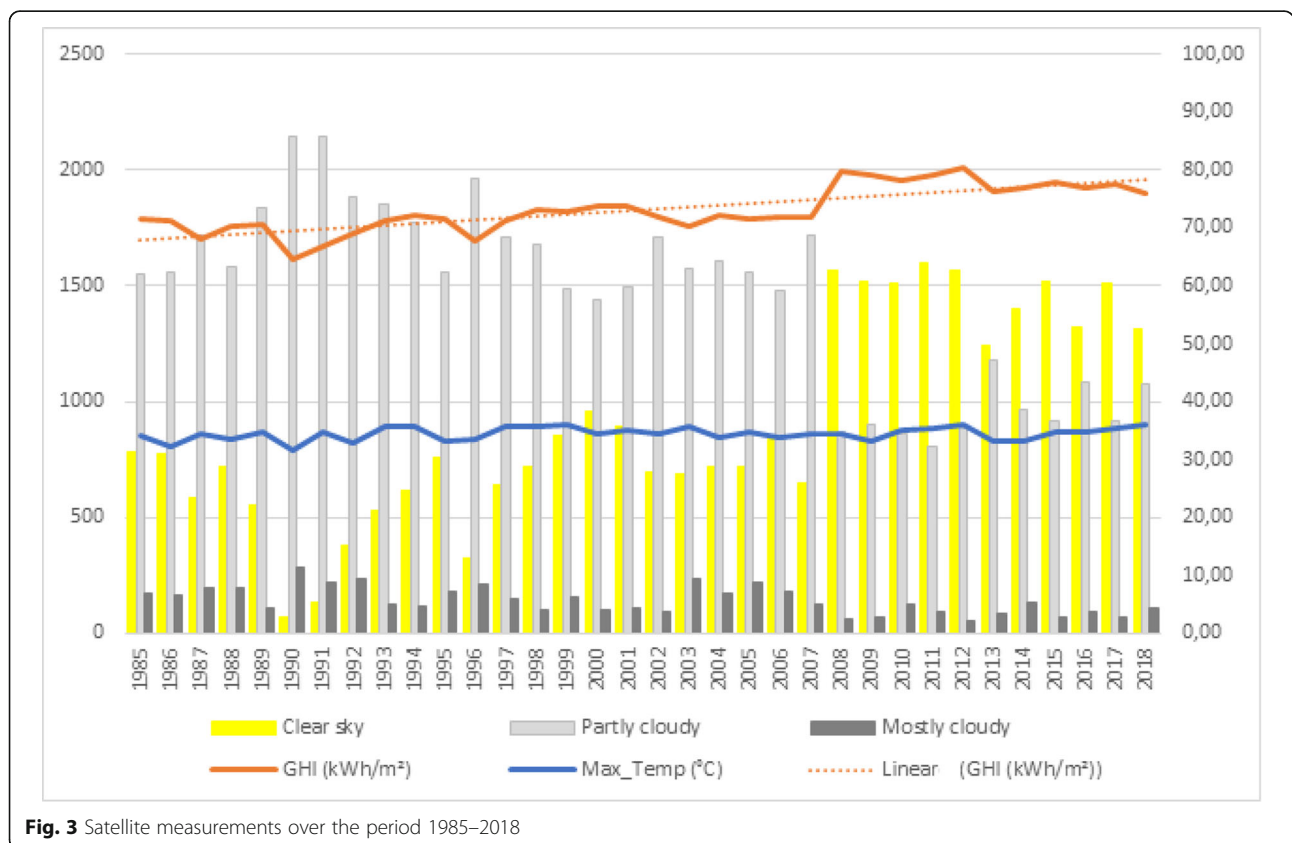
Fig. 3 Satellite measurements over the period 1985–2018

5. The totally cloudy days have decreased by about 13 days throughout the year:

6. ...

The data presented in the figures show significant interference of the climatic conditions in PV systems production. Currently, climatologists argue for the existence of a global anomaly in temperature degrees which might result in an increase by 5 °C by 2100.

Atmospheric warming will inevitably lead to snow cover and glaciers melting in addition to other severe climatic disturbances. The most frequently raised question is:

Will climatic disturbances in areas intended for future installations of solar power plants affect PV or CSP production?

Unlike previous research that has provided results related to answering the question above relying on mathematical or numerical models, we aim, in the following section, to yield estimates of the possible undesirable fluctuations in power production through using deep learning tools.

Methods
Deep learning analysis
Deep learning is the process of discovering patterns and finding anomalies and correlations in large data sets to predict outcomes [14]. It involves methods at the intersection of machine learning, statistics, and database systems. Using a software and some techniques, it is possible to use this information to reduce risks, increase revenues, cut costs, etc. [21–25].

Several advances in business processes and technology have led to a growing interest in data mining in the private and public sectors. Computer networks are one of these innovations which can be used to connect databases and the development of search related techniques such as neural networks as well as the spread of the client/server computing model. This model allows for a user's access to centralized data resources from the desktop and a researcher's increased ability to combine data from disparate sources into a single search source [26, 27].

Deep learning is a powerful set of learning techniques in neural networks. It is a technology in which computational models that are constructed of multiple processing layers are allowed to learn from the data set relying on multiple levels of abstraction (see Fig. 4).

There have been major advances in using these models in different fields such as speech recognition and object detection [28–30]. Metaphorically, neural networks represent a data processing brain, which means that the models are biologically inspired and are not just an exact replica of how the brain cells function (see Fig. 5). A full potential of neural networks might be exploited by using them in many applications. Thanks to their ability to "learn" from the data, their non-parametric nature, and their ability to generalize, they are considered a promising tool to improve data prediction and business classification [31].

Basically, neural networks are tools for modeling non-linear statistical data. They can be used to search for patterns in the data or to model complex relationships between inputs and outputs [32]. Using neural networks as a data warehousing tool, companies collect information from data sets in a process called data mining. The difference between ordinary databases and these data warehouses is that there is actual data manipulation and cross-fertilization that might allow users to make more informed decisions [33, 34].

Neural networks are more efficacious than mathematical models in data prediction. In fact, they can predict various components of solar radiation using some meteorological parameters as input [35–38]. Therefore, their use in many applications has gained increasing popularity as they proved to be one of the best tools for non-linear mapping. Research has previously used them to predict solar radiation properties, such as hourly diffuse, direct and global radiations, and daily global radiation [25, 39, 40].

Hidden layers
Although some feed-forward neural networks might contain five or six hidden layers, two hidden layers are often enough to speed up data convergence. Indeed, the network learns best when mapping is simple. Accordingly, a network designer would almost be able to solve
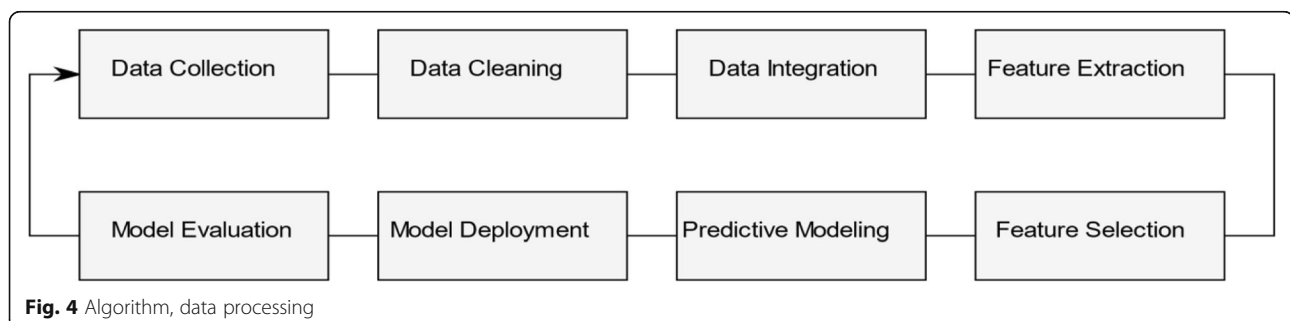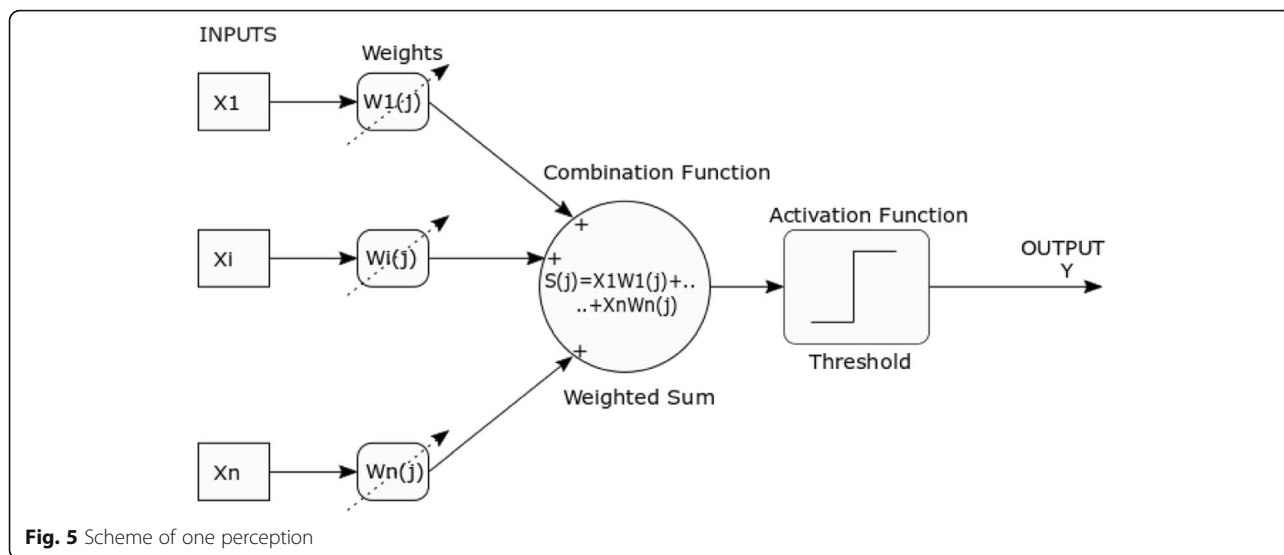

Fig. 4 Algorithm, data processing

Fig. 5 Scheme of one perception

any problem with one or two hidden layers. However, with deep learning, analytic platforms such as *Knime* that deals with complex data sets can deploy up to a hundred layers of neurons.

Hidden neurons

For less than five inputs, the used number of network inputs should approximately double the number of hidden neurons. As the number of inputs rises, the ratio of the hidden-layer neurons to inputs declines. To guarantee a small number of free variables, the weight of hidden neurons should be diminished through reducing their number. As a result, the need for large training sets will decrease. The regulation of the optimal number of hidden neurons for a given problem is always carried through validation set error. *Knime* can deploy up to a hundred of hidden neurons.

Transfer function

Almost any transfer function can be used in a feed-forward neural network. However, to use the

backpropagation training method, the transfer function must be differentiable. The most common transfer functions are the Gaussian, the logistic sigmoid, the linear transfer function, and the hyperbolic tangent. Linear neurons require very small learning rates to be properly trained. Gaussian transfer functions are employed in radial basis function networks which are often used to perform function approximation.

Training process

Supervised machine learning, used in *Knime* software, looks to predict a future outcome. For example, predicting the spread of viral diseases is based on the analysis of models of potentially relevant factors (predictors) such as virulence of the virus, speed of spread, and weather in past data.

As indicated in Fig. 6, the inputs of the prediction algorithm are the data collected from 1985 to 2018 concerning the hourly temperature (max, min, and average values), the wind speed, the value of *KT*, the value of GHI, and the values recorded during the tests of PV
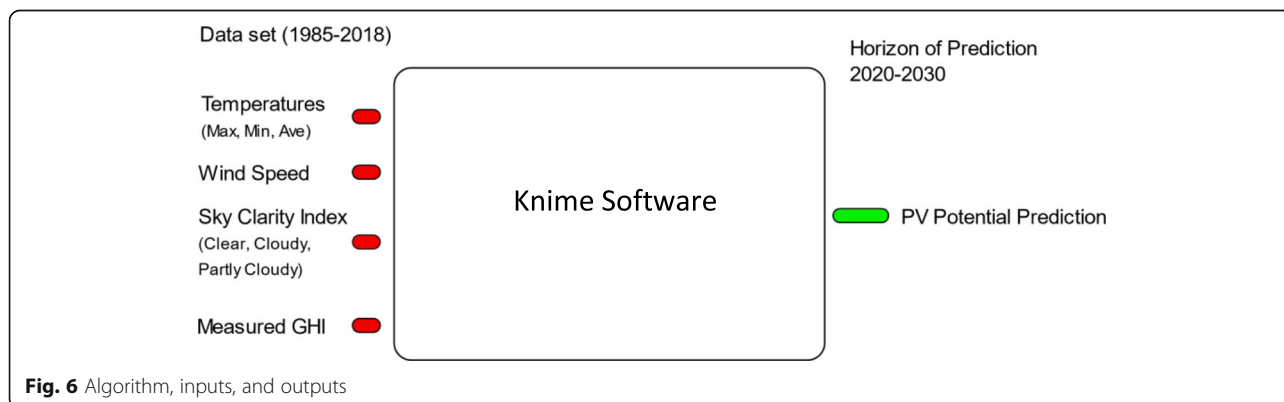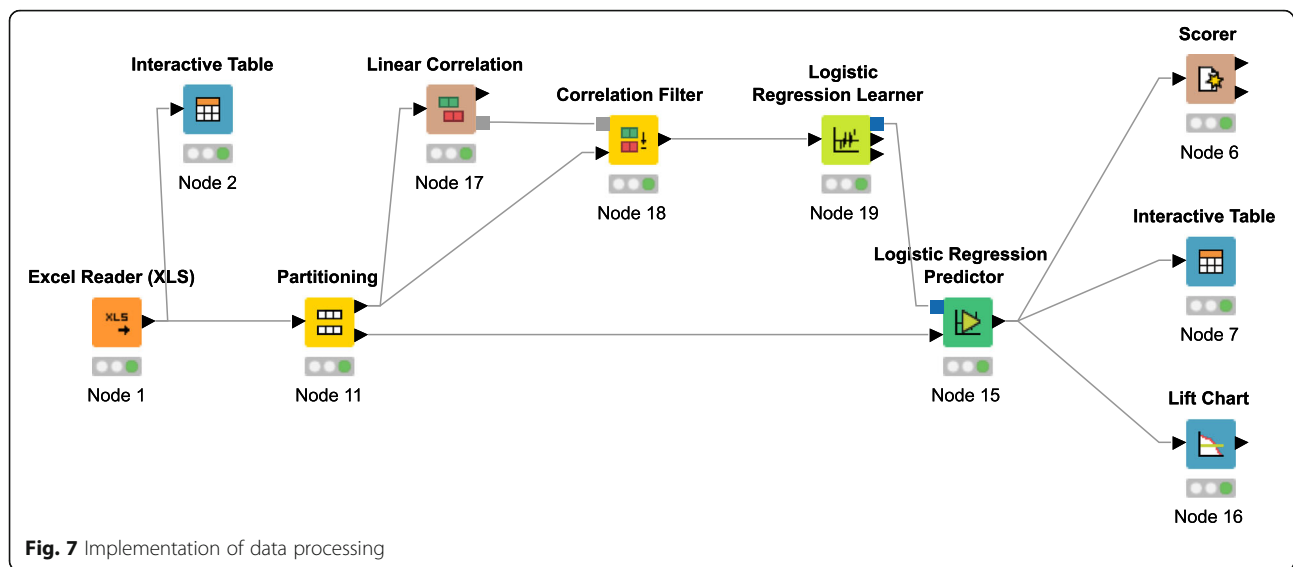

Fig. 6 Algorithm, inputs, and outputs

Fig. 7 Implementation of data processing

power production. The objective was to predict an output related to the next 10 years values of PV power production.

The prediction process occurred in 3 phases: the first phase consisted in formatting and filtering the data that will be introduced into the algorithm. We opted for a quadratic, autoregressive adaptive filtering. In the second phase, we used 60% of the data entered (from 1985 to 2005) to launch the prediction process and to compare the predicted results with the recorded ones (the 40% of the collected data from 2006 to 2018 was not used). This was a smoothly conducted training phase which generated very close results to the recorded ones. The error rate was 2%. The third and last phase involved using the entire database to predict production variations from 2020 to 2030. Figure 7 shows the workflow which is the implementation of the model on the *Knime* software [41, 42]. It is a graphic code.

A workflow is a sequence of nodes; it is the graphic equivalent of a script as a sequence of instructions. Nodes can be connected to each other through their input and output ports to form a workflow. It is a pipeline of the analysis process like read data, clean data, filter data, and train a model [21].

Node 1: Excel Reader

This node reads and provides a spread sheet at its output port.

Node10: Interactive table

Input table to display.

Node 3: Partitioning node

After sampling, the data is usually partitioned before modeling.

Node11: Linear correlation

It calculates for each pair of selected columns a correlation coefficient, i.e., a measure of the correlation of the two variables.

Node 12: Correlation filter

This node uses the model as generated by a correlation node to determine which columns are redundant (i.e., correlated) and filter them out. The output table will contain the reduced set of columns.

Node 5: Logistic regression learner

It performs a multinomial logistic regression. The solver combo box determines the solver that should be used for the problem.
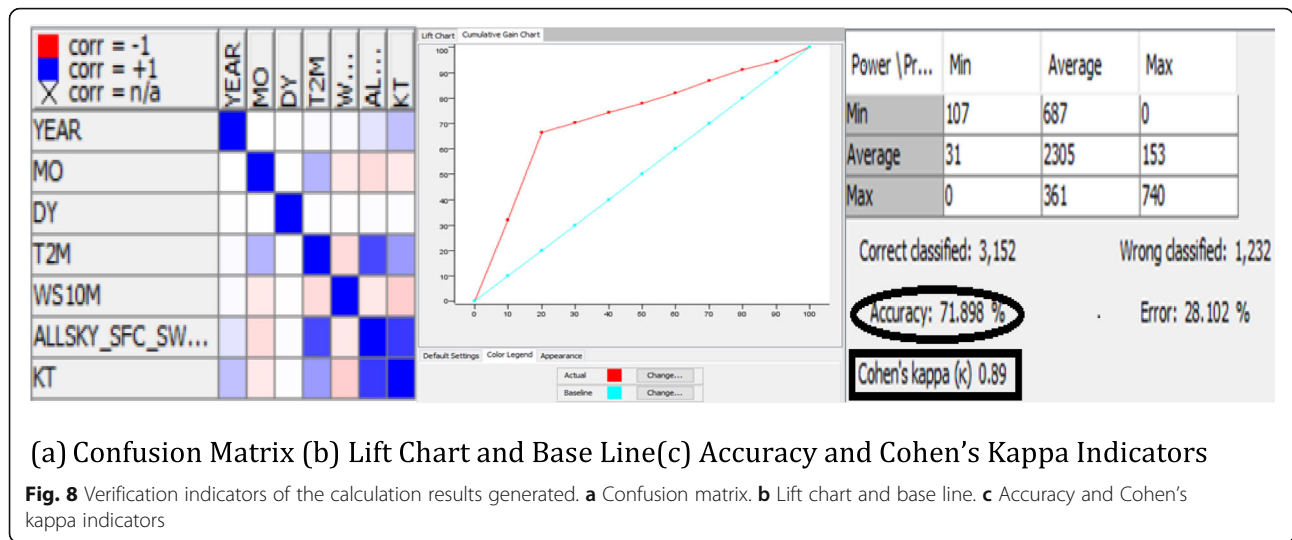
Node 6: Logistic regression predictor

It predicts the response using a logistic regression model. The node needs to be connected to a logistic regression node model and some test data. It is only executable if the test data contains the columns that are used by the learner model. This node appends a new column to the input table containing the prediction for each row.

Node7: Scorer

It compares two columns by their attribute value pairs and shows the confusion matrix, i.e., how many rows of which attribute and their classification match. The dialog allows the selection of two columns for comparison: the values from the first selected column are copied in the confusion matrix's rows and the values from the second column by the confusion matrix's columns. The output of the node is the confusion matrix with the number of matches in each cell. Additionally, the second out-port reports a number of accuracy statistics such as true positives, false positives, true negatives, false negatives, recall, precision, sensitivity, specificity, and *F*-measure, as well as the overall accuracy and Cohen's kappa.

(a) Confusion Matrix (b) Lift Chart and Base Line(c) Accuracy and Cohen's Kappa Indicators

Fig. 8 Verification indicators of the calculation results generated. **a** Confusion matrix. **b** Lift chart and base line. **c** Accuracy and Cohen's kappa indicators

Node 8: Interactive table

It displays data in a table view. In case the number of rows is unknown, the view counts their number when they get open. Furthermore, rows can be selected and highlighted.

Node 9: Lift chart

This node creates a lift chart. Additionally, a chart for the cumulative percent of responses generated is shown. A lift chart is used to assets a predictive model. The higher the lift gets (the difference between the "lift" line and the base line), the better performance the predictive model provides. The lift is the ratio between the results obtained with and without the predictive model. It is calculated as number of positive hits (e.g., responses) divided by the average number of positives without model.

Results and discussion

The obtained results must be verified using four of the most important verification criteria below to prove the accuracy and the reliability of the produced predictions [25, 41]:

1. The *accuracy* indicator
2. The *Cohen's kappa K* indicator
3. The area between the *base line* and the *lift curve*
4. The estimated correlations between the different parameters in the *confusion matrix*.

In this study, the accuracy indicator is 72%, a good percentage which gives the results a high degree of confidence (see Fig. 8c).

Let K be the Cohen's kappa coefficient used to measure the percentage of data values in the diagonal of the confusion matrix table and then adjusts these values to make aleatory agreement occurring possible. K can be classified as follows:

- $K \in [0\ 0,2)$ Poor agreement;
- $K \in [0,21\ 0,4)$ Fair agreement;
- $K \in [0,41\ 0,6)$ Moderate agreement;
- $K \in [0,61\ 0,8)$ Good agreement;
- $K \in [0,81\ 1)$ Very good agreement.

The correlation matrix (see Fig. 8a) is shown as a red-white-blue heatmap. Blue values show a high correlation, white values show correlation zero, and red values show an inverse correlation between the two columns.

The cumulative gains chart (see Fig. 8b) is a visual aid used to measure model performances. So, when the area between the *lift curve* and *the baseline* is great, the predictive model's performance is higher.

Table 2 shows that the effect of global warming is preponderant, that the production of a future power plant will increase by about 11% (see cases d, e, f, g, h, and i), and that climatic disturbances will not have a significant effect on its production.

Even in the case of increased climatic disturbances and temperature degrees decline, power production will increase by a small percentage (of 4%). This result is explained by the fact that the neural networks already internally memorized, which might have increased power production during the past 34 years. The hypothesis that

Table 2 Prediction of GHI evolution during 2020–2030 (main results)

T/KT	↘	→	↗
↘	Case a 1% (18 kWh/m²)	Case b 4% (78 kWh/m²)	Case c 3% (52 kWh/m²)
→	Case d 12% (12 kWh/m²)	Case e 11% (197 kWh/m²)	Case f 10% (186 kWh/m²)
↗	Case g 9% (158 kWh/m²)	Case h 11% (194 kWh/m²)	Case i 11% (11 kWh/m²)

future installed PV plants may risk loss of efficiency due to the climatic disturbances was rejected.

Conclusion

The overall objective of our work was to accurately predict the PV power potential in the case study area for the next 10 years, under changing climatic conditions. In this paper, we suggested the use of deep learning techniques to obtain a good estimate of the PV power which is often difficult to achieve with analytical models. We found that by applying this approach, the production value indicators will increase slightly due to global warming and that PV electricity production will not be affected by climate disturbance in Tunisia. The results generated are credible with an accuracy rate of 72% based on the assumption that solar activity remains the same as observed in past years.

In another way, some scientists will be probably expecting a cooler period in the next following years due to the sunspot activity. So, it will be judicious to analyze the variations in PV production in the coming years by integrating the sun's activity and the possible variations it may have on the climate.

Authors' contributions
ABO, AO, and MB wrote the manuscript. ABO and AO performed the experimental design and data acquisition. ABO, AO, and MB implemented the algorithm and the experiments. ABO and MB evaluated the results. MB comprehensively reviewed the manuscript. All authors read and approved the final manuscript.

Availability of data and materials
Not applicable.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]National School of Engineers of Carthage, University of Carthage, Tunis, Tunisia. [2]Laboratory of Robotics, Informatics and Complex Systems, University Tunis El Manar, Tunis, Tunisia. [3]Higher Institute of Environmental Sciences and Technologies, University of Carthage, Tunis, Tunisia. [4]Higher Institute of Information and Communication Technologies, University of Carthage, Tunis, Tunisia.

References
1. Derouich W, Besbes M, Olivencia JD (2014) Prefeasibility study of a solar power plant project and optimization of a meteorological station performance. J Appl Res Technol 12(1):72–79
2. Alam S, Kaushik SC, Garg SN (2009) Assessment of diffuse solar energy under general sky condition using artificial neural network. Appl Energy 86(4):554–564
3. Ahmad MJ, Tiwari GN (2010) Solar radiation models review. Int J Energy Environ 1(3):513–532
4. White WB, Lean J, Cayan DR, Dettinger MD (1997) Response of global upper ocean temperature to changing solar irradiance. J Geophys Res Oceans 102(C2):3255–3266
5. Feulner G, Rahmstorf S, Levermann A, Volkwardt S (2013) On the origin of the surface air temperature difference between the hemispheres in Earth's present-day climate. J Clim 26(18):7136–7150
6. Haigh ID, Wijeratne EMS, MacPherson LR, Pattiaratchi CB, Mason MS, Crompton RP, George S (2014) Estimating present day extreme water level exceedance probabilities around the coastline of Australia: tides, extra-tropical storm surges and mean sea level. Clim Dyn 42(1-2):121–138
7. Gupta R, Kumar P (2015) Cloud computing data mining to SCADA for energy management. In: 2015 Annual IEEE India Conference (INDICON). IEEE, pp 1–6
8. Fathallah MAB, Othman AB, Besbes M (2018) Modeling a photovoltaic energy storage system based on super capacitor, simulation and evaluation of experimental performance. Appl Phys A 24(2):120
9. Rao CN, Bradley WA, Lee TY (1984) The diffuse component of the daily global solar irradiation at Corvallis, Oregon (USA). Sol Energy 32(5):637–641
10. Cebecauer T, Suri M (2015) Typical meteorological year data: SolarGIS approach. Energy Procedia 69:1958–1969
11. Choi Y, Suh J, Kim SM (2019) GIS-based solar radiation mapping, site evaluation, and potential assessment: a review. Appl Sci 9(9):1960
12. Bechini L, Ducco G, Donatelli M, Stein A (2000) Modelling, interpolation and stochastic simulation in space and time of global solar radiation. Agric Ecosyst Environ 81(1):29–42
13. Belkilani K, Othman AB, Besbes M (2018) Assessment of global solar radiation to examine the best locations to install a PV system in Tunisia. Appl Phys A (2):122
14. Ammann CM, Joos F, Schimel DS, Otto-Bliesner BL, Tomas RA (2007) Solar influence on climate during the past millennium: results from transient simulations with the NCAR Climate System Model. Proc Natl Acad Sci 104(10):3713–3718
15. Belkilani K, Othman AB, Besbes M (2018) Estimation and experimental evaluation of the shortfall of photovoltaic plants in Tunisia: case study of the use of titled surfaces. Appl Phys A 124(2):179
16. Cahalan RF, Wen G, Harder JW, Pilewskie P (2010) Temperature responses to spectral solar variability on decadal time scales. Geophys Res Lett 37(7)
17. Camps J, Soler MR (1992) Estimation of diffuse solar irradiance on a horizontal surface for cloudless days, a new approach. Sol Energy 49(1):53–63
18. El Mghouchi Y, El Bouardi A, Choulli Z, Ajzoul T (2014) New model to estimate and evaluate the solar radiation. Int J Sustain Built Environ 3(2):225–234
19. Batlles FJ, Rubio MA, Tovar J, Olmo FJ, Alados Arboledas L (2000) Empirical modeling of hourly direct irradiance by means of hourly global irradiance. Energy 25(7):675–688
20. Davies JA, McKay DC (1982) Estimating solar irradiance and components. Sol Energy 29(1):55–64
21. Berthold MR, Cebron N, Dill F, Gabriel TR, Kotter T, Meinl T et al (2009) KNIME-the Konstanz information miner: version 2.0 and beyond. AcM SIGKDD Explorations Newsletter 11(1):26–31
22. Kaouther B, Othman A, Besbes M (2018) Estimation of global and direct solar radiation in Tunisia based on geostationary satellite imagery. In: 2018 IEEE PES/IAS Power Africa. IEEE, pp 190–194
23. Kopp G, Lean JL (2011) A new, lower value of total solar irradiance: evidence and climate significance. Geophys Res Lett 38(1)
24. Ngiam J, Khosla A, Kim M, Nam J, Lee H, Ng AY (2011) Multimodal deep learning. In: Proceedings of the 28th international conference on machine learning (ICML-11), pp 689–696

25. Nielsen MA (2015) Neural networks and deep learning 25. Determination press, San Francisco
26. Paugam-Moisy H (2006) Spiking neuron networks a survey (No. REP ORK). IDIAP
27. Wu AG, Dong RQ, Fu FZ (2015) Weighted stochastic gradient identification algorithms for ARX models. IFAC-Papers Online 48(28):1076–1081
28. Bertsekas DP (2010) Incremental gradient, subgradient, and proximal methods for convex optimization: a survey. Optimization Machine Learning 3:1–38
29. Gueymard C (2000) Prediction and performance assessment of mean hourly global radiation. Sol Energy 68(3):285–303
30. Hagan MT, Menhaj MB (1994) Training feedforward networks with the Marquardt algorithm. IEEE Trans Neural Netw 5(6):989–993
31. Almeida JS (2002) Predictive non-linear modeling of complex data by artificial neural networks. Curr Opin Biotechnol 13(1):72–76
32. Chen WS, Du YK (2009) Using neural networks and data mining techniques for the financial distress prediction model. Expert Syst Appl 36(2):4075–4086
33. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444
34. Mesejo P, Ibanez O, Fernandez-Blanco E, Cedron F, Pa-´zos A, Porto-Pazos AB (2015) Artificial neuron glia networks learning approach based on cooperative coevolution. Int J Neural Syst 25(04):1550012
35. Qahwaji R, Colak T (2006) Neural network-based prediction of solar activities. CITSA 2006, Orlando, pp 4–7
36. Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85–117
37. Sen Z (1998) Fuzzy algorithm for estimation of solar irradiation from sunshine duration. Sol Energy 63(1):39–49
38. Sivamadhavi V, Selvaraj RS (2012) Prediction of monthly mean daily global solar radiation using artificial neural network. J Earth Syst Sci 1216:1501–1510
39. Fadare DA, Irimisose I, Oni AO, Falana A (2010) Modeling of solar energy potential in Africa using an artificial neural network
40. Wang J, Ma Y, Zhang L, Gao RX, Wu D (2018) Deep learning for smart manufacturing: methods and applications. J Manuf Syst 48:144–156
41. Jagla B, Wiswedel B, Coppee JY (2011) Extending KNIME for next-generation sequencing data analysis. Bioinformatics 27(20):2907–2909
42. O'Hagan S, Kell DB (2015) Software review: the KNIME workflow environment and its applications in Genetic Programming and machine learning. Genet Program Evolvable Mach 16(3):387–391

Publisher's Note