Mohammad Ehteram
Akram Seifi
Fatemeh Barzegari Banadkooki

Application of Machine Learning Models in Agricultural and Meteorological Sciences

Springer

Application of Machine Learning Models
in Agricultural and Meteorological Sciences

Mohammad Ehteram · Akram Seifi ·
Fatemeh Barzegari Banadkooki

Application of Machine Learning Models in Agricultural and Meteorological Sciences

Springer

Mohammad Ehteram
Department of Water Engineering
and Hydraulic Structures
Faculty of Civil Engineering
Semnan University
Semnan, Iran

Akram Seifi
Department of Water Science
and Engineering, College of Agriculture
Vali-e-Asr University of Rafsanjan
Rafsanjan, Iran

Fatemeh Barzegari Banadkooki
Agricultural Department
Payame Noor University
Tehran, Iran

ISBN 978-981-19-9732-7 ISBN 978-981-19-9733-4 (eBook)
https://doi.org/10.1007/978-981-19-9733-4

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

Meteorological and agricultural predictions are required for water resource and planning management. Meteorological predictions can be used for predicting natural hazards such as flood and drought periods. Meteorological and agricultural predictions have a complex and nonlinear process. Thus, robust models are required for meteorological and agricultural predictions. This book uses robust machine learning algorithms for meteorological and agricultural predictions. Also, this book introduces new optimization algorithms for training machine learning models. First, the book introduces the structure of optimization algorithms for training machine learning models. Afterward, the structures of machine learning models are explained. Also, the structures of optimized machine learning models are explained. This book uses machine learning models to predict meteorological and agricultural variables. Different case studies are explained to evaluate the new machine learning models' ability to predict meteorological and agricultural variables. The decision-makers can use the current book for managing watersheds. Also, scholars can use the current book to develop hydrological science models.

Semnan, Iran — Mohammad Ehteram
Rafsanjan, Iran — Akram Seifi
Tehran, Iran — Fatemeh Barzegari Banadkooki
October 2022

Contents

Chapter 1
The Importance of Agricultural and Meteorological Predictions Using Machine Learning Models

Abstract This chapter reviews the applications of machine learning (ML) models for predicting meteorological and agricultural variables. The advantage and disadvantages of models are explained. This chapter also explains the importance of meteorological and agricultural predictions for water resource planning and management. The details of different machine learning models are explained. Afterward, the applications of these models are described. The ML includes different methods for learning predictive rules from historical datasets to predict unknown future data. Several studies have reported the superiority of ML techniques in agricultural and weather predictions that can maximize agricultural profit.

Keywords Optimization algorithms · Agriculture systems · Machine learning models · Water resource management

1.1 Introduction

Accurate meteorological prediction information is crucial to recognize the state of atmospheric conditions and prevent/reduce destructive effects of agricultural crises and disasters, such as tornadoes, floods, and major storms, from affecting people and property. Using meteorological variables for weather forecasting and climate prediction is essential for many industries, especially agriculture. The main aim of weather predicting models is to predict some meteorological variables, including air temperature, relative humidity, rainfall, dew point temperature, solar radiation, sunshine hours, and wind speed, based on the historical dataset as decision support systems in agriculture (Jaseena & Kovoor, 2020; Ghanbari-Adivi et al., 2022; Farrokhi et al., 2021).

Insufficient knowledge related to soil (soil type, soil temperature, water content, salinity, spatial and temporal variations) and crop (yield, water requirement, evapotranspiration, use of pesticides, harvesting, and marketing) properties, weather, and irrigation problems can cause farmers losing their farms and reduce expected efficiency. Following the famous saying "Information is Power", farmers may make better decisions by tracking crop, market, and environment information (Meshram et al., 2021; Farrokhi et al., 2020). Prior meteorological information assists farmers

M. Ehteram et al., *Application of Machine Learning Models in Agricultural and Meteorological Sciences*, https://doi.org/10.1007/978-981-19-9733-4_1

in improving crop yields by enabling them to make necessary decisions (Salman et al., 2015). Hence, predicting agriculture information, such as crop yield, early and accurately contributes significantly to facilitating decision-making and a country's growth (Sawasawa, 2003). Figure 1.1 shows the important parameters in studies related to agriculture, weather and climate, and irrigation problems.

In the last three decades, predicting agriculture, climate, and irrigation factors has been considered one of the most significant developments in achieving precision agriculture. It is possible to estimate the status of crops by monitoring environmental parameters (air quality, soil, weather, and water quality) for plants in the field and greenhouse. As a result, precision prediction of agricultural factors using powerful algorithms and intelligence heuristic models provides detailed and valuable information that supports farming practices (Mancipe-Castro & Gutiérrez-Carvajal, 2022;

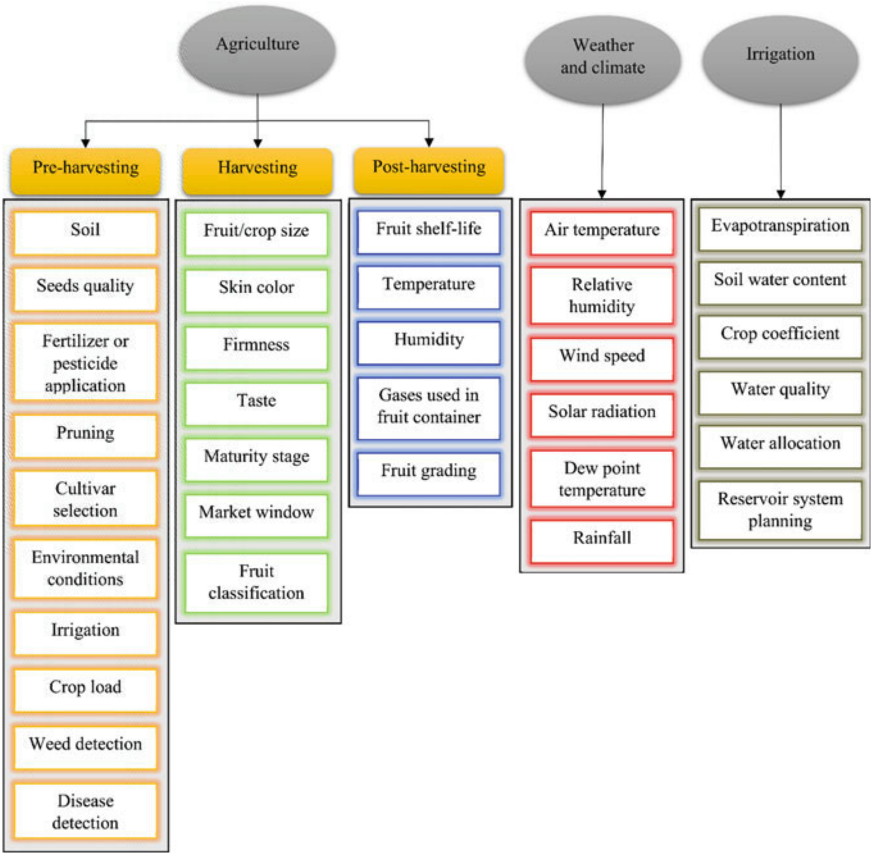

Fig. 1.1 Different parameters considered in studies related to agriculture, weather and climate, and irrigation problems

Nwokolo et al., 2022; Patil and Deka, 2016). One of the main challenges of prediction in agriculture studies is to develop robust techniques that can produce more accurate predictions using limited data. Data scarcity, data missing, ungagged data, or limitations are important issues in agriculture and climate predictions that need to be addressed in precise agriculture predictions by developing forecasting systems.

In the past few decades, innovative technologies have revolutionized the world and human life to solve new, complex, challenging problems. In terms of the models or methodologies used for meteorological and agriculture factors predicting, forecasting systems can be categorized into three types, including statistical models, artificial intelligence (AI) models, and hybrid models. Statistical and mathematical models such as auto-regressive moving average (ARMA), auto-regressive integrated moving average (ARIMA), and their variants are widely used over time (Table 1.1). Solving complex problems is difficult with these models due to nonlinear behavior and relationships between the components. Hence, some methods have been developed to exhibit "intelligence behavior". By using these methods, valuable information can

Table 1.1 Stochastic models application

Researchers	Study description	Applied model
Aghelpour and Norooz-Valashedi (2022)	Predicting daily reference evapotranspiration rates in a humid region	Auto-regressive (AR), moving average (MA), ARIMA, ARMA, least square support vector machine (LSSVM), adaptive neuro-fuzzy inference system (ANFIS), and generalized regression neural network (GRNN)
Elsaraiti and Merabet (2021)	Wind speed prediction	ARIMA, artificial neural networks (ANNs), recurrent neural networks (RNNs), and long short-term memory (LSTM),
Praveen and Sharma (2020)	Climate variability and its impacts on agriculture production	ARIMA
Wen et al. (2019)	Linear modeling of agricultural management time series	ARIMA and SVM
Dwivedi et al. (2019)	Forecasting monthly rainfall	ARIMA and ANN
Katimon et al. (2018)	Predicting river water quality and hydrological variables	ARIMA, AR, and MA
Mossad and Alazba (2015)	Drought forecasting in a hyper-arid climate	ARIMA
Kumari et al. (2014)	Rice yield prediction	ARIMA
Fernández-González et al. (2012)	Estimation of atmospheric vineyard pathogens	ARIMA

be extracted and used to recognize patterns and gain a deeper understanding of the problem (Yaghoubzadeh-Bavandpour et al., 2022).

AI techniques are powerful methods for solving complex problems in several fields of science. Engelbrecht (2007) defined AI as "*the study of adaptive mechanisms to facilitate intelligent behavior in complex and changing environments*". The artificial intelligence models can be further classified into machine learning (ML) and deep learning predictors. Machine learning is defined as "*it gives computers the ability to learn without being explicitly programmed*" (Samuel, 1967). ML is a learning process by itself from experience and given data to perform a prediction (new or future data). Technologies like ML can be used to get information and process it to increase crop production, improve crop quality, reduce crop disease, and increase the profitability of farmers. Precision learning is critical to improving harvesting yields in agriculture (Meshram et al., 2021). Deep learning (DL) is a subset of ML that has high performance in dealing with real-life, complex, and challenging problems (Fig. 1.2).

Some of ML models are artificial neural networks (ANNs), Bayesian models (BM), deep learning (DL), dimensionality reduction (DR), decision trees (DT), ensemble learning (EL), random forest (RF), and support vector machines (SVMs). The most important ML algorithms employed for weather forecasting are adaptive neuro-fuzzy inference systems (ANFIS), bootstrap aggregating (bagging), backpropagation network (BPN), generalized regression neural network (GRNN), K-nearest neighbor (KNN), multiple linear regression (MLP), etc. (Jaseena & Kovoor, 2020; Liakos et al., 2018). Nonlinear datasets can be handled better using ML and deep learning models. Rapid development in AI and ML techniques has improved the precision of future weather, environmental, and agriculture predictions (Gheisari et al., 2017). These techniques provide predictions based on real-world data (Meenal

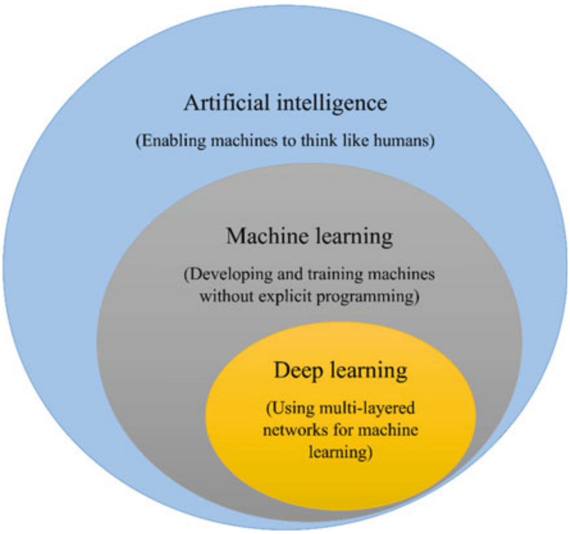

Fig. 1.2 Hierarchy diagram of artificial intelligence, machine learning, and deep learning (Arumugam et al., 2022)

et al., 2022). Table 1.2 shows the applications of ML and DL algorithms in agriculture and weather studies.

Table 1.2 Machine learning and deep learning models application

Researchers	Study description	Applied model
Band et al. (2022)	Estimation long-term mean monthly wind speed	ANN, RF, gene expression programming (GEP), and multivariate adaptive regression spline (MARS)
Chaudhary et al. (2022)	Soil moisture estimation	SVM, RF, MLP, ANFIS, radial basis function (RBF), Wang and Mendel's (WM), subtractive clustering (SBC), hybrid fuzzy interference system (HyFIS), and dynamic evolving neural fuzzy inference system (DENFIS)
Zongur et al. (2022)	Detection of the diameter of the inhibition zone of gilaburu (*Viburnum opulus* L.) extract against eight different Fusarium strains isolated from diseased potato tubers	SVM, KNN, RF, ensemble algorithms (EA), AdaBoost (AB) algorithm, and gradient boosting (GBM) algorithm
Buyrukoğlu (2021)	Prediction of Salmonella presence in agricultural surface waters	SVM, ANN, RF, and Naïve Bayes (NB)
Yoosefzadeh-Najafabadi et al. (2021)	Predicting soybean (*Glycine max*) seed yield	MLP, SVM, and RF
Nuruzzaman et al. (2021)	Identifying the potato species	RF, linear discriminant analysis (LDA), logistic regression, SVM, CART, NB, and KNN
Wu et al. (2019)	Estimation of monthly mean daily reference evapotranspiration	ANN, RF, gradient boosting decision tree (GBDT), extreme gradient boosting (XGBoost), multivariate adaptive regression spline (MARS), SVM, and kernel-based nonlinear extension of Arps decline (KNEA)
Fan et al. (2018)	Predicting daily global solar radiation	SVM and XGBoost
Mokhtarzad et al. (2017)	Drought forecasting	ANFIS, ANN, and SVM
Goel and Sehgal (2015)	Estimation of the ripeness of tomatoes based on color	Fuzzy rule-based classification approach (FRBCS)

Two or more models are combined in hybrid models to improve forecasting performance (Jaseena & Kovoor, 2020). Hybrid models provided more accurate predictions than conventional and other standalone artificial intelligence-based forecasting models (Seifi et al., 2021a, 2021b). Prediction can be more accurate and effective with the advancement of hybrid models using ML, optimization techniques, and suitable data visualization methods (Gheisari et al., 2017). Recently, meta-heuristic/swarm intelligence optimization algorithms have been developed to find the optimal solutions of a problem. These algorithms provide an opportunity to overcome problems related to the stability of individual models. Table 1.3 shows the applications of hybrid algorithms in agriculture and weather studies.

There is a need to increase cereal production by 3 billion tons and meat production by over 200% by 2050 to meet the needs of the global population (Trendov et al., 2019). Weather uncertainty, climate change crises, rising food prices, and the rapid growth of the world's population have all led to agricultural sectors looking for new ways to maximize harvests. Hence, the use of ML techniques is increasing

Table 1.3 Hybrid models application

Researchers	Study description	Applied model
Bazrafshan et al. (2022)	Tomato yield prediction	ANFIS and MLP hybridized with multiverse optimization algorithm (MOA), particle swarm optimization (PSO), and firefly algorithm (FFA)
Seifi et al. (2022)	Prediction of pan evaporation	ANFIS hybridized with seagull optimization algorithm (SOA), crow search algorithm (CA), firefly algorithm (FFA), and PSO
Sahoo et al. (2021)	Prediction of flood	Radial basis function–firefly algorithm (RBF-FA) and support vector machine–firefly algorithm (SVM-FA)
Seifi et al. (2021a)	Global horizontal irradiation predictions	ANFIS and ELM hybridized with multiverse optimization (MVO), sine cosine algorithm (SCA), and salp swarm algorithm (SSA)
Seifi et al. (2021b)	Soil temperature prediction	ANFIS, SVM, RBFNN, and MLP hybridized with sunflower optimization (SFO), FFA, salp swarm algorithm (SSA), and PSO
Ehteram et al. (2021c)	Infiltration rate prediction	ANFIS hybridized with SCA, PSO, and FFA
Ashofteh et al. (2015)	Irrigation allocation policy under climate change	Genetic programming
Noory et al. (2012)	Irrigation water allocation and multicrop planning optimization	Discrete PSO algorithm

in agriculture as technology advances (Shaikh et al., 2022). In preharvesting, the ML models predict crop, agriculture, and environmental conditions, including seed quality, genetics, soil, pruning, fertilizer application, temperature, and humidity. By considering each component, production losses can be minimized. Also, ML techniques are helping farmers reduce harvesting and post-harvesting losses (Meshram et al., 2021). A key principle of ML in agriculture is its adaptability, speed, precision, and cost-effectiveness. Farmers can increase yields and improve quality with fewer resources by using AI and ML techniques in agriculture (Shaikh et al., 2022). This chapter focuses on ML applications in agriculture, weather and climate, and irrigation problems.

1.2 The Necessity of Meteorological Variables Prediction

Weather and climate are two parts of meteorology that can largely influence human life. Identifying and predicting weather and climate are important for some studies, such as agriculture management (Parasyris et al., 2022). Also, weather prediction and climate considerations are highly correlated with decision-making for agriculture, water resource management, irrigation management, drought and flood conditions management, etc.

The spatial and temporal distribution of most meteorological variables is an important factor that is commonly used in a variety of scientific fields, including climatology (Philandras et al., 2015; Varquez & Kanda, 2018), environmental and ecological studies (Paul et al., 2019; Zanobetti & Schwartz, 2008), hydrological studies (Wang et al., 2009), agriculture (Wakamatsu, 2010), epidemiology studies (Bunker et al., 2016), and many other studies. Meteorological variables information is often limited in time and space, and there is a critical need to predict these variables (Kloog et al., 2014). Also, as global climate change continues, developing reliable and accurate models is imperative, especially using limited data. Hanoon et al. (2021) presented that using the univariate ARMA models, the low temperature values are overestimated, while high values are underestimated, resulting in poor water resources planning and management. Therefore, developing reliable prediction models with low uncertainty that avoid these shortages is necessary for predicting meteorological variables (Ehteram et al., 2021a; El-Shafie et al., 2014).

A great deal of spatial variability, time variation, and stochastic behavior is associated with meteorological variables, making it difficult to predict with simple models (Adnan et al., 2021). The classification of different models for predicting meteorological variables is given in Fig. 1.3. Despite the importance of conceptual models for identifying meteorological processes, they have encountered many practical challenges, particularly when accuracy is the most important factor (El-Shafie et al., 2011). Rather than building conceptual models, exploring and developing data-driven models could be more beneficial. In different studies, models based on data have been found to provide accurate predictions (Ehteram et al., 2021b; Seifi et al., 2022). However, due to the nonlinear dynamics inherent in meteorological phenomena, most models may not always perform well and may lead to undesirable results in many cases (Hanoon et al., 2021). Since ML approaches include both effective structure and efficient learning, these approaches have been proposed as an alternative modeling technique for a dynamic system with nonlinear behavior (Ahmed et al., 2019).

ML models require less information to make accurate predictions of future time series. The internal network parameters are determined based on the available time series and applying a suitable tuning algorithm. Furthermore, this could include adjusting the initially selected network structure for assessing the better model structure while studying a special problem (Palit & Popovic, 2006). Recently, ML techniques have remarkably advanced in modeling dynamic and nonlinear systems for different science applications (Ehteram et al., 2021c; Panahi et al., 2021; Seifi et al., 2021b, 2022) (Fig. 1.3). Thus, it could be used as an efficient approach for modeling meteorological processes. The main advantage of this strategy is its ability to succeed when explicit knowledge of the internal meteorological process is unavailable. Although these ML models have proven effective, it is still not known which of

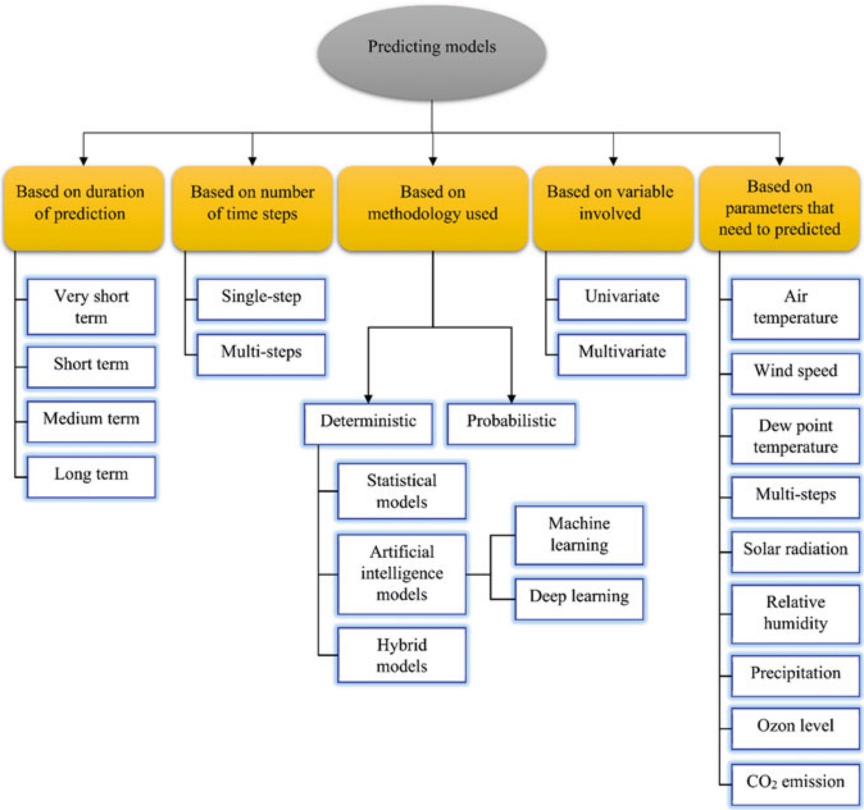

Fig. 1.3 Classification of predicting models related to weather and climate studies

the ML models would be the most appropriate option for certain system processes. Therefore, evaluating and comparing various ML modeling approaches are necessary to select the most appropriate one (Hanoon et al., 2021).

Overall, many studies approved the efficiency of ML techniques for predicting meteorological variables (Table 1.4). For example, Mohammadi et al. (2015) used extreme learning machine (ELM)-based, SVM, and ANN models to predict daily dew point temperature. The findings demonstrated that the applied ML models have a high potential for predicting daily dew point temperature. The performance of the ELM model was higher than those of SVM and ANN techniques. Azad et al. (2019) investigated the ability of standalone ANFIS to predict monthly air temperature. In another attempt, the ANFIS model was optimized using genetic algorithm (GA), PSO, ant colony optimization for continuous domains (ACOR), and differential evolution (DE). The results indicated that the hybrid ANFIS-GA model has the best performance in predicting maximum air temperature. Pham et al. (2020) presented that hybrid ANFIS-PSO, SVM, and ANN models are useful for predicting daily rainfall using other meteorological variables (maximum temperature, minimum temperature, wind speed, relative humidity, and solar radiation). Among applied ML models, the SVM was more efficient in predicting rainfall. Qadeer et al. (2021) predicted relative humidity using RF and SVM ML models. Also, a commercial process simulator called Aspen HYSYS® V10 was used to create a data mining environment. The prediction accuracy of RF model was 74.4% higher than those of the SVM. Compared to Aspen HYSYS, the ML models of RF and SVM predicted relative humidity with a mean absolute deviation of 1.1% and 4.3%, respectively. Jin et al. (2022) applied an ensemble model using integrating two regression models of the generalized additive model (GAM), and generalized additive mixed model (GAMM), and two ML models of RF, and extreme gradient boosting (XGBoost) to estimate daily mean air temperature from satellite-based land surface temperature. The results showed that the ensemble model has the highest performance ($R^2 = 0.98$, RMSE = 1.38 °C), and two ML models [RF ($R^2 = 0.97$), XGBoost ($R^2 = 0.98$)] showed better accuracy than the two regression models [GAM ($R^2 = 0.95$), GAMM ($R^2 = 0.96$)]. Malik et al. (2022a, 2022b) reviewed the ANN-based model in predicting solar radiation and wind speed. They demonstrated that using the artificial neural system is the best way to anticipate extremely nonlinear meteorological data.

1.3 The Necessity of Agricultural Factors Prediction

Agriculture plays a vital role in the global economy. The increasing human population will increase pressure on the agricultural system. Precision farming, also known as digital agriculture, is a new field of science that uses data-intensive methods to maximize agricultural productivity and minimize its environmental impact. A possible way to deal with ecological, social, and economic problems in agriculture is through intelligent systems, which offer the opportunity to develop intelligent systems. The

Table 1.4 Application of machine learning in the prediction of meteorological variables and weather

Researcher	Model(s)	Predictive variable								
		Air temperature	Relative humidity	Solar radiation	Wind speed	Dew point temperature	Rainfall	Air pressure	CO_2 emission	Ozone level
Başakın et al. (2022)	ANN, GPR, SVM, MARS				×					
Namboori (2020)	SVM								×	
Tran Anh et al. (2019)	ANN						×			
Khosravi et al. (2018)	ANN, SVR, FIS, ANFIS			×						
Liu et al. (2018)	ANN	×								
Rasel et al. (2017)	ANN, SVR	×					×			
Piri et al. (2016)	ANN, SVM, ANFIS			×						
Khajure and Mohod (2016)	ANN	×	×		×	×		×		
Shabariram et al. (2016)	SVM						×			
Ahmed (2015)	RF	×			×		×			
Mohammadi et al., (2015)	ANFIS					×				
Luna et al. (2014)	ANN, AVM									×

(continued)