

Hybrid Tree-Ensemble and Metaheuristic Dust Aware Frameworks of Solar PV Power Prediction in Arid Regions: Scoping Review Evidence, Gaps, and Future Directions From the Case of Iraq and the MENA Region (2020–2025)

Fahad Ahmed Hadi^{1*}, Mohanad Hasan Ali²

¹² Electrical Engineering Techniques Department, Mussaib Technical College, Al-Furat Al-Awsat Technical University, Babylon, Iraq

*Correspondence: Fahad Ahmed Hadi; Electrical Engineering Techniques Department, Mussaib Technical College, Al-Furat Al-Awsat Technical University, Babylon, Iraq; Email: fahad.ahmed.tcm27@student.atu.edu.iq (F.A. Hadi); com.mhn@atu.edu.iq (M.H. Ali); Tel: +9647704882232 (F.A. Hadi); +9647735401438 (M.H. Ali)

Received 20/04/2026, Received in revised form 19/06/2026

Accepted 03/07/2026, Available online 10/10/2026

FJIAS 2026, 2(S1): 24-67

Special Issue: Selected Papers from the 3rd International Conference on Sustainable Development and Climate Change (ICSDCC 2026). Organized by Al-Mussaib Technical College / Al-Furat Al-Awsat Technical University, Iraq, in cooperation with Delta University for Science and Technology, Egypt. Held at Delta University for Science and Technology, Mansoura, Egypt, on 18 July 2026.

ABSTRACT

Background: In arid and hyper-arid areas, solar photovoltaic (PV) power generation prediction is hampered by high aerosol optical depth (AOD), frequent dust events, high temperatures, and sparse ground-based measurement networks. In Iraq, the frequency and intensity of dust storms are especially acute, especially under these conditions. inflict quantifiable effects on the solar irradiance supply and PV systems output. Purpose: This scoping review systematically identifies evidence on hybrid generation prediction frameworks (2020-2025) that are peer-reviewed. hybridizing tree-based ensemble machine learning models with metaheuristic optimization algorithms on solar PV. prediction of power generation with specific emphasis on the applicability of the approach to arid climate and the gaps in Iraq-specific research. Methods: A scoping review in adherence to PRISMA-ScR reporting principles was conducted in accordance across four databases (Scopus, Web of Science, IEEE Xplore, ScienceDirect) of a scientific character. Studies were screened against five pre-established eligibility criteria and categorized under an eight-level taxonomy (T1-T8). N=25 Core studies. were held in evidence to synthesize evidence, with N=26 Contextual references to give methods, environmental, or geographic supporting evidence. Results: Tree-ensemble models prevail on the corpus: Random Forest appeared in N=7 of 21 empirical Core studies (33.3%) and XGBoost in N=6 of 21 (28.6%). Only N=4 experiments regarding the larger literature affirmed Tier-1 hybrid arrangements the use of metaheuristic tree optimizers. hyperparameter tuning the core arid-climate ensemble of N=1 solar generation prediction in four of the twenty-one empirical Core studies (19.0%) dust and aerosol variables were directly used as inputs to the model, but with 54.5 percent Solar irradiance reduction which can be related broadly down to Iraqi measurement over time and 62.3 percent reduction distantly from

American reference point level respectively we bring attention years between 2004-2012. Impact of extreme periods of dust events on PV power loss Tree-Ensemble Forecasting Directly : Iraq Gap: N=1 Core study was the only such study. to an Iraqi dataset, validating a MERRA-2 satellite data channel at Mosul but absent metaheuristic optimization. or dust-aware feature engineering. Research Agenda: There are four suggested priority directions: (i) dust-aware feature. ML-retrieved or MAIAC-corrected AOD engineering; (ii) systematic metaheuristic evaluator of optimizers. models in a data-sparse arid environment; iii) validation against satellite-derived forest-cover estimates specific to Iraq. inputs; (iv) the use of common metric protocols, as well as a dust-event subset evaluation framework dedicated to the standardization of benchmarking procedures for enablement consistent cross-study comparability and reproducibility. This scoping review was not prospectively registered; although the PROSPERO database accepts registrations for scoping reviews, no formal protocol was archived or publicly available before the start of our work. However, we strictly followed the PRISMA-ScR guidelines and also the JBI methodology at all steps of a scoping review to ensure methodological rigor and transparency. Potential limitations of the present study include limited dual-independent literature screening, which is expected to add some selection bias; for instance five references were detected after-search and beyond an a priori defined window period (ie. (2025), October 2025: Taha et al. (2025), 02:00 pm November 2025; Shehab et al. (2026); and Oruç et al. (2026) were intentionally kept solely as additional contextual evidence instead of integrated into the main analysis. Additional limitations include conducting our literature search only in English-language databases, which may mean missing relevant studies published in other languages (see Table 1 for justification of inclusion and exclusion criteria), as well as the lack of a pre-archived, prospectively registered protocol to enhance replicability and transparency of the review process.

Keywords: Photovoltaic power prediction; Tree-ensemble machine learning; Metaheuristic optimization; Aerosol optical depth; Dust storms; Arid climate

1. Introduction

Solar photovoltaic energy has emerged as one of the fastest-growing renewable technologies worldwide, with global installed capacity surpassing 2.2 terawatts by end of 2024 and annual additions exceeding 550 gigawatts — the highest annual deployment on record (IEA 2025). The integration of large-scale PV generation into electrical grids requires accurate forecasting of solar irradiance and PV power output to support scheduling, reserve planning, and energy management (Di Leo et al. 2025). The main research problem is to enhance the accuracy of forecasts.

particularly as renewable energy is increasingly penetrated, and grid operators require a good predictability of generation. in different climatic conditions (Al-Dahidi et al. 2024). Forecasting is much more complicated in arid and hyper-arid environments. Atmospheric, in these places there is atmospheric not only variability in clouds but also high levels of aerosol, high turbidity. frequently dust events, and temperature spikes (Koprulu et al. 2025; Saini et al. 2023). Aerosols and dust in the air. scatter and absorb the incoming solar radiation, accumulate dust resulting in surface soiling that. immediately lowers the PV electrical power (Maghami 2025). A meta-analysis on the research of dust deposition on large scale. in 860 publications it was confirmed that dust is one of the most important environmental stress factors of PV system.

Are more impacted in arid regions around the globe (Maghami 2025; Saini et al. 2023). A long-term climatological study of meteorological records in Baghdad and Nasiriya recorded 3.1 major dust storm days and 7.8 days, respectively, in 1980–2015 (Attiya and Jones 2020), and

132 dust storm events occurred in Iraq during 2020-2023 alone based on satellite-based analysis (Ismail et al. Field measurements in Baghdad established that severe dust storm events decrease the PV output power by about 62.3% when compared to baseline conditions before the storm (Chaichan et al. 2023), and economic modelling of four Iraqi cities revealed that insufficient cleaning frequency raises Levelized Cost of Energy (LCOE) by over 30% (Hameed et al. 2024). In many arid regions where ground-based measurements are limited, satellite-derived datasets provide the primary data pathway for solar resource assessment. The Prediction of Worldwide Energy Resources (POWER) project of NASA.

is in a position to offer globally standardized datasets of solar radiation (NASA POWER 2025) whereas the AERONET network is in place. provides quality-controlled AOD measurements commonly applied in the investigation of atmospheric and solar energy (Holben et al.). (1998). Al-Rikabi et al. (2025) and Hameed (2025) note that in spite of the high solar potential and imperative of renewable energy in Iraq, the country has not capitalized on its potential as a solar nation to date. The supply of good quality ground-based irradiance measurements is also scarce (et al. 2024), and the majority of measurements in Iraq were of low quality. The solar resource analysis is based on the data collected by satellites or reanalysis (Al-Rikabi et al. 2025). There are a few extensive review studies that have analysed PV forecasting approaches on a general level. Al-Dahidi et al. (2024) has conducted a review of 257 publications and suggested a seven-module data-driven forecasting framework in which the data is recorded. tree-ensemble (XGBoost, Random Forest, Gradient Boosting) and metaheuristic optimizers (GA, PSO, GWO, WOA) as new hybrid elements; but their survey dealt not with the particular combination. of tree-ensemble models optimized by the metaheuristic nor dust-filled or dry-climate situations as the main ones. variables. Di Leo et al. (2025) conducted a survey of the progress in PV forecasting and reported the increased use of PSO, GA, and Differential Evolution to hyperparameter of ANN, SVM and LSTM; metaheuristic. independent validation of optimization of SVM architectures to predict solar generation has also been tested.

But coverage was still skewed towards temperate and data-rich environments and not addressed. High aerosol loading (dust) combined with relatively few well-distributed ground-based observational stations gives rise to typical methodological limitations associated with application of available methodologies in hyper-arid and rainless regions like Iraq. Alcañiz et al. conducted a wide-scope scoping review, It was found that of N=100 PV forecasting studies published over 5 years from the year of 2015–2020, only N=22 (22.0%) dealt with dry/arid climatic conditions [10]. Importantly, within this sub-set there were none where the directly relevant studies had been performed on Iraq and hence provide unambiguous quantitative evidence for a substantive geographic and climatic research gap that is fundamental to support and motivate the rationale for the present review.

Al-Dahidi et al. (2023) This series of three reviews on protocols for Cyberbullying Research is a comprehensive review of established and potential cyberbullying related research. 2024; Di Leo et al. 2025; Alcañiz et al. By additional complementary and supplementary analyses done by Iheanetu (2022), Mohamad Radzi et al. (2023). (2023), Chu et al. Collectively, a series of critically important, interrelated research gaps warranting deeper exploration are revealed and then left unresolved (i.e., the general performance/rigor as well as real-world agility of tree-ensemble models—specifically improved and parameterized through sophisticated metaheuristic algorithms—when employed under arid conditions characterized by extreme energy feedback versus data sparsity [pseudodust] or other similar aspects characteristic of harsh, dust-soaked environments; ii) the functional role and quantitative contribution facilitated specifically by aerosol/dust-related variables—including Aerosol Optical Depth (AOD),

particulate matter with diameter less than 2.5 micrometers (PM_{2.5}), atmospheric turbidity—as vital input parameters within hybrid photovoltaics generation models; and (iii) the current state, extent and scope of solar generation forecasting research relating to Iraq and other geopolitically similar countries across the Middle East & North African (MENA) region. No scoping review has yet quantitatively and systematically mapped this complex three-way intersection within the limits of our systematic search strategy, which defines the scope of a retained corpus of literature.

In order to thoroughly engage with the identified gaps, three specific and interlinked research questions are constructed to shape the ensuing review:

RQ1: In which tree-ensemble models and metaheuristic optimizers have been used in solar PV power generation?

prediction in peer-reviewed literature (2020-2025) and the frequency with which they are mixed in proven hybrid.

configurations?

RQ2: How much of the reviewed studies include the dust-related atmospheric variables (AOD, PM_{2.5}, and so on)?

as a predictor of atmospheric turbidity) as inputs into a model, and what are the corresponding improvement in the generation prediction performance.

with their inclusion?

RQ3: What is the state of tree-ensemble/metaheuristic-based solar generation prediction study on?

What are the methodological gaps that are to be given priority consideration in Iraq and other arid and hyper-arid areas?

The goals of the review are to: (i) map systematically through the landscape of tree-ensemble and metaheuristic solar.

generation prediction published 2020-2025; (ii) describes the application of dust and aerosol variables in more than one place.

the corpus that is retained; (iii) discover gaps in Iraq-related research and generalizable methodological discoveries; and (iv)

Suggest a four-part research agenda of the development of dust-aware hybrid generation prediction.

The rest of this paper is structured in the following way. Section 2 explains the scoping review process, protocol, eligibility criteria, search strategy, study selection process, and data extraction procedures. Section 3 summarizes the evidence that has been retained in four dimensions of analysis in accordance with the TCCM (Theory, Context, Characteristics, and Methods) framework: theoretical landscape (Section 3.2), contextual and environmental factors (Section 3.3), model characteristics and performance (Section 3.4), and methodological approaches and data source patterns (Section 3.5). Section 4 is a structured discussion in the framework of the seven-move review, with comparison to previous reviews, unexpected findings, theoretical and practical implications, a structured research agenda, and a statement of review limitations. Section 5 gives some concluding remarks.

2. Material and methods

2.1. Review Design and Protocol Registration

This research is designed as a scoping review, based on PRISMA-ScR framework (Tricco et al., 2018) and JBI guidelines (Peters et al., 2020). Scoping reviews are appropriate in those cases when the literature is rather haphazard in terms of scope and focus- the idea is to get an idea of what is available, how scientists have already dealt with the issue, and where the apparent gaps exist, not to conduct a statistical pooling or subject individual studies to a formal quality filter. It was reduced to the actual appearance of the literature. There is a great deal of variation among the studies in this field, different models, different datasets, different climates, and that sort of variability does not make meta-analysis a good fit. The review is gap characterisation oriented and a scoping approach is more suited to that purpose than any other approach would be.

The first limitation that should be mentioned initially is that no protocol was registered before the beginning of this review. Although PROSPERO has been accepting scoping review registrations since 2018, the lack of a prospectively archived protocol is a known methodological limitation, and is further discussed in Section 4.7. The PRISMA-ScR guidelines (Tricco et al. 2018) and JBI methodology (Peters et al. 2020) were followed in all phases of the review. The initial electronic search was done in March 2025. Additional backward citation tracking had been performed up to April 2025. Later, when preparing the manuscript and revising it (August 2025–November 2025), three new references were found that were published after the end of the main search window: Kut and Pietrucha-Urbanik (2025) (published 18 August 2025), Yang et al. (2025) (published 20 October 2025), These were not located during the initial search or citation-tracking procedure: they were discovered by chance during the revision of the manuscript and were included as contextual supplementary evidence to recognize the emerging relevant literature. Since they are not Core studies according to Criterion D and do not appear in all Core N-counts as well as all frequency calculations based on denominator and all empirical synthesis results, they do not qualify as Core N-counts at all. Likewise, the article Shehab et al.. (2026) has been found after the search and will be considered contextual evidence outside of the 2020-2025 prime search period; it will not be counted in any of the Core N-counts accordingly. Also, Oruc et al.. (2026) was mentioned in the revision of the introduction as background information about the use of metaheuristic optimization in SVM architectures; it is also not included in all core N-counts and denominator-related calculations of the same nature. The main results of this review are limited to the articles that include all five inclusion criteria during the primary search between March 2025.

2.2. Eligibility Criteria

The studies were evaluated against five preset eligibility criteria, arranged along the dimensions needed to be included in the Core analytical sample (Table 1). Those studies that passed Criterion A (tree-ensemble) but failed Criterion B (metaheuristic) were kept as Core T2 references, provided that they also passed Criteria C through E, and provided a single-component tree-ensemble baseline against which hybrid T1 configurations are compared. Those studies that satisfied Criterion B but not A were also retained as Core T3 references, provided that they met Criterion C–E. The retained literature did not find any T3 studies that met all of Criteria B through E. This is in itself a finding (see Section 3.1). Prior review studies (T8) were stored in a different non-empirical synthesis layer within the Core corpus, but not counted in any frequency counts or performance metric calculations, which are computed using empirical forecasting studies only. T8 studies that met Criterion D were retained because of their theoretical and distributional value but not counted in any frequency counts or performance metric calculations.

On Criterion E: empirical studies whose main aim is to model atmospheric proxies (such as AOD, determined using ground irradiance) or aerosol time series (such as prediction of aerosol time series) were retained in the empirical Core, as long as their results are directly consumed by solar irradiance or PV forecasting pipelines. The logic behind this is simple: the aerosol-forecasting interface is sufficiently close to the core issue of the review that a strict delimitation of it would exclude work that actually is of interest in comprehending the effects of dust on PV output. AOD retrieval accuracy was used to maintain, studies that met Criterion A or B, but failed to meet one or more of Criterion C, D, or E as Contextual references in which they provided methodological, environmental or geographic evidence to the review purposes. This included dust characterization, AOD retrieval, Iraq-specific solar resource assessments, and PV soiling.

Criterion	Dimension	Inclusion	Exclusion
A	Algorithmic	≥ 1 tree-based ensemble model (RF, XGBoost, GBM, LightGBM) as primary or integrated forecasting component	No tree-ensemble model present
B	Optimization	≥ 1 nature-inspired metaheuristic (GA, PSO, GWO, WOA, SSA, BBOA, or equivalent) for HPO, feature selection, or model configuration	Bayesian optimization or grid search only (not classified as metaheuristic)
C	Environmental	Arid, semi-arid, or dust-influenced environment; OR incorporates AOD, PM _{2.5} , or turbidity as input variables	Exclusively temperate or maritime climate without dust-related variables
D	Temporal	Published 2020–2025 inclusive	Published before 2020 or after 2025
E	Domain	Direct solar irradiance prediction or PV power generation prediction application; or forecast-relevant atmospheric proxy modelling (AOD retrieval, aerosol time-series prediction) where outputs directly support solar generation prediction pipelines	PV system sizing, economic modelling, or grid design without forecasting component; atmospheric studies with no demonstrated link to solar irradiance or PV forecasting

Table 1. Predefined eligibility criteria for study inclusion and exclusion. HPO = hyper parameter optimization; RF = Random Forest; GBM = Gradient Boosting Machine; GA = Genetic Algorithm; PSO = Particle Swarm Optimization; GWO = Grey Wolf Optimizer; WOA = Whale Optimization Algorithm; SSA = Salp Swarm Algorithm; BBOA = Brown Bear Optimization Algorithm.

2.3. Search Strategy and Database Selection

To conduct the literature search, we searched four databases in March 2025, Scopus, Web of Science, IEEE Xplore and ScienceDirect. The query was constructed based on what we were seeking: literature that used tree-based forecasting with some sort of metaheuristic tuning, preferably in dusty or arid environments. The complete Boolean strings are given in the Supplementary Material. We also conducted a more generalized form of the same query without the environmental restrictions, as not all methodologically relevant hybrid studies explicitly identify dust or aridity, although they might be conducted under similar conditions. Articles that survived the first screening were subsequently subjected to hand checks to verify those articles that had not been retrieved by the database queries.

2.4. Study Selection Process

The four database records were imported and reduplicated, and then the screening process started. In the four databases, 520 records in total was retrieved by the search: the largest number of records was in Scopus (210), then Web of Science (140), then IEEE Xplore and ScienceDirect (95 and 75 respectively). After the removal of duplicates, 350 records was transferred to title and abstract screening. Title and abstract screening was handled by the first author (F.A. Hadi). Where a record sat on the fence, the second author (M. H. Ali) stepped in to review it, and the two authors talked through any cases they could not immediately agree on. Records that could not be confidently excluded at the title and abstract stage were pulled for full-text review, with final calls on ambiguous cases confirmed by the second author. The studies that survived were subsequently sorted into a set of eight, tiered taxonomy (T1–T8). The levels address various facets: T1, T2 and T3 are about algorithmic structure; T4, T7 are about the types of data and variables; T5, T6 are about the location or operational setting; and T8 is about previous review research that constitutes the non-empirical theoretical layer (Section 2.2). Each study was designated one main tier, and in cases involving it, a secondary tier, no study has more than two labels. Records that were excluded were recorded with reasons recorded according to the criterion (A-E). Overall, 25 Core studies that fit all the relevant criteria and 26 Contextual references that fit some criteria were retained. The entire selection procedure, such as the number of records at each phase, is documented in Section 3.1 in accordance with PRISMA-ScR guidelines (Tricco et al., 2018).

2.5. Data Extraction and Charting

The extraction of data was done in a structured charting template that was used in all 25 Core and 26 Contextual studies. In cases where a necessary field was not reported in the source paper, it was marked as NR instead of estimated or inferred.. Each eligible study was extracted on the following fields: Author(s) and year of publication; Location and climate zone of the study; Type(s) of tree-ensemble model; Type and functional role of a metaheuristic optimizer (feature selection, hyperparameter tuning, or both); Dust, aerosol, or soiling-related input variables; Forecast horizon and temporal resolution; Reported performance measures (RMSE, MAE, MAPE, R-squared). The synthesis of charted data was done in a narrative manner in four dimensions of analysis; Theoretical landscape (Section 3.2), Section 3.3, Section 3.4, and Section 3.5; Fields were charted instead of using narrative synthesis or formal quality scoring. No weighting or appraisal scores were based on the extraction template. The above methodology forms the systematic structure of evidence retrieval and classification. Section 3 shows the synthesized results of the N = 25 Core studies, which are arranged based on the TCCM (Theory, Context, Characteristics, and Methods) evidence-synthesis framework.

2.6. Evidentiary Boundaries and Decision Rules

The review should be considered a systematic evidence-mapping activity, rather than an effectiveness assessment. The records that are kept are classified based on some eligibility criteria. Frequency calculations were made using the denominator, which only included empirical core studies unless otherwise specified. The past reviews are kept to provide context, but not to be part of the empirical algorithm-frequency analyses. The results can be considered descriptive evidence of trends, gaps, and priorities in the retained corpus due to partial dual-independent screening, the absence of formal critical evaluation and large variations in datasets, climates, forecast horizons, and reported metrics. They do not make conclusive assertions regarding model excellence or applicability in various contexts. All the measures in this review that are based on the denominator are only calculated on the retained empirical core synthesis subset unless otherwise indicated. One of the core studies (Muramshetti et al., 2025) is published in a journal of minor-database-indexing (J Intell Syst Internet Things, ASPG). It

meets all five eligibility criteria and is included in the core corpus, but the frequency-count contributions of this study should be interpreted with due caution due to this indexing limitation..

3. Results and Evidence Synthesis

3.1. Overview of Included Studies

After deduplication and two-stage screening, N=25 Core studies and N=26 Contextual references were selected from an initial pool of 350 unique records. The temporal distribution of the N=25 Core studies was skewed towards 2024–2025 (N=14 of 25, 56.0%), which aligns with the increasing rate of hybrid machine learning research. Geographically, N=13 of 25 Core studies (52.0%) took place in Europe or East Asia, and N=4 of 25 (16.0%) in Iraq locations or datasets. The primary classification tier included N=1 study (4.0%) Tier-1 Hybrid (tree-ensemble + metaheuristic) in the Core arid-climate sample; N=12 of 25 (48.0%) T2 (Tree-Ensemble only); N=0 T3 (Metaheuristic only, no qualifying studies identified); N=4 of 25 (16.0%) The rest of the N=4 Core studies are primarily designated as T5, which is the Iraq-specific context. They also possess secondary tier labels, which are indicative of subordinate algorithmic or contextual factors. Each study has at most two tier labels. The frequency calculations in Section 3.4 and 3.5, which are dependent on the denominator, including the number of algorithm studies, the number of studies in the horizon distribution, and the reporting rate of performance metrics, are all based on the N=21 empirical forecasting studies, which are calculated as N=25 minus the N=4 T8 review studies, unless otherwise specified. The previous review articles were retained to give a theoretical and distributional background but not included in the denominator-based empirical analyses. The four T8 prior reviews are discussed first in Section 3.2 to lay out the theoretical landscape before showing empirical findings. The study selection process went through four sequential stages, following the PRISMA-ScR reporting standard (Tricco et al. 2018). Four databases, including Scopus (210), Web of Science (140), IEEE Xplore (95), and ScienceDirect (75), were searched using databases, and found 520 records. Following deduplication, 350 unique records were sent to title and abstract screening, where 289 records were eliminated due to a clear failure to meet one or more of the eligibility criteria (A, B, D, or E). The rest of the 61 records were reviewed in full-text; 10 were excluded with reasons recorded according to criterion. N = 25 core studies and N = 26 contextual references were retained and N=51 records were obtained to synthesize the evidence.

In addition to the descriptive counts, there are four structural patterns that should be highlighted and put the evidence base in context in which the following sections should be read. First, the temporal acceleration (N=12 of 21 empirical studies in 2024-2025, 57.1% of the corpus) indicates that hybrid tree-ensemble forecasting is a sub-field that is not a mature one, but a maturing one, and the gap in Iraq is increasingly becoming consequential as global approaches are being developed without regional testing. Second, the geographic concentration of Europe and East Asia (52.0) versus N=4 Iraq-based studies (16.0) is not a mere sampling artefact - it is due to the paucity of high-quality ground-based measurement networks in arid MENA regions, which in turn drives the satellite-derived pathways (MERRA-2, MAIAC) in Sections 3.4 and 4 Third, N=1 of 21 empirical studies (4.8%) is a Tier-1 hybrid according to Criterion B, which proves that metaheuristic optimization of tree-ensembles is not an established but an emerging practice in arid-climate forecasting, despite its reported advantages in related engineering fields (Mirjalili and Lewis 2016; Sansine et al. 2022). Fourth, the co-occurrence of T4 (dust-focused) and T5 (Iraq-specific) labels are not found in the corpus - no retained study uses dust-aware inputs and Iraqi datasets, a structural intersection further described in Section 3.3 and quantified as the Iraq Gap Index (IGI = 20.0).

Figure 1. PRISMA-ScR Flow Diagram: Study Selection Process

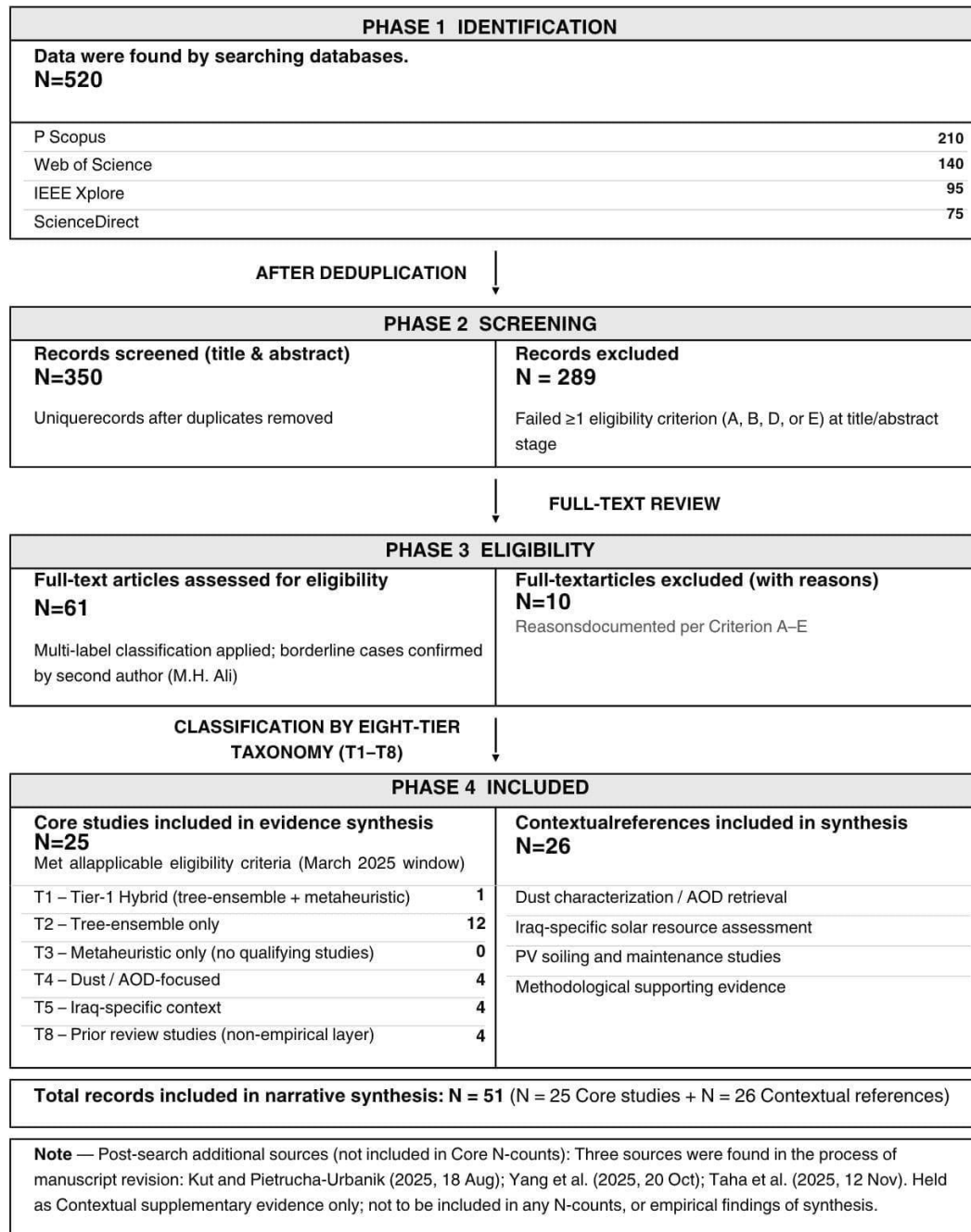


Figure 1 caption. PRISMA-ScR flow diagram illustrating the four-stage study selection process. Identification: 520 records retrieved across four databases (Scopus, Web of Science, IEEE Xplore, ScienceDirect). Screening: 350 unique records screened; 289 excluded. Eligibility: 61 full-text records assessed; 10 excluded with documented reasons. Included: N=25 Core studies and N=26 Contextual references (N=51 total). Screening by first author (F.A. Hadi); borderline cases reviewed by second author (M.H. Ali). PRISMA-ScR = Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (Tricco et al., 2018). AOD = aerosol optical depth.

Figure 1. PRISMA-ScR flow chart showing four-step study selection process database identification (n=520),

deduplication (n=350), title/abstract screening, full-text review, and N=25 Core + N=26 Contextual references identified.

3.2. Theoretical and Review Landscape (T — Prior Review Evidence)

The theoretical framework of evidence synthesis was based on four T8 prior review studies in the Core corpus. Al-Dahidi et al. (2024), sifted through 257 articles and proposed a seven-module framework of data-driven PV forecasting. Emerging hybrid components included tree-ensemble models, including XGBoost, Random Forest, GBM, and metaheuristic optimizers, including GA, PSO, GWO, and WOA, but dust and arid-climate conditions were not included in the picture.. Di Leo et al. (2025) followed the increasing adoption of PSO, GA, and Differential Evolution to hyperparameter optimization in ANN, SVM, and LSTM models, with the majority of the coverage in temperate, data-rich environments. Alcañiz et al. (2023) have found an eye-opening statistic: 22 of 100 PV forecasting studies were about dry or arid climates, and none of them focused on Iraq, which is perhaps the most obvious numerical manifestation of the uneven geographical distribution of the field. Iheanetu (2022) compared physical, statistical, and ML forecasting models, with competitive accuracy of KNN-similarity and NARX models, with MAE as low as 3.13% and RMSE of 2.3 among the best, but no systematic framework of dust-laden or data-scarce conditions.

There are three other reviews that provide a helpful background. Mohamad Radzi et al. (2023) surveyed ML-based PV forecasting over a variety of forecast horizons; tree-ensemble methods were common. However, studies of arid climates were scarce. The same observation was made by Chu et al. (2021) in another perspective: considering intra-hour irradiance forecasting, dust and aerosol dynamics hardly appeared, which is somewhat unexpected, considering the importance of these factors in short-term solar variability. The image that Maghami (2025) paints is even more grim, as the author has reviewed 860 papers on the impact of dust deposition on PV systems, and it was concluded that ML-based detection and prediction systems are far behind the magnitude of the very issue. These seven review studies, four of which are T8 Core studies (Al-Dahidi et al. 2024; Di Leo et al. 2025; Alcañiz et al. 2023). Iheanetu (2022) and three Contextual T8 references (Maghami 2025; Mohamad Radzi et al. 2023).

3.3. Contextual and Environmental Evidence from Arid and Dust-Affected Regions

Core and contextual studies were heading in the same direction, as far as dust, aerosol and soiling effects on solar systems in Iraq and similar arid conditions were concerned. Chaichan et al. (2023) recorded the actual effect of a severe dust storm on a PV system in Baghdad - output power decreased by approximately 62.3 percent compared to the baseline before the storm, solar irradiance was reduced by 54.5 percent, and the concentration of fine particles shot up. Radhi et al. (2024) used ANN models optimized with GA and GWO to short-term PV power prediction in Iraq - the only work in the literature that unites metaheuristic optimization and Iraq-based solar data. ANN-GA was superior in all the measured metrics, including the standard ANN and ANN-GWO versions, which is a fair sign that metaheuristic tuning does pay off in this specific data setting.

The context of the dust impact is further defined by supporting Iraq-specific environmental and operational evidence in the wider retained corpus. Hameed et al. (2024), whose main emphasis is on economic modelling, not forecasting, measured that poor cleaning frequency raised LCOE by over 30% in four Iraqi cities, dust-related soiling was the leading economic stressor to grid-connected PV, contextual motivation to dust-aware forecasting. Al-Humairi et al. (2025) approached the problem of predicting faults in PV systems under arid conditions in a maintenance-oriented manner, where the model was trained using ML-based techniques to

achieve 92.1% accuracy, which contributes to the increasing evidence that ML tools can be used successfully in hot-arid PV settings. Koprulu et al. (2025) conducted a field experiment under conditions similar to those encountered by Iraqi sites and investigated the effect of temperature, dust, and cement-coal dust as a combination of factors that reduce the efficiency of PV modules—the soiling losses they observed were large. These sources are maintained due to their environmental and operational applicability and are not included in the core forecasting denominators.

When one considers the entire retained corpus covering Iraq specifically, the image that comes out is geographically disjointed. Baghdad appears in seven papers — including dust impact quantification, soiling economics, and ML-based air quality prediction — but none of them use tree-ensemble forecasting with metaheuristic optimization and dust-aware inputs simultaneously. Mosul is mentioned in one study that confirms the MERRA-2 data pathway to XGBoost solar irradiance prediction (Sulaiman et al. 2025), but without integrating metaheuristic or dust-variable. A single ANN-metaheuristic study (Radhi et al. 2024) based on GA and GWO optimization, but on ANN architectures instead of tree-ensemble models, is provided in Amarah (Misan). A number of other Iraqi cities, such as Kirkuk, Fallujah, Basra, are also found in the Contextual corpus, primarily via soiling or local environmental research, and not via integrated forecasting pipelines. What there city-level data aggregation reveals are that the Iraq research gap are not merely a geographic gap, but rather a structural gap of the particular methodological mix that are of interest: no Iraqi site has so far been the focus of a study that combines tree-ensemble forecasting, metaheuristic optimization, and dust-sensitivity feature engineering.

To quantify this gap, this review proposes an Iraq Gap Index (IGI) which is the ratio of non-Iraqi to Iraq-specific empirical Core studies. $IGI = 20 / 1 = 20.0$. The only Core study that uses tree-ensemble methods directly on Iraqi solar generation data is the sole Iraqi study Sulaiman et al. (2025). The other three studies based in Iraq deal with dust effects, soiling economics or metaheuristic-ANN techniques and hence are not included in the IGI denominator. A 20.0 IGI is a harsh figure — it implies that every study that uses tree-ensemble forecasting on Iraqi solar data has twenty others that have done the same on other locations. Reducing that ratio to 5.0 or less would be a fair measure of actual improvement.

Similar arid and aerosol-rich areas provide additional evidence. Alonso-Pérez et al. (2025) used Random Forest for daily and hourly PV prediction in the Canary Islands in case of Saharan dust intrusion, with an accuracy of 97.58% at the daily level and 95.86% at the hourly level, having only four meteorological variables. This is promising, but the outcome is based on one dust episode and a source that is yet to be fully indexed on major databases. Bashir et al. (2025) approached the problem in a different way, creating a hybrid CNN-RF pipeline to detect dust on PV surfaces automatically by transforming the I–V curve parameters into RGB image matrices, enabling identification of five working states without aerosol measurements. A similar argument was expressed by Meng and Song (2020) in another direction: when the PM_{2.5} was directly inserted into a weather-classified Random Forest model, the daily PV forecasting accuracy was significantly enhanced in China when the conditions were polluted or cloudy. What all these studies are indicating is that aerosol and particulate variables are worth considering in tree-ensemble models - they have forecasting information that standard meteorological inputs are likely to overlook.

Sarafian et al. (2023), developed a multi-task learning model in depth to warn of dust events in the Middle East and North Africa. The most interesting finding of their results was that AOD had a greater predictive value than any other variable in predicting the onset of dust events with

useful lead times as long as 24 hours, something to remember when anyone is developing dust sensitive PV forecasting systems. Logothetis et al. (2024), conducted a comparison among five ML algorithms, namely LightGBM, Random Forest, MARS, KNN, and ANN to retrieve AOD based on ground irradiance data. LightGBM emerged as the obvious winner, reducing RMSE by two or three factors compared to MERRA-2 and CAMSRA reanalysis products. Lin et al. (2023) tackled the same issue on the input side: MAIAC MODIS AOD at 1 km resolution (instead of 2 km MERRA-2 AOD) reduced AOD RMSE by 0.03-0.08 at most AERONET locations, and the benefits were also transferred to DNI simulation accuracy. Another piece of information was provided by Panicker and Valarmathi (2025), a PSO-optimized SARIMA-SVR hybrid to predict AOD time-series, in which the metaheuristic tuning produced a stark contrast between the forecasting accuracy and the metaheuristic tuning. The combination of the evidence in these twelve studies provides a strong argument that dust and aerosol variability do imprint a quantifiable signature on both PV output and solar irradiance estimation in the MENA region and similar arid environments, and that the instruments to describe these effects, whether in the field or through ML, are already available.

3.4. Model Characteristics and Performance (C — Characteristics)

The 21 empirical core studies, with four T8 previous reviews, that set the theoretical background, had a considerable architectural variety in tree-ensemble implementations to predict solar generation. Rathika et al. (2025) used the Brown Bear Optimization Algorithm (BBOA) to automatically optimize the hyperparameters of a Random Forest model to achieve an MSE of 0.16614 and an R^2 of 0.818, which outperformed PSO-optimized and Firefly Algorithm-optimized baselines in ten independent runs of the model in confirmed hybrid configurations. Three other hybrid forms were found within a Contextual corpus: Sansine et al. (2022), used PSO to optimize XGBoost and LSTM hyperparameters simultaneously to predict probabilistic solar irradiance; Xu et al. (2025) developed an LSTM-XGBoost architecture with energy entropy decomposition and Seagull Optimizer; and Xiong et al. (2024) The Whale Optimization Algorithm (WOA), introduced by Mirjalili and Lewis (2016), is another possible candidate to integrate into metaheuristic-tree-ensemble with the previously verified algorithms.

Together, these four hybrid configurations reveal that tree-ensemble models form effective foundations of metaheuristic augmentation, and their underrepresentation in the arid climate Core corpus (one of 21 empirical studies (4.8%)) highlights the main research gap that this review will fill. Kut and Pietrucha-Urbanik (2025) reported an R^2 of 0.94 when using XGBoost to predict short-term PV energy under temperate conditions, with XGBoost performing better than comparators of Random Forest and linear regression, though this study is not counted in Core counts (Section 2.1) as it was published post-search and is considered Contextual evidence only. Sulaiman et al. (2025) used XGBoost and rolling window training schemes to GHI forecasting of Mosul with MERRA-2 satellite-derived inputs, with optimal 2-day rolling window settings of $R^2 = 0.907$ and $MAE = 546.6$ Wh. It is the only tree-ensemble forecasting study in the Core corpus that is directly applied to an Iraqi dataset. In a predictive maintenance workflow to model PV performance and detect faults, Marangis et al. (2024) used XGBoost, with a fault detection sensitivity of 96.9% — this is not a direct forecasting outcome, but contextual evidence of the predictive performance of XGBoost in utility-scale PV settings. Rinesh et al. (2024) introduced an XGB-SPPGP model to predicting solar PV power generation with the assistance of the IoT in data logging, with the overall accuracy of 99.3% and the accuracy ratio of 98.7%.

Models based on LightGBM yielded good results consistently. Yang et al. (2025) used Bayesian optimization and time-series cross-validation on LightGBM to predict PV over the medium

term, with $R^2=0.89$, $MAE=37.49$, and $RMSE=64.67$, showing the ability to capture high-dimensional nonlinear meteorological interactions, as a post-search reference, (Yang et al., 2020; Ye et al., 2023) introduced LX-PF, a pretrain-finetune LightGBM-XGBoost paradigm, with an accuracy of 96.46 percent over LSTM (79.02) and SVR (80.88) baselines. Boujoudar et al. stated that LightGBM was the most successful model (out of four models, CatBoost, LightGBM, XGBoost and RF) as per Chen et al., Abstract. Suanpang and Jamjuntr (2024) discovered that LightGBM is more effective than KNN in grid-scale comparisons under various forecast scenarios, Nguyen et al. (2025) evaluated Gradient Boosting, XGBoost, and LightGBM as models for predicting solar energy using boosting-based architectures, which showed performance improvements over ANN baselines.

In several studies, ensemble architectures that integrate tree-based models were studied. FRD-TLEL was presented by Wang et al. (2023) as the two layer ensemble that uses XGBoost, RF, CatBoost and decision tree base learners and has an MAPE of 2.81% on all test days (Table 5). Massaoudi et al. (2021) suggested an Enhanced Random Forest with Voted Feature Importance based on LIME, ElasticNet, and XGBoost, which is better than regular RF in all their weather subsets. RF, Extra Tree, and XGBoost showed $R^2 = 0.975$, $MAPE = 0.69$, and $MAE = 0.0130$ (units as reported in source) in rooftop PV forecasting. Yang et al. (2021) have developed an Improved RF with peak correction and DGM(1,1) grey prediction which is even more accurate by 1.48 percent and RMSE by 2.94 percent than the regular RF. Li et al. (2022) used XGBoost in an architecture of probabilistic solar irradiance, and the deterministic and probabilistic accuracy were better on a standard benchmark dataset (Contextual reference; temperate benchmark dataset, not part of Core N-counts). Gevorgian et al. (2024) showed that RF and XGBoost showed consistent performance in the state of data scarcity due to feature engineering and cross-validation techniques. Scott et al. (2023) compared various ML models on PV generation prediction and discovered that tree-ensemble methods were always competitive with deep learning methods at the same data scales. A hybrid AI ensemble framework that combines various tree-based model elements to achieve real-time adaptive optimization in solar energy systems was proposed in Muramshetti et al. (2025) and reported performance improvements over the baseline configurations; this work falls into T2 category within the Core corpus and the quantitative contributions of the study should be viewed with necessary caution due to the limited major-database indexing of the journal the results are based on (see The initial assessment of tree-based ensemble methods (Random Forest, Gradient Boosting, bagging) to solar radiation modelling at daily and hourly scales was conducted by Hassan et al. (2017), who showed that ensemble methods reliably performed better than MLP, SVR, and single decision tree comparators at multiple climatic stations; this baseline result established tree-ensemble methods as the main ML structure to be utilized in solar radiation Being published prior to 2020, (Hassan et al. 2017) is only referenced as a foundational reference and not included in any Core N-counts and denominator-based calculations (Criterion D).

3.5. Methodological Patterns and Data Source Characteristics (M — Methodological)

Empirical Cores Studies $N=21$ (without $N=4$ T8 precedent reviews) cross-study analysis demonstrates that there are a number of methodological patterns that reoccur. In terms of data sources: $N=18$ (85.7) of 21 empirical Core studies used ground-based measurement data as the input data pathway, and $N=3$ (14.3) empirical Core studies used a satellite derived/reanalysis product as the input data pathway. This pathway has been confirmed by $N=1$ Core study (Sulaiman et al. 2025), which explicitly used MERRA-2 reanalysis data as the main meteorological input to validate the data-scarce Iraqi conditions. $N=4$ of 21 empirical Core studies (19.0%) (Alonso-Pérez et al. 2025; Logothetis et al. 2024; Meng and Song 2020;

Panicker and Valarmathi 2025) used aerosol or AOD data as model inputs or study targets, and N=0 Core studies used tree-ensemble. In terms of forecast horizons in the N = 21 empirical Core studies: N = 1 of 21 (4.8%) studies covered ultra-short-term (minutes to 1 hour); N=13 of 21 (61.9%) covered short-term (day-ahead to 48 hours); and N = 7 of 21 (33.3%) covered medium-to-long-term. Regarding performance metrics: The RMSE was reported in N = 19 of 21 (90.5%) empirical Core studies; MAPE in N=15 of 21 (71.4%); R^2 in N=17 of 21 (81.0%); and MAE in N=16 of 21 (76.2%). This metric heterogeneity, along with the lack of standardized benchmark datasets to forecast in arid regions, restricts direct cross-study performance comparison, which is reported in previous reviews (Alcañiz et al. 2023; Al-Dahidi et al. 2024).

The most common type of tree-ensemble was the Random Forest, used in N=7 of 21 empirical core studies (33.3%). As a primary model or an important ensemble component (Alonso-Pérez et al. 2025; Rathika et al. 2025; Meng and Song 2020; Yang et al. 2021; Massaoudi et al. 2021). Random Forest as an independent model is also used in AOD retrieval (Logothetis et al. 2024). XGBoost was the second most common, featuring in N=6 of 21 (28.6%) (Sulaiman et al. 2025; Wang et al. 2023; Rinesh et al. 2024; Ye et al. 2023; Gevorgyan et al. 2024; Nguyen et al. 2025). LightGBM appeared as the primary model in N=2 of 21 (9.5%) studies (Boujoudar et al. 2024; Ye et al. 2023). Note: empirical studies that use multiple types of tree-ensemble models, like Gevorgyan et al. (2024) (RF and XGBoost) and Ye et al. (2023) (LightGBM and XGBoost), are counted in each of the relevant model types; the frequencies are not mutually exclusive and do not add to each other to the total number of 21 empirical core studies. The temporal distribution of the empirical Core corpus indicates a growing trend in the number of publications: N=4 in 2020-2021, N=5 in 2022-2023, and N=12 in 2024-2025, which is a 3-fold increase in the annual production of publications in the five-year period examined, and indicates the rapid development of tree-ensem Importance and interpretability analysis (SHAP values (Nguyen et al. 2025; Allal et al. 2024), or Voted Feature Importance (Massaoudi et al. 2021)) was used in N=3 of 21 empirical Core studies (14.3%), showing an emerging but nascent trend of model interpretability in tree-ensemble solar generation prediction

The above evidence synthesis reveals an apparent and recurrent structural gap in the retained corpus: no research has concurrently validated a tree-ensemble model with metaheuristic hyperparameter optimization in dust-laden arid environments with AOD or particulate matter as explicit model inputs. Section 4 puts these results into perspective in the context of the overall research environment and suggests a specific research agenda.

4. Discussion

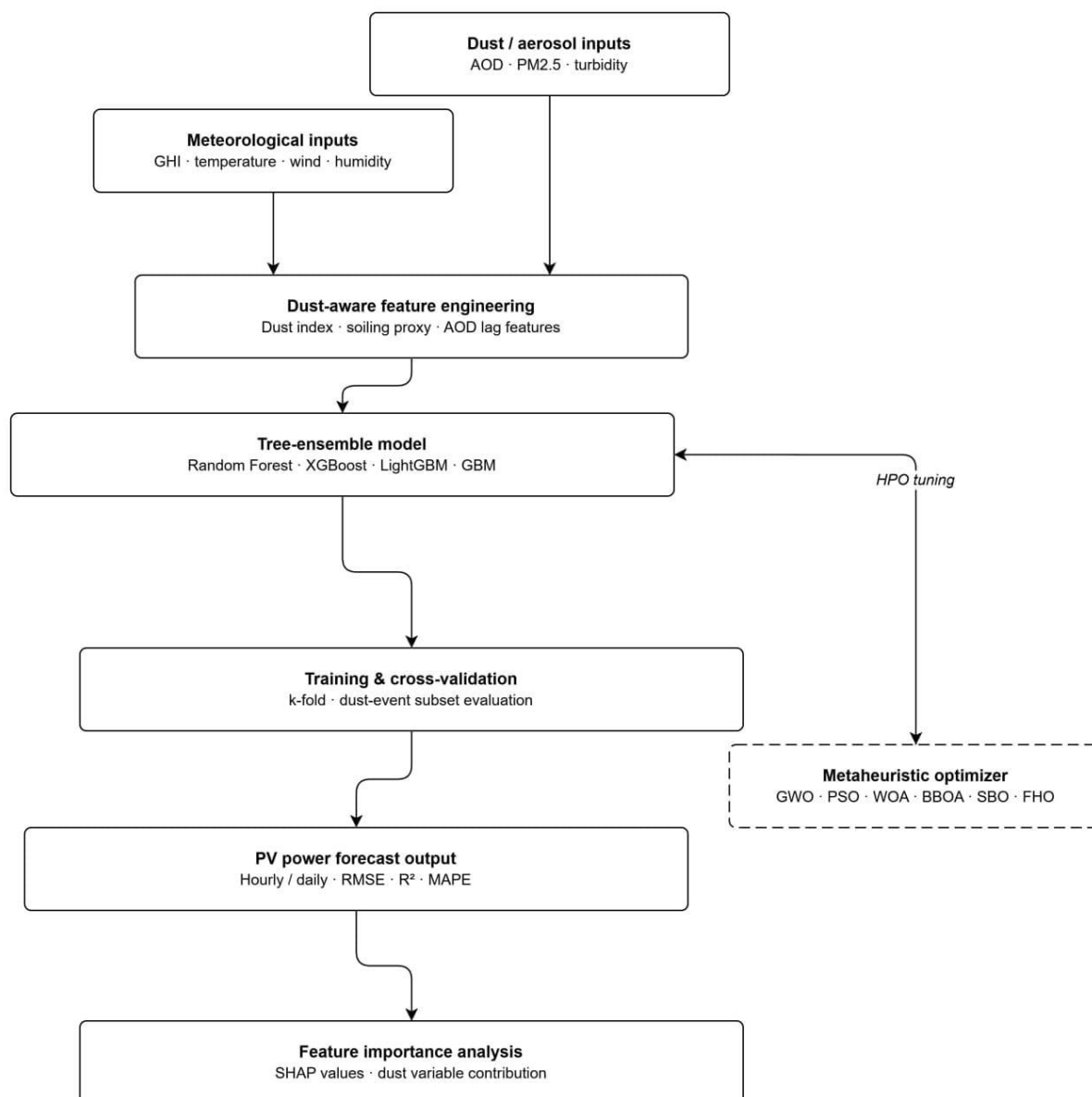


Figure 2. Conceptual diagram of the dust-aware hybrid tree-ensemble and metaheuristic solar PV power forecasting framework proposed as a research direction in this review.

4.1. Summary of Key Findings

N = 25 core studies were synthesized in this scoping review, and published within 2020–2025 on the topic of hybrid tree-ensemble and metaheuristic frameworks for predicting solar PV power generation and answering three research questions. To address RQ1, tree-ensemble algorithms, specifically Random Forest (N=7 of 21 empirical core studies (33.3%), XGBoost (N=6 of 21, 28.6%)) are the most common algorithm selections in the literature reviewed, but their combination with nature-inspired metaheuristic optimizers is uncommon. In answer to RQ2, N=4 of 21 empirical core studies (19.0%), which include dust storms as model inputs, are consistent with the field evidence of a limited number of measurement studies in Iraq dust storms significantly reduce the availability of solar irradiance by up to 54.5% and PV power generation by up to 62.3% (Chaichan et al. (2020)). To address RQ3, Iraq-specific tree-

ensemble forecasting is still in its infancy, with N=1 core study (Sulaiman et al., 2025) using XGBoost directly on an Iraqi solar irradiance dataset through MERRA-2 inputs, without metaheuristic optimization or dust-aware feature fusion. Reported performance values in retained studies are to be interpreted descriptively, in the sense that datasets, climates, forecast horizons, and evaluation protocols were never directly comparable across the corpus.

4.2. Comparison with Prior Reviews

The results of this review are more or less similar to, and go beyond, previous review evidence in three ways. To start with,, the prevalence of Random forest and XGBoost here is in line with the general PV forecasting taxonomy of Al-Dahidi et al. (2024), and Iheanetu (2022), who reported the use of the tree-ensemble models as the main constituents of current forecasting frameworks. But both previous reviews did not estimate the rate of metaheuristic combination with the tree-ensemble models in particular, which is bridged by this review by providing a baseline rate of N=1 of 21 empirical core studies (4.8%) in the arid-relevant corpus. Second, the geographic imbalance reported by Alcañiz et al. (2023) - who discovered that only 22.0% of pre-2020 PV forecasting studies had considered dry or arid climates in general, and none had considered Iraq specifically - continues to exist in the current cor Third, the observation that N=0 Core studies concurrently used tree-ensemble forecasting, metaheuristic optimization, and aerosol inputs is beyond what any previous review has measured, making this three-way lack a structural characteristic of the existing research space. A result that is somewhat contradictory to the evidence of previous reviews is the relative ranking of LightGBM in the corpus. Both Di Leo et al. (2025) and Mohamad Radzi et al. (2023) forecasted a rise in LightGBM use, but LightGBM is the main model in only N=2 of 21 empirical Core studies (9.5%) in this review, which is significantly lower than XGBoost (28.6%) and Random Forest (33.3%). This could be due to the arid-climate and dust-context filtering used in this review, which could prefer established models with a longer history of validation in data-limited environments, or it could be that LightGBM is not being adopted as quickly in arid-region forecasting as it is being adopted in temperate contexts.

4.3. Unexpected Findings

This review had three findings that were surprising compared to the previous literature. To begin with, the almost total lack of validated Tier-1 hybrid configurations (N=1 of 21 empirical Core studies, 4.8%) was worse than expected. This result is quantitatively confirmed by a systematic cross-tabulation of the retained corpus: of the N=51 retained studies (N=25 Core and N=26 Contextual), N=0 also validated a tree-ensemble model with a nature-inspired metaheuristic optimizer when dust or aerosol variables (AOD, PM2.5, turbidity) were explicitly modeled. The nearest analogue in the Iraqi context is the study by Sarafian et al. (2023), who used AOD-based dust event prediction with deep learning, and Sulaiman et al. (2025), who used XGBoost to predict Iraqi solar irradiance (GHI) data using MERRA-2 - but neither of them used all three components. It is noteworthy that a hybrid COA-optimized Random Forest model was used to predict PM2.5 concentration in Baghdad (Shehab et al. 2026, retained as Contextual evidence outside the 2020-2025 primary search window) - showing that metaheuristic tree-ensemble optimization can be technically applied to the Iraqi atmospheric context - but this was applied to air quality prediction Since metaheuristic optimization of tree-ensemble models is well-reported in related fields of engineering (Mirjalili and Lewis 2016), and some earlier reviews (Al-Dahidi et al. 2024; Di Leo et al. 2025) explicitly reported metaheuristic approaches and tree-ensemble models as a new hybrid element, it was anticipated that a higher proportion of studies would have The observation that Bayesian optimization and grid search are the most popular hyperparameter tuning strategies in the corpus, though also

effective, are not metaheuristic methods, by the predefined criterion B - means the difference between how these methods were classified in previous reviews and how this review used a more demanding criterion of metaheuristic methods.

Second, the good performance of the BBOA-RF hybrid (Rathika et al. 2025) in comparison to PSO-RF and Firefly-RF baselines in a single study indicates that newer nature-inspired algorithms might provide competitive or superior hyperparameter tuning to tree-ensemble models than well-established benchmarks (PSO, GA) - a result that should be systematically replicated across multiple tree. Third, the finding that XGBoost with MERRA-2 inputs achieves R^2 of 0.907 for Mosul, Iraq GHI forecasting (Sulaiman et al. 2025) without any Iraq-specific meteorological calibration or dust-variable inputs suggests that the MERRA-2 pathway may be more viable for Iraqi forecasting than the sparse ground-station literature alone would suggest, providing a potentially transferable baseline that warrants further validation for future dust-aware model development.

4.4. Theoretical Implications

The structural gap identified in this review; i.e. the absence of a study that combines tree-ensemble forecasting, metaheuristic optimization, and dust-aware aerosol inputs, has theoretical implications for the field. It implies that the benefit of metaheuristic optimization that has been reported in similar contexts (Sansine et al., 2022; Xu et al., 2025; Rathika et al., 2025; Mohamad Radzi et al., 2023) has not been applied to arid-climate tree-ensemble forecasting. This propagation failure can be indicative of the compound data needs of such a system: both demanding high-quality aerosol inputs and enough training data to generalize tree-ensembles and metaheuristic convergence with small sample sizes. In theory, this inspires the development of sample-efficient optimization models and transfer learning methods that may reduce the data burden of metaheuristic integration in data-sparse arid settings. The fact that ML-retrieved AOD is a much better model input than reanalysis AOD (Logothetis et al., 2024) also indicates that the quality of aerosol input, and not the presence of aerosol variables, can be a determining factor in the performance of dust-aware forecasting models.

4.5. Practical Implications

The practical implications of this review are concentrated in three operational domains. For solar energy planners and grid operators in Iraq and the MENA region, the finding that XGBoost with MERRA-2 inputs achieves R^2 of 0.907 for Mosul GHI forecasting (Sulaiman et al. 2025) without requiring ground measurement infrastructure provides a potentially transferable baseline that warrants further validation under Iraq-specific conditions. The results of Koprulu et al. (2025), Al-Humairi et al. (2025), and Hameed et al. (2024) are all indicative of the notion that optimization of cleaning schedules based on the predictions of dust events informed by ML models would be beneficial to reducing LCOE in Iraqi PV systems, although the direct validation of operational pipeline is an unfinished research area. To the designers of PV systems, the observation that dust events decrease the irradiance by 54.5% (Chaichan et al. 2023) and output by 62.3% under favorable conditions suggests that dust-conservative sizing margins must be used during the design of systems in arid regions where forecasting-based dispatch is to be implemented. To algorithm developers, the good BBOA-RF and PSO-XGBoost performance (Rathika et al. 2025; Xu et al. 2025) and the SO-XGBoost results (Sansine et al. 2022) offer candidate starting configurations to metaheuristic-tree-ensemble hyperparameter optimization. These settings should be tested on Iraq-specific and dust-prone data conditions and only when these settings are validated then they can be widely used.

4.6. Research Agenda

Table 2 is a structured four-component research agenda, based on the identified gaps, to predict dust-aware hybrid tree-ensemble solar generation in Iraq and similar arid regions.

Priority	Research Action	Evidence Base	Timeline
RA-1	Dust-aware feature engineering: systematically evaluate AOD, PM2.5, PM10, and turbidity as inputs to tree-ensemble forecasting models trained on Iraqi datasets.	Logothetis et al. (2024) and Lin et al. (2023) show ML-retrieve and MAIAC-corrected AOD outperforms MERRA-2 AOD; Sarafian et al. (2023) confirm AOD as leading dust event predictor.	Near-term (1–2 years)
RA-2	Metaheuristic optimizer evaluation: systematically compare GWO, PSO, WOA, BBOA, and hybrid swarm algorithms for XGBoost, RF, and LightGBM hyperparameter tuning under arid-region data conditions.	Rathika et al. (2025) (BBOA-RF), Sansine et al. (2022) (PSO-XGBoost), Xu et al. (2025)(SO-XGBoost), Mirjalili and Lewis(2016)(WOA) establish viable starting configurations.	Near-term (1–2 years)
RA-3	Iraq-specific dataset development: extend the Mosul MERRA-2 XGBoost pipeline (Sulaiman et al. 2025) to Baghdad, Basra, and Nasiriya, incorporating MAIAC AOD inputs and dust event metadata from satellite records (Chaichan et al. 2023; Logothetis et al. 2024).	Sulaiman et al. (2025) validate MERRA-2 as a viable data pathway for Iraq; Attiya and Jones (2020) and Ismail et al. (2024) provide comprehensive dust event climatology for multi-location validation.	Medium-term (2–3 years)
RA-4	Standardized benchmarking: develop a benchmark dataset and evaluation protocol for dust-aware arid-region PV forecasting, incorporating standardized metric reporting (RMSE, MAPE, R^2) and dust-event subset evaluation.	Metric heterogeneity across N=21 empirical Core studies (RMSE reported N=19/21 (90.5%), MAPE N=15/21 (71.4%)) and absence of shared benchmark datasets limit cross-study comparison (Alcañiz et al. 2023; Al-Dahidi et al. 2024).	Long-term (3–5 years)

Table 2. Dust-conscious hybrid tree-ensemble solar generation prediction in Iraq and arid areas. RA = Research Action; HPO = hyperparameter optimization; MAIAC = Multi-Angle Implementation of Atmospheric Correction.

To support the research agenda presented in Table 2, and to help future researchers in Iraq and other data-scarce arid regions, Table 3 presents six recommended open-access datasets, which collectively provide solar irradiance, aerosol optical depth, dust event records, and PV performance data. All six datasets are cited in this review, and have shown appropriateness for solar resource assessment for Iraq or the MENA region.

Dataset	Provider	Key variables	Access	Citation in this review
NASA POWER	NASA	GHI, DNI, DHI, temperature, wind speed, RH	power.larc.nasa.gov	NASA POWER 2025; Section 1

Dataset	Provider	Key variables	Access	Citation in this review
MERRA-2	NASA GSFC	GHI, AOD, T2M, wind speed, humidity	disc.gsfc.nasa.gov	Sulaiman et al. 2025; Sections 3.4, 3.5
MAIAC/MODIS AOD	NASA	AOD at 550 nm, 1 km resolution	ladsweb.modaps.eosdis.nasa.gov	Lin et al. 2023; Section 3.3, RA-1
AERONET	NASA	AOD Level 2.0, Ångström exponent	aeronet.gsfc.nasa.gov	Holben et al. 1998; Section 1
PVGIS	JRC / EU	GHI, PV yield, SARA-2 irradiance	re.jrc.ec.europa.eu/pvgis	Section 3.3 (dataset context)
Global Solar Atlas	World Bank	GHI, DNI, PVOU long-term averages	globalsolaratlas.info	World Bank / Global Solar Atlas 2023

Table 3. Recommended open-access datasets for dust-aware hybrid tree-ensemble solar PV forecasting in Iraq and arid regions. All datasets are cited within this review. GHI = Global Horizontal Irradiance; DNI = Direct Normal Irradiance; DHI = Diffuse Horizontal Irradiance; AOD = Aerosol Optical Depth; RH = Relative Humidity; T2M = Temperature at 2 metres; PVOU = PV power output potential.

4.7. Limitations of This Review

There are five methodological limitations of this review. First, the first author did primary title and abstract screening; only borderline cases were subjected to dual-independent screening, which was done by the second author (M.H. Ali). This incomplete design increases the chances of selection bias compared to completely independent two-screening regimes. Second, five sources that were published after the main search period were detected and included as contextual supplementary evidence only; they are not included in any of the Core N-counts and empirical synthesis results, as it is made known in Section 2.1. They include three references found during the revision of the manuscript (Kut and Pietrucha-Urbanik (2025); Yang et al. (2025); Taha et al. (2025)), one reference that was found outside the 2020-2025 search (Shehab et al. (2026)), and one reference that was referenced as contextual evidence during the revision of the introduction. Third, the English-only, four-database search might not have identified publication or grey literature on Iraq-specific solar monitoring data, which may have underestimated the Iraq-specific evidence base. Fourth, the quality of individual studies was not critically evaluated, which is also consistent with the scoping review approach (Peters et al. 2020); comparisons of performances are to be considered evidence mapping, but not claims about effects. Fifth, the conservative metaheuristic criterion (Criterion B, and excluding both Bayesian optimization and grid search) is by definition conservative and would have a Tier-1 count of N=1 in the arid-relevant Core corpus; a more generous one that allows Bayesian optimization would most probably have a larger count of N=4-5. Across all five limitations, the principal finding — that no study within the retained corpus simultaneously validated a dust-aware hybrid tree-ensemble and metaheuristic framework under Iraq-specific conditions — remains consistent, and is supported by convergent evidence from N=25 Core and N=26 Contextual references.

The complete study selection process, including sequential screening stages and record counts consistent with the PRISMA-ScR reporting standard (Tricco et al. 2018), is described in Section 3.1.

5. Conclusion

This scoping review has identified N=25 Core studies and N=26 Contextual references that discuss hybrid tree-ensemble and metaheuristic architectures to solar PV power generation forecasting, with a specific focus on arid-climate applicability and research gaps specific to Iraq. Four key findings are made. To begin with, the most common algorithmic selection in the literature reviewed is tree-ensemble models, specifically Random Forest (N=7 of 21 empirical Core studies, 33.3%) and XGBoost (N=6 of 21, 28.6%): both exhibit consistently competitive performance in heterogeneous, non-comparable study settings. Second, metaheuristic optimization is still significantly underused in conjunction with tree-ensemble approaches to solar generation prediction, where N=1 of 21 empirical Core studies (4.8%) report a Tier-1 hybrid configuration in the arid-relevant corpus, although it has been shown to be beneficial in more general engineering settings (Sansine et al. 2022; Xu et al. 2020). Third, the model includes dust storms and aerosol variables as model inputs in only N=4 of 21 empirical Core studies (19.0%), even though field measurements show that dust storms cut solar irradiance by 54.5% and PV output by 62.3% in Iraq (Chaichan et al. 2023) and that ML-retrieved AOD is a Fourth, tree-ensemble predictions of Iraq-specific data are still in their early phases, with N=1 Core study (Sulaiman et al. 2025) defining the MERRA-2 data pathway of Mosul solar irradiance predictions but not optimized by metaheuristics or dust-aware feature combination.

The review has four contributions to the field. First, it offers a systematic scoping evidence map of tree-ensemble and metaheuristic solar generation prediction literature that is specifically filtered to be arid-climate and dust-aware (2020-2025), which has not been previously reviewed (Alcañiz et al. 2023; Al-Dahidi et al. 2024; Di Leo et al. 2025). Second, it proposes and applies an eight-level analytical taxonomy designed to aid this review (T1-T8) to classify hybrid generation prediction architectures which can be reused in systematic or scoping reviews of this area in future. Third, it measures the gap in the evidence of Iraq by decomposing the gap by city and provides a reproducible gap metric (Iraq Gap Index, IGI = 20.0), thus providing a baseline in terms of which research can be tracked. Fourth, it officially recognizes a structural three-way gap - no paper has implemented tree-ensemble forecasting in conjunction with metaheuristic optimization and aerosol inputs - and suggests a four-part actionable research agenda to address this gap. All these results indicate that there exists an apparent recurrent gap in the retained corpus: none of the studies identified in this review have validated a dust-conscious hybrid tree-ensemble and metaheuristic solar generation prediction model in the Iraq-specific environmental and data conditions. Table 2 provides a systematic research agenda that would provide a prioritized line of penetration into this gap by dust-aware feature engineering, systematic metaheuristic optimizer testing, creation of Iraq-specific datasets, and standardized benchmarking. This review should be viewed in the context of the five methodological limitations outlined in Section 4.7: partial dual-independent screening; five post-search and post-window references included as supplementary contextual evidence only; English-only database search; no formal quality appraisal; and the strict metaheuristic criterion that can undercount hybrid configurations in a broader definition of optimization. This review presents a systematic descriptive evidence base on the need to focus future dust-conscious forecasting studies in Iraq and other arid settings.

Acknowledgments

The authors acknowledge the support of the Electrical Engineering Techniques Department, Mussaib Technical College, Al-Furat Al-Awsat Technical University, Kufa, Iraq.

Disclosure

The authors report no conflicts of interest in this work. No external funding was received for this study.

Ethics Statement

This is a secondary scoping review of peer-reviewed literature that is publicly available. It did not entail human subjects, animal subjects or primary data gathering. None of the ethics approval were needed.

Data Availability Statement

The evidence to support the results of this review is all provided by publicly-available peer-reviewed literature and publicly available institutional data sources (such as NASA POWER, AERONET, and Global Solar Atlas) listed in the reference list. The data extraction scheme and the informative charted data can be obtained on request to the respective author.

References

- Al-Dahidi S, Madhiarasan M, Al-Ghussain L, Abubaker AM, Ahmad AD, Alrbai M, et al. Forecasting solar photovoltaic power production: a comprehensive review and innovative data-driven modeling framework. *Energies*. 2024;17(16):4145.
- Al-Humairi A, Khalis E, Al Hemyari ZA, Jung P. Machine learning-based predictive maintenance for photovoltaic systems. *AI*. 2025;6(7):133.
- Al-Rikabi IJ, Omara AAM, Abuelnour MA, Abdul-Zahra AS, Al Jubori AM, Alsaad H. Energy landscape in Iraq: current status, research review, and policy insights. *Energy Sci Eng*. 2025;14:625–678. <https://doi.org/10.1002/ese3.70359>
- Alcañiz A, Grzebyk D, Ziar H, Isabella O. Trends and gaps in photovoltaic power forecasting with machine learning. *Energy Rep*. 2023;9:447–471.
- Allal Z, Noura HN, Chahine K. Machine learning algorithms for solar irradiance prediction: a recent comparative study. *e-Prime Adv Electr Eng Electron Energy*. 2024;7:100453. <https://doi.org/10.1016/j.prime.2024.100453>
- Alonso-Pérez S, López-Solano J, Montes C, Ocaña L, Llarena E, González-Díaz B. Assessment of random forest algorithm efficacy for photovoltaic production prediction during Saharan dust intrusion episodes: a case study including wet deposition. *Renew Energy Environ Power Qual J*. 2025;3:36–43.
- Attiya AA, Jones BG. Climatology of Iraqi dust events during 1980–2015. *SN Appl Sci*. 2020;2(5):845.
- Bashir SB, Farag MM, Hamid AK, Adam AA, Bansal RC. Innovative dust detection and efficient cleaning of photovoltaic panels using CNN–RF machine learning. *Energy Convers Manag*. 2025;27:101079.
- Boujoudar M, Bouarfa I, Dadda A, Elydrissi M, et al. Power PV forecasting using machine learning algorithms based on weather data in semi-arid climate. *BIO Web Conf*. 2024;109:01024. <https://doi.org/10.1051/bioconf/202410901024>

- Chaichan MT, Kazem HA, Al-Waeli AH, Sopian K, Fayad MA, Alawee WH, et al. Sand and dust storms' impact on the efficiency of photovoltaic modules installed in Baghdad: a review study with an empirical investigation. *Energies*. 2023;16(9):3938. <https://doi.org/10.3390/en16093938>
- Chu Y, Li M, Coimbra CFM, Feng D, Wang H. Intra-hour irradiance forecasting techniques for solar power integration. *iScience*. 2021;24:103136.
- Di Leo P, Ciocia A, Malgaroli G, Spertino F. Advancements and challenges in photovoltaic power forecasting: a comprehensive review. *Energies*. 2025;18(8):2108.
- Gevorgian A, Pernigotto G, Gasparella A. Addressing data scarcity in solar energy prediction with machine learning and augmentation techniques. *Energies*. 2024;17(14):3365. <https://doi.org/10.3390/en17143365>
- Global Solar Atlas / World Bank. Iraq: solar irradiation and PV power potential maps [Internet]. 2023 [cited 2026 Mar 11]. Available from: <https://globalsolaratlas.info/download/iraq>
- Hameed MA, et al. Economic consequences based on reversible and irreversible degradation of PV park in the harsh climate conditions of Iraq. *Energies*. 2024;17(11):2652.
- Hassan MA, Khalil A, Kaseb S, Kassem MA. Exploring the potential of tree-based ensemble methods in solar radiation modeling. *Appl Energy*. 2017;203:897–916. <https://doi.org/10.1016/j.apenergy.2017.06.104>
- Holben BN, Eck TF, Slutsker I, Tanre D, Buis JP, Setzer A, et al. AERONET — a federated instrument network and data archive for aerosol characterization. *Remote Sens Environ*. 1998;66(1):1–16.
- Iheanetu KJ. Solar photovoltaic power forecasting: a review. *Sustainability*. 2022;14(24):17005.
- International Energy Agency. Global Energy Review 2025 [Internet]. Paris: IEA; 2025 [cited 2026 Apr 8]. Available from: <https://www.iea.org/reports/global-energy-review-2025>
- Ismail RR, Halos SH, Al-Abudi BQ. Detection of the most frequent sources of dust storms in Iraq during 2020–2023 using space tools. *Kuwait J Sci*. 2024;52(1):100328.
- Koprulu A, Kareem YK, Yaseen A, Awad AT. The effects of temperature, dust, and cement-coal particles on PV panel performance: a case study in Iraq. *Eurasian J Sci Eng*. 2025;11(2):357–367.
- Kut P, Pietrucha-Urbanik K. Forecasting Short-Term Photovoltaic Energy Production to Optimize Self-Consumption in Home Systems Based on Real-World Meteorological Data and Machine Learning. *Energies*. 2025;18(16):4403. <https://doi.org/10.3390/en18164403>
- Li X, Ma L, Chen P, Xu H, et al. Probabilistic solar irradiance forecasting based on XGBoost. *Energy Rep*. 2022;8:1087–1095.
- Lin C-A, Zhang Y, Heath G, Henze DK, Sengupta M, Lu C-H. Improvement of aerosol optical depth data for localized solar resource assessment. *Sol Energy*. 2023;249:457–466. <https://doi.org/10.1016/j.solener.2022.11.047>
- Logothetis S-A, Salamalikis V, Kazantzidis A. A machine learning approach to retrieving aerosol optical depth using solar radiation measurements. *Remote Sens*. 2024;16(7):1132. <https://doi.org/10.3390/rs16071132>

- Maghami MR. Impact of dust deposition on photovoltaic systems and mitigation strategies. *Technologies*. 2025;14(1):15. <https://doi.org/10.3390/technologies14010015>
- Marangis D, Livera A, Tziolis G, Makrides G, Kyprianou A, Georghiou GE. Trend-based predictive maintenance and fault detection analytics for photovoltaic power plants. *Solar RRL*. 2024;8(24):2400473. <https://doi.org/10.1002/solr.202400473>
- Massaoudi M, Chihi I, Sidhom L, Trabelsi M, Refaat SS, Oueslati FS. Enhanced random forest model for robust short-term photovoltaic power forecasting using weather measurements. *Energies*. 2021;14(13):3992. <https://doi.org/10.3390/en14133992>
- Meng M, Song C. Daily photovoltaic power generation forecasting model based on random forest algorithm for North China in winter. *Sustainability*. 2020;12(6):2247. <https://doi.org/10.3390/su12062247>
- Mirjalili S, Lewis A. The whale optimization algorithm. *Adv Eng Softw*. 2016;95:51–67.
- Mohamad Radzi PNL, Akhter MN, Mekhilef S, et al. Review on the application of photovoltaic forecasting using machine learning for very short to long-term forecasting. *Sustainability*. 2023;15:2942.
- Muramshetti SC, Kunal K, Murugadoss R, Madeshwaren V. Hybrid AI ensemble for real-time adaptive optimization in solar energy systems. *J Intell Syst Internet Things*. 2025;17(2):278–294. <https://doi.org/10.54216/JISIoT.170218>
- NASA POWER Project. Prediction of Worldwide Energy Resources (POWER) [Internet]. 2025 [cited 2026 Mar 11]. Available from: <https://power.larc.nasa.gov/>
- Nguyen HN, Tran QT, Ngo CT, Nguyen DD, Tran VQ. Solar energy prediction through machine learning models: a comparative analysis of regressor algorithms. *PLOS ONE*. 2025;20:e0315955.
- Oruç S, Hınıs MA, Tuğrul T. Forecasting photovoltaic power in high-latitude regions via support vector machine optimized by meta-heuristics. *Sci Rep*. 2026;16:3438. <https://doi.org/10.1038/s41598-025-33415-7>
- Panicker NK, Valarmathi J. Time series prediction of aerosol optical depth across the northern Indian region: integrating PSO-optimized SARIMA-SVR based on MODIS data. *Acta Geophys*. 2025;73:2097–2126. <https://doi.org/10.1007/s11600-024-01472-7>
- Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. *JBIM Evid Synth*. 2020;18(10):2119–2126.
- Radhi SM, Al-Majidi SD, Abbod MF, Al-Raweshidy HS. Machine Learning Approaches for Short-Term Photovoltaic Power Forecasting. *Energies*. 2024;17(17):4301. <https://doi.org/10.3390/en17174301>
- Rathika SK, Meera PS, Lavanya V, Hemamalini S. Brown bear optimized random forest model for short-term solar power forecasting. *Results Eng*. 2025;25:104583. <https://doi.org/10.1016/j.rineng.2025.104583>
- Rinesh S, Deepa S, Nandan RT, Sachin RS, Thamil SV, et al. Prediction and classification of solar photovoltaic power generation using extreme gradient boosting regression model. *Int J Low-Carbon Technol*. 2024;19:2420–2430.
- Saini RK, Saini DK, Gupta R, Verma P, Saini HS. Effects of dust on the performance of solar panels: a review update from 2015–2020. *Energy Environ*. 2023;34(6):2110–2162.

- Sansine V, Ortega P, Hissel D, Hopuare M. Solar irradiance probabilistic forecasting using machine learning, metaheuristic models and numerical weather predictions. *Sustainability*. 2022;14(22):15260. <https://doi.org/10.3390/su142215260>
- Sarafian R, Nissenbaum D, Raveh-Rubin S, Agrawal V, et al. Deep multi-task learning for early warnings of dust events in the Middle East and North Africa. *npj Clim Atmos Sci*. 2023;6:23.
- Scott C, Ahsan M, Albarbar A. Machine learning for forecasting a photovoltaic (PV) generation system. *Energy*. 2023;278:127807. <https://doi.org/10.1016/j.energy.2023.127807>
- Shehab ZN, Algamal ZY, Faisal RM. Forecasting PM2.5 daily concentration in Baghdad, Iraq based on improving random forest algorithm. *Iraqi Natl J Earth Sci (INJES)*. 2026;26(1):180–192. <https://doi.org/10.33899/injes.v26i1.60200>
- Singh U, Singh S, Gupta S, Alotaibi MA, Malik H. Forecasting rooftop photovoltaic solar power using machine learning techniques. *Energy Rep*. 2025;13:3616–3630. <https://doi.org/10.1016/j.egyr.2025.03.005>
- Suanpang P, Jamjuntr P. Machine learning models for solar power generation forecasting in microgrid application implications for smart cities. *Sustainability*. 2024;16(14):6087. <https://doi.org/10.3390/su16146087>
- Sulaiman A, Mahmood FE, Hamzah SK. Evaluating the impact of window size on model accuracy in solar irradiance forecasting using XGBoost. *Math Model Eng Probl*. 2025;12(9):3177–3185.
- Taha A, Makeen P, Nazih N. Short-term and long-term solar irradiance forecasting with advanced machine learning techniques in Zafarana, Egypt. *Sci Rep*. 2025;15:39553. <https://doi.org/10.1038/s41598-025-24853-4>
- Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med*. 2018;169(7):467–473.
- Wang H, Yan S, Ju D, Ma N, Fang J, Wang S, et al. Short-term photovoltaic power forecasting based on a feature rise-dimensional two-layer ensemble learning model. *Sustainability*. 2023;15(21):15594. <https://doi.org/10.3390/su152115594>
- Xiong G, Zhang J, Fu X, Chen J, Mohamed AW. Seasonal short-term photovoltaic power prediction based on GSK–BiGRU–XGBoost considering correlation of meteorological factors. *J Big Data*. 2024;11:164. <https://doi.org/10.1186/s40537-024-01037-x>
- Xu Y, Ji X, Zhu Z. A photovoltaic power forecasting method based on the LSTM-XGBoost-EEDA-SO model. *Sci Rep*. 2025;15:30177. <https://doi.org/10.1038/s41598-025-16368-9>
- Yang M, Zhao M, Liu D, Ma M, Su X. Improved random forest method for ultra-short-term prediction of the output power of a photovoltaic cluster. *Front Energy Res*. 2021;9:749367. <https://doi.org/10.3389/fenrg.2021.749367>
- Yang Y, Lee S-H, Choi Y-S, Lee K-M. LightGBM medium-term photovoltaic power prediction integrating meteorological features and historical data. *Energies*. 2025;18(20):5526. <https://doi.org/10.3390/en18205526>
- Ye J, Zhao B, Deng H. Photovoltaic power prediction model using pre-train and fine-tune paradigm based on LightGBM and XGBoost. *Procedia Comput Sci*. 2023;224:407–412. <https://doi.org/10.1016/j.procs.2023.09.056>