ORIGINAL ARTICLE

A Comprehensive Machine and Deep Learning Approach for Aerosol Optical Depth Forecasting: New Evidence from the Arabian Peninsula

Ahmad Qadeib Alban[1] · Ammar Abulibdeh[2] · Lanouar Charfeddine[3] · Rawan Abulibdeh[4] · Abdelgadir Abuelgasim[5]

Abstract

Accurate forecasting of environmental pollution indicators holds significant importance in diverse fields, including climate modeling, environmental monitoring, and public health. In this study, we investigate a wide range of machine learning and deep learning models to enhance Aerosol Optical Depth (AOD) predictions for the Arabian Peninsula (AP) region, one of the world's main dust source regions. Additionally, we explore the impact of feature extraction and their different types on the forecasting performance of each of the proposed models. Preprocessing of the data involves inputting missing values, data deseasonalization, and data normalization. Subsequently, hyperparameter optimization is performed on each model using grid search. The empirical results of the basic, hybrid and combined models revealed that the convolutional long short-term memory and Bayesian ridge models significantly outperformed the other basic models. Moreover, for the combined models, specifically the weighted averaging scheme, exhibit remarkable predictive accuracy, outperforming individual models and demonstrating superior performance in longer-term forecasts. Our findings emphasize the efficacy of combining distinct models and highlight the potential of the convolutional long short-term memory and Bayesian ridge models for univariate time series forecasting, particularly in the context of AOD predictions. These accurate daily forecasts bear practical implications for policymakers in various areas such as tourism, transportation, and public health, enabling better planning and resource allocation.

✉ Ammar Abulibdeh
aabulibdeh@qu.edu.qa

Lanouar Charfeddine
lcharfeddine@qu.edu.qa

1 Research Assistant in Machine Learning, College of Engineering, Qatar University, Qatar, Doha, Qatar

2 Applied Geography and GIS Program, Department of Humanities, College of Arts and Sciences, Qatar University, Doha, Qatar

3 Center for Entrepreneurship and Organizational Excellence, College of Business and Economics, Qatar University, P.O. Box 2713, Doha, Qatar

4 Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada

5 Rabdan Academy, 65 Alinshirah Street, Alsa'adah, 22401 Abu Dhabi, United Arab Emirates

1 Introduction

Atmospheric aerosols are a blend of fine solid, liquid, gaseous or mixed particles ranging in size (particle sizes of $10-3$ to $100\ \mu m$) suspended in the air due to natural processes and anthropogenic activities such as dust, smoke, and pollution (Mushtaq et al. 2022; Almazroui 2019; Huang et al. 2020; Putaud et al. 2010; Li, et al. 2022). These particles come from both natural sources, like volcanic eruptions, dust storms, and sea spray, as well as human activities, such as industrial processes, combustion of fossil fuels, and agriculture (Mushtaq et al. 2022; Arfin et al. 2023). Large concentrations of these particles have a significant impact on the management of the Earth's atmosphere, ecosystems, climate system, regional and global climate change, ambient air quality, agricultural production, and human health (Arden Pope, et al. 2011; Zhang et al. 2016; Song et al. 2018). Aerosols play a key role in climate by scattering and absorbing

solar radiation, which is the energy balance between the Earth and its atmosphere and can lead to changes in the hydrological cycle patterns, affect the formation and optical properties of clouds, reduce atmospheric visibility, reduce the amount of sunlight that reaches the ground, affect the radiation equilibrium of Earth, and alter atmospheric chemistry (Chen et al. 2013; Abuelgasim and Farahat 2019; Bilal et al. 2013; Levy et al. 2007). The interaction between the presence of atmospheric aerosols in the atmosphere and the incoming solar radiation significantly affect the earth radiation budget (Ali and Assiri 2019; Ali et al. 2017, 2018).

Aerosol Optical Depth (AOD) is a measure of how much aerosols in the atmosphere prevent the transmission of light by absorption or scattering of light (Zhang et al. 2020; Wei et al. 2020). A higher AOD value indicates a higher concentration of aerosols, which can lead to more significant scattering and absorption of sunlight (Ranjan et al. 2021). This measurement is critical in understanding the impact of aerosols on Earth's radiation balance and climate. AOD data, often estimated from satellite observations and ground-based measurements, help in assessing the clarity of the atmosphere and the global distribution of aerosols (Almazroui 2019; Li et al. 2020). This information is vital for climate modeling and predicting changes in Earth's energy balance due to aerosols.

Studying aerosols and their optical depth is essential for several reasons. Firstly, aerosols significantly affect Earth's climate system; they can affect surface temperatures, influence weather patterns, and alter the energy balance of the atmosphere (Li, et al. 2022; Hansen et al. 2011; Zhou et al. 2021). Secondly, aerosols have direct implications for human health (Oh et al. 2020; Sahu et al. 2020; Chowdhury et al. 2022). Fine particulate matter, a type of aerosol, is linked to respiratory and cardiovascular diseases, posing a significant public health risk. Understanding the sources, composition, and distribution of aerosols is crucial for effective environmental and public health policies. Moreover, accurate knowledge of aerosol optical properties is essential for developing strategies to mitigate climate change and protect human health. Therefore, ongoing research and monitoring of aerosol characteristics and dynamics are critical in addressing environmental and health challenges at both local and global scales.

In the environmental literature, it is well acknowledged that a full understanding of the data generating process of the environmental degradation proxies is essential to help policy makers better anticipating any possible deviation from the international standards required to meet the COP21 requirements. However, while most of the previous studies have focused on the analysis of a single or local environmental degradation proxies such as dioxide carbon emissions, CH4 and SO2 due to their data availability (Abulibdeh 2022), little attention have been given to the case of a more global,

sophisticated and comprehensive measures. This study tried to fill this gap in the empirical literature by analyzing daily data of the AOD of the Arabian Peninsula (AP) region over the period 2003–2019. The choice of the AOD as a measure of environmental degradation is mainly motivated by its important features in terms of providing a more global understanding of the level of environmental degradation.

In this study, we analyzed and modeled the aerosol in the AP region. The choice of the AP region as the region for analysis was mainly motivated by its large contribution to the global sand dust aerosols (Abuelgasim and Farahat 2019; Kumar et al. 2018). The region is known to experience high levels of aerosol pollution due to a combination of natural and human-made sources, such as dust storms, industrial activities, and transportation, which can have a significant impact on regional and global climate patterns (Ali et al. 2017). Dust storms are prevalent throughout the AP in which the Arabian Desert is a significant contributor to natural dust, with over half of the world's annual average dust emissions originating from this region (Ali and Assiri 2019; Nichol and Bilal 2016; Butt et al. 2017). These storms can be generated locally, transported to large distances, or a combination of both (Engelstaedter et al. 2006). Typically, instability in the atmosphere on a large scale and strong winds at ground level over the AP lead to the onset of dust storms in different regions. During the spring and summer months, these storms tend to happen more often in the eastern and southern areas of the AP (Klingmüller et al. 2016).

Studies have shown that the AP is particularly vulnerable to the effects of aerosols, which can have significant impacts on regional weather patterns, climate change, and public health (Abuelgasim and Farahat 2019; Meo et al. 2013; Farahat et al. 2015). Additionally, due to the importance of the AP as a hub of oil production and industrial development, aerosol levels above the region may be higher than in other parts of the world (Abuelgasim and Farahat 2019; Levy et al. 2007; Kumar et al. 2018). Studies conducted earlier have indicated a strong correlation between the fluctuation of aerosol levels in the AP and the frequency and strength of dust storms. Specifically, when there is a rise in either the number or intensity of dust storms, it is anticipated that there will be a corresponding increase in atmospheric aerosol concentration (Farahat et al. 2015; Esmaeil et al. 2014). Some of these studies focused on comparing satellite-based measurements with ground-based observations from the Aerosol Robotic Network (AERONET) over the AP. For example, Almazroui (Almazroui 2019) assessed the accuracy of the Moderate-resolution imaging spectroradiometer (MODIS) Deep Blue (DB) algorithms, specifically the Collection-51 (C-51) and Collection-06 (C-06), in measuring AOD at 550 nm over the period 2002–2013 and compared these satellite-based measurements with ground-based observations from the Aerosol Robotic Network (AERONET). The results indicate

that while MODIS algorithms generally capture the AOD patterns observed by AERONET, there are notable discrepancies. Both algorithms show significant uncertainties and errors, highlighting the need for further research to improve AOD estimation over Saudi Arabia. Ali and Assiri (2019) analyzed the spatiotemporal variations of aerosols over the Arabian Peninsula from 2003 to 2017 utilizing collection-06 MODIS-Merged Dark Target (DT) and Deep Blue (DB) aerosol products (550 nm) from the Terra and Aqua satellites. The study also employs AERONET-derived AOD data to evaluate the accuracy of these satellite-derived AOD measurements. The results show that the Terra satellite recorded the highest AOD over the southern Red Sea and the west coast of the Arabian Gulf, extending to central Saudi Arabia. In contrast, the Aqua satellite displayed the highest AOD mostly over the southern Red Sea. The study concludes that further qualitative research is needed to enhance the aerosol retrieval efficiency of the DB and DT algorithms over bright-reflecting surfaces. Ali et al. (2017) analyzed the spatiotemporal variations of AOD using the MODIS- DB algorithm from the Aqua satellite and AERONET to understand the seasonal distribution and variability of AOD over the region and to assess the accuracy of the MODIS- DB algorithm in capturing these patterns. They found significant seasonal variations in AOD across different regions of Saudi Arabia. The correlations between MODIS DB AOD and AERONET AOD varied by season, with higher correlations in spring and winter. The study concluded that inaccuracies in aerosol model selection and surface reflectance calculation were the primary reasons for reduced correlation between satellite and ground-based AOD measurements.

Several quantitative modelling approaches have been used to model and forecast environmental degradation proxies (Abulibdeh 2022; Jerrett, et al. 2005). The simplest and most meaningful way to categorize the different approaches used in previous studies is to classify them under econometric time series models versus machine and deep learning models (Charfeddine et al. 2023). Under the first class of models, the econometric time series model, several models have been used to model and forecast the AOD time series, such as autoregressive integrated moving average (ARIMA), seasonal ARIMA (SARIMA), and exponential smoothing models (Meo et al. 2013). Most of these studies have demonstrated the usefulness of time series models in analyzing and predicting AOD (Farahat et al. 2015). Taneja et al. (2016) used ARIMA to model the monthly average AOD levels over New Delhi. They found a seasonality patterns in the AOD time series. Li et al. (2020) used ARIMA to predict and reproduce AOD variability over China and the United States between 2003 and 2015. They found that the concentration values of AOD are high in the eastern parts and during the summer in these two countries. Abuelgasim et al. (Abuelgasim et al. 2021) examined the spatiotemporal

trends of AOD over the United Arab Emirates using long term time series analysis for the period 2003–2018. Data were obtained from MODIS and Multi-Angle Implementation of Atmospheric Correction (MAIAC). The results show the presence of significant annual seasonal variation in AOD between the summer, spring, and winter seasons. The increase in AOD results in increasing the air temperature and humidity causing scattered events of haboobs and intense dust storms. Spatially, the values of the AOD were found higher over the desert and coastal areas. In terms of the time series modelling, they found that SARIMA forecast model provides a reliable and accurate monthly forecast for the AOD concentration.

The second class of models that have been widely used in recent literature is machine learning and artificial intelligence techniques. These models have been used to explore and model the different patterns and dependencies within historical data to predict future environmental behavior. Examples of such techniques applied to analyze environmental time-series data encompass random forest regression, support vector regression, and various neural network architectures (Zaheer et al. 2023; Nath et al. 2022; Zbizika et al. 2022; Jing et al. 2017). For instance, Yang et al. (Zhen and Shi 2023) assessed the accuracy of monitoring atmospheric composition and climate aerosol optical depth values at a wavelength of 550 nm by employing ground-based observational data from AERONET stations in China for a period spanning from 2003 to 2007. They developed a data fusion correction model using a random forest regression technique, yielding a significant reduction in mean absolute error from 0.225 to 0.047 and a decrease in root mean square error from 0.400 to 0.120. However, the utilization of random forest regression may be limited in capturing the intricate relationships between independent variables and model errors. Similarly, Komal et al. (Zaheer et al.2023) explored the feasibility of employing diverse machine learning structures for aerosol optical depth prediction, including support vector regression and multi-linear regression. They proposed a hybrid model by integrating support vector regression with the gray wolf optimizer meta-heuristic algorithm. Notably, the gray wolf optimizer algorithm optimized support vector regression hyperparameters, resulting in an overall improvement in the model performance.

In the realm of machine learning methods, neural network models have gained prominence due to their flexibility and robust capacity to model complex patterns concealed within data (Jing et al. 2017). Various neural network architectures have been applied to AOD predictions, encompassing traditional backpropagation neural networks (Jing et al. 2017), deep neural networks for aerosol optical depth retrieval (Jing et al. 2017), generative adversarial networks to enhance AOD estimation (Hoyne et al. 2019), convolutional neural networks coupled with the MERRA-2 reanalysis dataset for

AOD prediction (Qin et al. 2018), and long short-term memory networks (Hochreiter and Schmidhuber 1997), which constitute a specialized class of recurrent neural networks. While recurrent neural networks are adept at predicting time series data, they struggle to model long-term dependencies in AOD time series due to the vanishing gradient problem. To address the high dependency issues between observations, various hybrid models based on the long short-term memory (LSTM) algorithm have been utilized to assess predictive capabilities (Daoud et al. 2021; Tian and Chen 2022; Han et al. 2021). For instance, Daoud et al. (2021) found that Conv-LSTM outperforms the basic LSTM and the hybrid convolutional neural network (CNN)-LSTM model in predicting AOD. Han et al. (2021) also found that the CNN- LSTM model provides better forecasting capability than ARIMA, LSTM, Conv-LSTM and deep residual network-LSTM (ResNet-LSTM) models. In a recent paper, Tian and Chen (2022) propose an AT-CONVLSTM, a new attention-based spatial–temporal information extraction network, and found that it outperforms the multilayer perceptron, LSTM, CNN-LSTN, and Conv-LSTM (see Table 1 for more detailed summary of previous studies).

This paper follows this second methodology approach by utilizing different machine and deep learning models including basic, hybrid and combined models to predict the AOD time series. The use of these kinds of models is mainly motivated by the large sample size of our data which makes these types of models outperform traditional econometric models. Moreover, the different features that characterize the aerosol data (seasonality, outliers, nonlinearity) make from these types of models the appropriate candidates to model and forecast these types of series (see (Charfeddine et al. 2023)).

The primary objectives of this study is the development of advanced AOD forecast models using machine learning techniques and advanced computational algorithms. This is achieved through threefold approaches. First, in contrast to some previous studies, this paper used an advanced statistical technique to de-seasonalize the AOD time series, e.g. the daily seasonal adjustment technique (DSA). This approach had the advantage of adjusting for several patterns such as intra-weekly, intra-monthly, intra-annual. Moreover, the daily seasonal adjustment approach had the advantage of adjusting for the calendar effects, cross-seasonal effects, and outliers. Second, we make use of a large set of machine learning and deep learning model architectures for AOD forecasting where we performed hyper-parameter tuning to find the best model configurations for predicting future AOD values. This is particularly important since the proposed specifications allowed us to construct models that can effectively capture the patterns and characteristics of the AOD data for the most optimal time-series forecasting. Third, we combined the two best-performing models using a rolling prediction scheme for

more accurate performance. The main contribution of this paper is to investigate the role of feature extraction for the prediction of AOD values in the AP region using various machine-learning models, with the addition of two combined models to enhance prediction results.

The rest of the paper is organized as follows. Section 2 presents the study area and how the data was extracted. Section 3 presents the materials and methods used to deal with missing data, deseasonalization, normalization and the different machine and deep learning models used in the empirical literature to forecast the aerosol time series. Section 4 presents and discusses the empirical findings. Finally, Sect. 5 concludes and proposes some policies recommendations.

2 Study Area and Data Extraction

2.1 Study Area

The AP is situated in the southwestern region of Asia, covering an area of approximately a million square kilometers and is home to around 77 million people (DiBattista et al. 2020; Patlakas et al. 2019; Abulibdeh and Zaidan 2020). The region encompasses Saudi Arabia, Oman, Qatar, the United Arab Emirates (UAE), Bahrain, and Yemen, situated between 12 and 32°N latitudes and 30–60°E longitudes (Abulibdeh et al. 2019a). The AP is a vast region with diverse geographical features and climate zones, including deserts, mountains, coastal plains, and wetlands (Abuelgasim et al. 2021; Watson-Parris et al. 2019; Abulibdeh et al. 2021). The environmental conditions in this region are challenging and have played a significant role in shaping the cultures, lifestyles, and economies of the people who live there. It is surrounded by three major water bodies, namely the Red Sea to the west, the Arabian Gulf to the east, and the Arabian Sea to the south (Fig. 1). One of the most striking environmental features of the AP is its vast deserts. The region is predominantly desert, with the exception of the southwestern region, which is characterized by mountainous terrain and receives more rainfall than other parts of the peninsula (Abulibdeh et al. 2019b). The AP is mostly arid and is characterized by vast deserts, including the Rub' al Khali or the "Empty Quarter," which is one of the largest sand deserts in the world. This desert is characterized by its towering sand dunes, high temperatures, and harsh winds, making it one of the harshest environments on earth and promoting significant sand and dust movements. The mountainous regions of the AP are also notable for their unique environmental conditions. The region is known for its extreme temperatures, with some locations experiencing temperatures as high as 54 °C during the summer months (Abulibdeh 2021). In contrast, the spring and autumn seasons are mild, while winters are relatively cold. The

Table 1 Summary of previous studies investigating various approaches for AOD forecasting

Study	Aim of the study	Models Used	Results
Jing et al. (2017)	Comparing and contrasting the performance of two methods, namely SVM and back propagation (BP) neural network, in forecasting the AOD values using PM concentration as an estimation model	- Support vector machine SVM - Backpropagation (BP) neural network	- Both SVM and BP models exhibited robust nonlinear fitting capabilities and simplicity in implementation - BP model took less time computation than the SVM model
Zhang et al. (2021)	Enhancing the estimation of (AOD) by applying a conditional adversarial network architecture (AeroCGAN) and merging multi-sensor AOD products	- Conditional generative adversarial network architecture (AeroCGAN), where the generator in AeroCGAN used an encoder-decoder network for initial resolution improvements - Three-dimensional convolutional neural network (3D-CNN) was adopted for environmental feature extraction and residual network for improving accuracy	- The AeroCGAN model, trained using MODIS Multiple Angle Implication of Atmospheric Correction (MAIAC) AOD data, provided 1-km-resolutin, hourly Himawari AOD products - The generator in AeroCGAN displayed satisfactory results in initial resolution improvements
Qin et al. (2018)	Enhancing the estimation of daily aerosol optical depth (AOD) and aerosol radiative effect (ARF) using an optimized artificial neural networks (ANN) model	An optimized back propagation (BP) artificial neural network (Genetic_BP) was constructed for enhancing AOD values, using meteorological measurements	The optimized (BP) artificial neural network (Genetic_BP) displayed high accuracy in estimating (AOD) values, with correlation coefficient R of 0.866 for CASNET stations and 0.865 for AERONET
Yeom et al. (2021)	Comparing the performance of the deep neural networks (DNN) with traditional machine learning models (ML), including RF and SVR, and selected physical models	- Deep Neural Network (DNN) model was applied to estimate the spatiotemporal dependencies of the hourly (AOD) using (GOCI) over Northeast Asia - Random Forest RF and support Vector Regression (SVR) models were compared with the DNN model to evaluate their accuracy in estimating AOD values - Physical models were also compared with the DNN model	DNN model demonstrated higher accuracy in estimating the hourly AOD values compared to traditional ML models, including RF and SVR, and the selected physical models. The DNN model had RMSE of 0.112, MBE of 0.007, and correlation coefficient R of 0.863 in the hold-out validation analysis
Nabavi et al. (2018)	- Conduct Forecasting for AOD values using deterministic models and machine learning algorithms - Identifying the factors causing prediction errors of both DMs and MLs, such as the resolution of variable datasets, omission of influential predictors, and scarcity of extreme cases	- Deterministic weather prediction models (DMs) - Machine Learning Algorithms (MLAs) - Linear Models and Neural Networks	MLAs, especially Multivariate Adaptive Regression Splines (MARS) and Support Vector Machines (SVM), outperformed deterministic weather prediction models (DMs) in forecasting AOD in West Asia on a monthly time scale
Foued et al. (2020)	- Investigate the dependency of AOD on wavelength and develop a mathematical model of global solar radiation based on AOD data between two wavelengths: 550 and 1250 nm - Establishing a relationship between routine lighting measurements and AOD, to use it as an alternative tool to estimate AOD at radiometric stations	- Mathematical model of global solar radiation based on AOD - Methods and Algorithms implemented to examine the performance of cloud and radiation forecasts in Helsinki	A comparative study between theoretical results and experimental data displayed a satisfactory correlation with a relative error of 2–15%
Nemanja et al. (2016)	Building a semi-supervised method for Compiling Forecasts of (AOD) from multiple satellite instruments into a more precise forecast, using a small number of accurate ground-based instruments as labels	- Semi-supervised learning method for compiling forecasts - Expectation–Maximization (EM) algorithm to optimize the parameters that maximize the likelihood function	The proposed technique effectively harnessed labelled and unlabeled data to provide accurate compiled AOD estimations

Table 1 (continued)

Study	Aim of the study	Models Used	Results
Daoud et al. (2021)	Assess prediction of temporal monthly of AOD over four dust sources	Long-short term memory (LSTM), Convolutional neural networks-long-short term memory (CNN-LSTM) and Convolutional long-short term memory (ConvLSTM)	ConvLSTM algorthims shows the least RMSE

peninsula has receives low amounts of annual rainfall ranging from 77 to 102 mm.

The AP is experiencing rapid urbanization and industrialization, leading to an increase in air pollution levels (Abulibdeh 2021; Seddon et al. 1794). Understanding the sources and characteristics of aerosols in the region can help in developing effective strategies to improve air quality. The AP is located in a region of the world that is highly susceptible to climate change. This region is known for its arid and semi-arid climate, and any changes in precipitation patterns can have significant implications for water resources and agriculture (Abulibdeh et al. 2019b). Understanding the sources and characteristics of aerosols in the region is therefore important for predicting future climate changes and developing strategies to mitigate their impacts. Finally, the AP is an important global oil-producing region, and the oil and gas industry is a significant source of aerosol pollution. Studying aerosols from this region can therefore help to develop more sustainable practices and reduce the environmental impact of these industries.

2.2 MODIS and Image Processing

This study utilized the daily MODIS MAIAC AOD product acquired from the Level-1 and Atmosphere Archive & Distribution System (LAADS) Distributed Active Archive Center (DAAC) for the period of 2003–2019 for the AP. The MODIS sensor is an effective tool for capturing the spatiotemporal variability of aerosols globally (Remer et al. 2008; Levy et al. 2015). The MODIS instrument, which is based on satellite technology, generates global AOD data at a 10 km resolution over land using two algorithms—Dark Target (DT) (Levy et al. 2013) and Deep Blue (DB) (Hsu et al. 2013). The DT algorithm has been tested only on vegetated and moist-soil surfaces, and has limitations in detecting aerosol levels on bright-reflecting surfaces (Levy et al. 2013). The accuracy of the DT algorithm is affected by errors in selecting aerosol model schemes and surface reflectance calculations (Nichol and Bilal 2016; Bilal, et al. 2019). To overcome these limitations, the DB algorithm has been developed, which can estimate AOD levels over desert areas (Almazroui 2019). The blue channel in the DB algorithm is used to calculate aerosol density on bright-reflecting surfaces where low reflectance is needed. Passive information regarding aerosols that was obtained from MODIS sensors has been utilized in various studies on global aerosol distribution (Watson-Parris et al. 2019; Bellouin et al. 2020; Wei et al. 2019; Remer, et al. 2020), radiative forcing (Yuan et al. 2019; Subba et al. 2020) and how aerosols impact regional climate (Zhao et al. 2021; Qin, et al. 2021). It has been noted that the dark blue AOD products from MODIS Aqua typically offer greater precision when it comes to the AP (Butt et al. 2017).
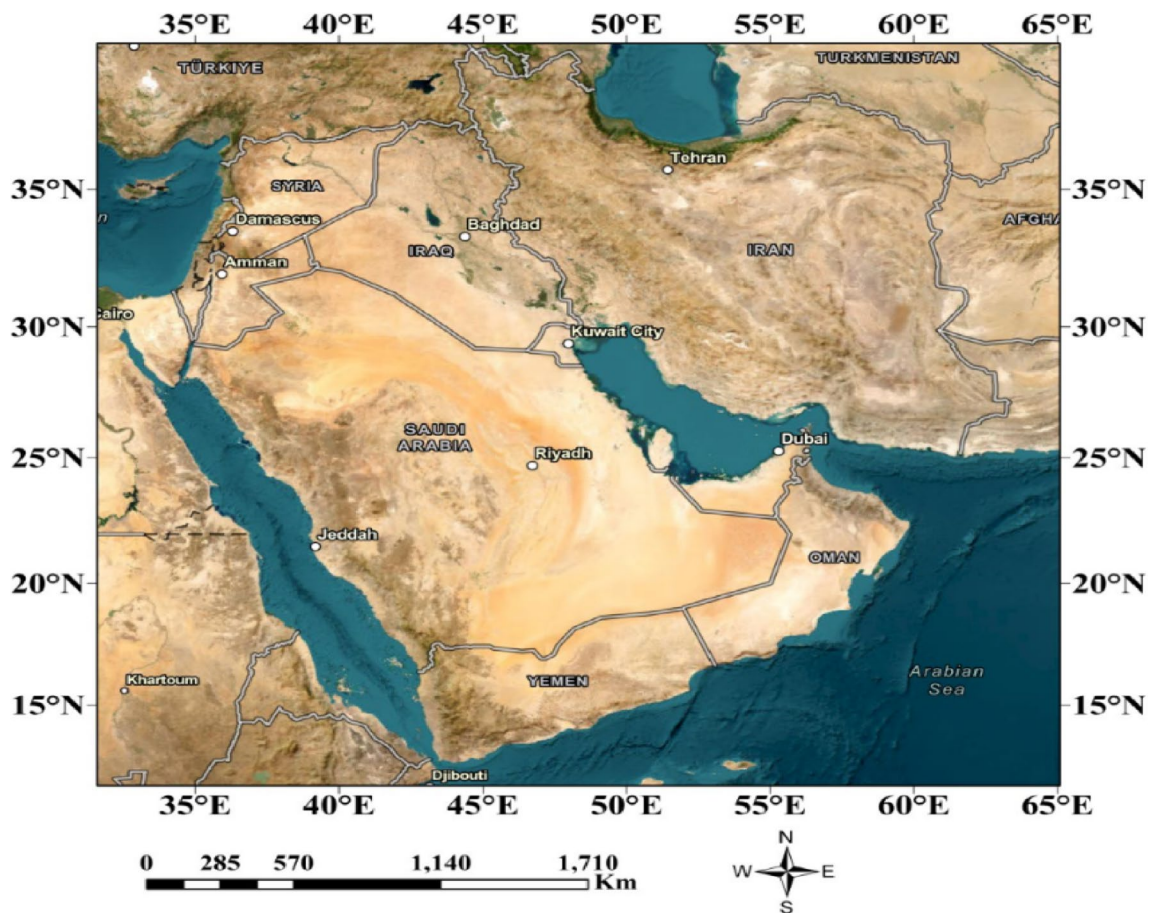
Fig. 1 The study area

Other techniques have been developed for estimating AOD from optical remote sensing imagery using radiative transfer theory (Levy et al. 2007). The Dense Dark Vegetation (DDV) method is the most popular among them, and it has been successfully used to measure AOD in areas with dense vegetation canopies (Levy et al. 2007). This method effectively removes surface reflectance by utilizing the correlation between shortwave infrared (2100 nm) and the red (660 nm) and blue (470 nm) bands, where the effect of aerosols is negligible (Levy et al. 2013). However, there are still uncertainties associated with sensor calibration drift when deriving AOD from space-borne observations. The DDV method has some limitations when used in bright surfaces, such as urban areas or arid/semi-arid regions. To overcome these limitations, a new method called Multi-angle Implementation of Atmospheric Correction (MAIAC) algorithm has been developed (Mhawish et al. 2018). This algorithm is designed to retrieve both the surface reflectance and atmospheric products simultaneously at a spatial resolution of 1 km. It achieves this by using time-series observations from MODIS and climatology data from Aerosol Robotic Network (AERONET) to estimate aerosol properties over both dark vegetated and

bright surfaces. The MAIAC algorithm utilizes regression coefficients to derive the surface parameterization based on the surface bidirectional reflectance distribution function in the MODIS bands of Blue (470 nm), green (550 nm), and shortwave infrared (2130 nm). It retrieves both aerosol optical thickness and surface bi-directional reflection factor using seven regional aerosol models for different regions across the globe. For dust-related aerosols, the algorithm uses either the dust or background models, similar to the DT algorithm (Levy et al. 2013). The choice to use the MAIAC algorithm for this study was driven by the reported studies of its effectiveness in estimating AOD over bright targets (Eibedingil et al. 2021; Lyapustin, et al. 2011; Chen et al. 2021).

To process MODIS data, this study combined different titles that covered the AP by mosaicking them and then removed areas outside the boundaries. This study also used statistical parameters like standard deviation, maximum, and minimum AOD to calculate the average daily AOD concentrations during the study period. A short model code was developed to expedite the mosaicking and statistical calculations using ArcGIS 10.6.1. The study computed the average annual AOD concentrations per season and per year
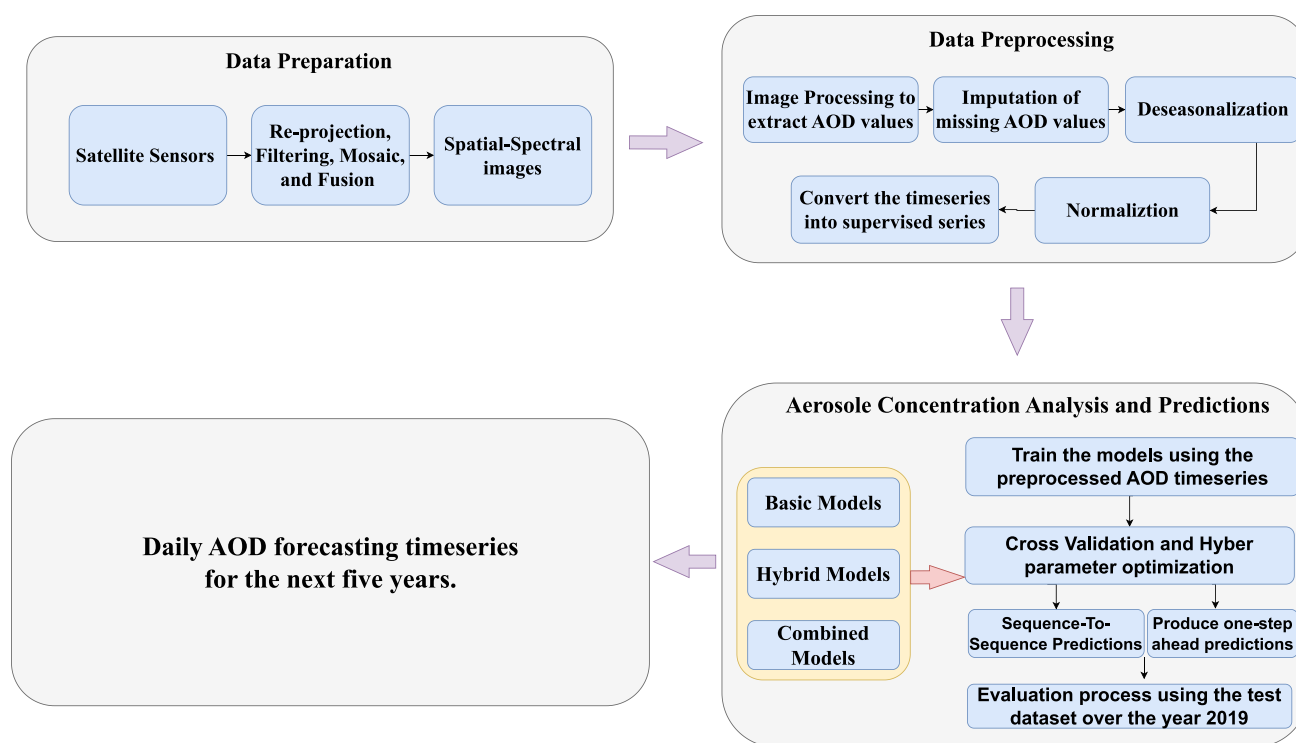
Fig. 2 The flowchart of the empirical methodology

by utilizing MODIS image data, which helped analyze the spatiotemporal variability of AOD. The statistical information obtained for each day was analyzed to identify multitemporal trends on a daily, weekly, monthly, and annual basis, which were then used as inputs for the time series analysis models.

It is worth noting that sandstorms and the existence of clouds can affect the quality of the data specifically during certain seasons such as in summer. These environmental events can introduce noise and errors into the data, which can lead to the presence of outliers (Ali et al. 2017).

3 Materials and Methods

This section aims to present the methodology used to model and forecast the AOD time series. Specifically, it can be summarized in four steps starting with data preparation and ending with daily AOD forecasting (as shown in Fig. 2). Data preparation included filtering, MOSAIC, fusion, and re-projection of the data from the satellite sensors. The second step corresponds to the pre-processing of the data by imputing any missing values, deseasonalizating, normalizating, and converting the data into supervised series. Next, the aerosol analysis and prediction step consists of training the models, cross validation and selection of hyperparameters. In what follows, we will present all the three steps of data

preprocessing, aerosol analysis and prediction, and Daily AOD forecasting.

3.1 Data Preprocessing

3.1.1 Data Pre-processing and Normalization

The AOD data, obtained from MODIS spanning January 2003 to December 2019, contained occasional missing values due to satellite coverage gaps and cloudy conditions. Given the relatively smooth nature of AOD changes over time and the absence of abrupt daily fluctuations, we employed a simple imputation technique (Nelson et al. 1999). Missing values were estimated by taking the average of the AOD values from one day before and one day after.

After addressing missing values, we deseasonalized and normalized the AOD time series. Deseasonalization, also known as seasonal adjustment, is vital for understanding the fundamental components of a time series, including trends and cyclical patterns. This process enhances prediction accuracy compared to using raw data (Golbraikh et al. 2003). Consequently, the models were trained and tested on deseasonalized AOD data. To further enhance model performance, we applied normalization to center the deseasonalized AOD time series. Normalization ensures that all features are on a similar scale, preventing feature bias due to differing magnitudes. In our study, we utilized MaxAbsScaler to normalize
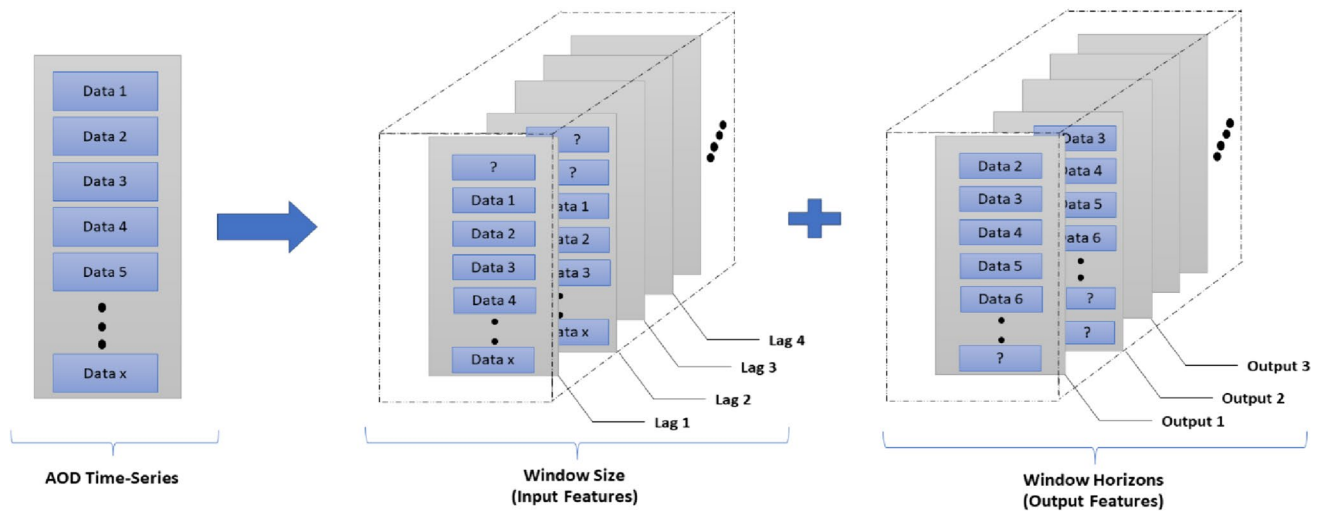
Fig. 3 Convert AOD data into Supervised Time-series

the deseasonalized AOD data. Each data point was divided by the maximum absolute value, transforming the values into the range [− 1, 1], as expressed by Eq. (11),

$$x_{norm} = \frac{x}{|x_{max}|} \tag{1}$$

where x_{max} represents the maximum value in the time series, x is the value to be normalized, and x_{norm} is the resulting normalized value. This normalization approach ensures consistent and comparable scaling of AOD values while preserving their distribution and relationships. It contributes to the meaningfulness and accuracy of our experiments and data analysis.

Following the preprocessing steps, the raw AOD data was transformed into supervised time series, as depicted in Fig. 3. For our study, we systematically experimented with lags up to 30, 60, or 90 lags. These data lags will be used as input features and the window horizons was set to 7. These configurations were strategically adopted to strike a balance between capturing relevant historical patterns and producing accurate future forecasts.

3.2 Daily Seasonality Adjustment of Aerosol Time Series

In the econometric literature, it is well documented that data preprocessing is a necessary step for a better understanding of the mechanism generating the time series under study (Ollech 2021; Bandara et al. 2021). Seasonality can affect the performance of forecasting models by introducing predictable patterns and dependencies in the data. Therefore, to address this issue deseasonalization must be performed on the data to preprocess it before feeding it to the models. This process removes the seasonal component from the data, making it easier for the models to capture the underlying trend and irregularities. The reasoning behind this is that if regular and irregular fluctuations that are due to seasonality, outliers and/or holidays are not removed they can obscure and hide important features of the data generating process of the time series under examination. Moreover, in some cases the type and extend of these regular and irregular fluctuation can be themselves of particular interest for policymakers and government when proposing policies that will help in improving environmental quality.

While there exists several approaches for extracting the trend and seasonality from low frequency data (month or quarterly), only a few of these approaches were proposed to deal with high frequency data (Ollech 2021; Bandara et al. 2021). Recently, the daily seasonal adjustment (DSA) and the multiple seasonal-trend decomposition using loess (MSTL) methods developed by Ollech (2021); Bandara et al. 2021) respectively have shown their high performance in better deseasonalizing high frequency time series, with application ranging from daily/hourly electricity consumption, daily currency in circulation, and daily NO_2 pollution.

In this paper, we used the DSA method as the main approach to remove the seasonality from the aerosol time series. In what follows, we briefly introduce the DSA method and we refer readers to Ollech (2021) for more details. Both methods, the DSA and MSTL, have as starting points the seasonal-trend decomposition using Loess (STL), where the original time series is decomposed into three components as follow:

$$Y_t = TC_t + S_t + I_t, \quad \forall t = 1, \ldots, T \tag{2}$$

where Y_t represents the original time series (the aerosol time series in our case), TC_t is the trend-cycle, S_t is the seasonal component and I_t is the adjusted time series (irregular component). While, this STL technique is widely used when adjusting for single seasonal patterns, it cannot accommodate for multiple seasonal patterns (such as intra- weekly, intra-monthly and intra-annual effects). To accommodate for multiple seasonality, Ollech (2021) proposed the DSA technique which extends the standard STL model given in Eq. (1) to include the four new components as follow:

$$Y_t = TC_t + S_t^{(7)} + S_t^{(31)} + S_t^{(365)} + C_t + I_t \quad \forall t = 1, \ldots, T \tag{3}$$

$S_t^{(7)}, S_t^{(31)}, S_t^{(365)}$ and C_t represent the intra-weekly, intra monthly, intra-annual and moving holiday effects, respectively. The DSA procedure is a four-step algorithm approach that combines the STL technique of Cleveland et al. (1990) with the regression ARIMA model (RegARIMA model).

The four-step algorithm of Ollech (2021) is given by,

- First, use the STL to adjust for intra-weekly periodic patterns.
- Second, adjust for the calendar effects, cross-seasonal effects, and outliers using the RegARIMA model.
- Third, use the STL method to adjust intra-monthly periodic effects.
- Finally, use the STL method to adjust for intra-annual effects.

For more technical details and empirical applications of the DSA procedure, see (Ollech 2021; Webel et al. 2023) and the DSA-vignette online R-procedure.

3.3 Basic, Hybrid and Combined Machine and Deep Leaning Models

3.3.1 Basic Models

The first group of models, known as the basic models, comprised of XGBoost, support vector regression (SVR), back propagation neural network (BPNN), long short-term memory (LSTM), bidirectional-LSTM, stacked-LSTM, and Bayesian ridge regression (BRR). These models directly take the raw data and learn how to map it to the corresponding output features using the various regression techniques. Such models, including subsequent architectures, are designed for time series data and have structures that can capture temporal dependencies. For instance, the LSTM model, and its variants, have memory cells that can learn and memories long-term dependencies in the data. Therefore, the violation of the assumption of independence among observations

due to autocorrelation in time-series forecasting can be accounted for through the structure of these models. These models analyzed and forecasted the future values of AOD time series implicitly without an explicit feature extraction component. This group hypothesizes that the models do not have a separate part responsible for the feature extraction process, but rather it is performed implicitly within the structure of the regression models. These models directly take the raw data and learn how to map it to the corresponding output features. General mathematical representation for these basic models can be represented as follow:

$$y = f(X;\Theta) + \varepsilon \tag{4}$$

where, y represents the target daily AOD concentration feature, X represents the feature matrix (including the independent variables), $f(X;\Theta)$ represents the forecasting function of the model driven by the list of parameters, Θ. The forecasting function can capture the complex dependencies between the independent and dependent features. The form of the forecasting function varies based on the type of the model. For SVR, the aim of the forecasting function f is to define a high dimensional hyperplane or decision boundary with the parameters Θ to minimize the prediction. For Bayesian ridge, f represents a linear regression model with the assumption of the Gaussian prior distribution on the coefficients:

$$f(X;\Theta) = X\beta \tag{5}$$

where β denotes the regression coefficients. For BPNN, f represents a multi-layer neural network with weights and bias represented by Θ. The regression output is obtained using the forward propagation method through the network:

$$f(X;\Theta) = \sigma\left(W^{(L)}\sigma\left(W^{(L-1)}\left(\ldots \sigma\left(W^{(1)}X + b^{(1)}\right)\ldots + b^{(L-1)}\right) + b^{(L)}\right)\right) \tag{6}$$

where σ is the activation function, and $W^{(i)} and b^{(i)}$ are the weights and biases of the layer i. In the case of XGBoost, $f(X;\Theta)$ represents an ensemble of decision trees with specific tree structures associated with a set of weights Θ. The final prediction output is obtained by combining the predictions of these trees. For LSTM, Bi-LSTM, and stacked-LSTM, the regression function $f(X;\Theta)$ includes adding more flexible techniques such as recurrent or convolutional layers to the neural networks parameterized by Θ. Finally, ε represents the error term, caused by the noise or residual in the observed dependent variable.

3.3.2 Hybrid Models

The second group, known as the hybrid models, involves stacking two models, one for feature extraction and one for regression, based on the extracted features. This group

Table 2 A concise summary of the feature extraction and regression components, hyperparameters, and other relevant details for each of the hybrid models

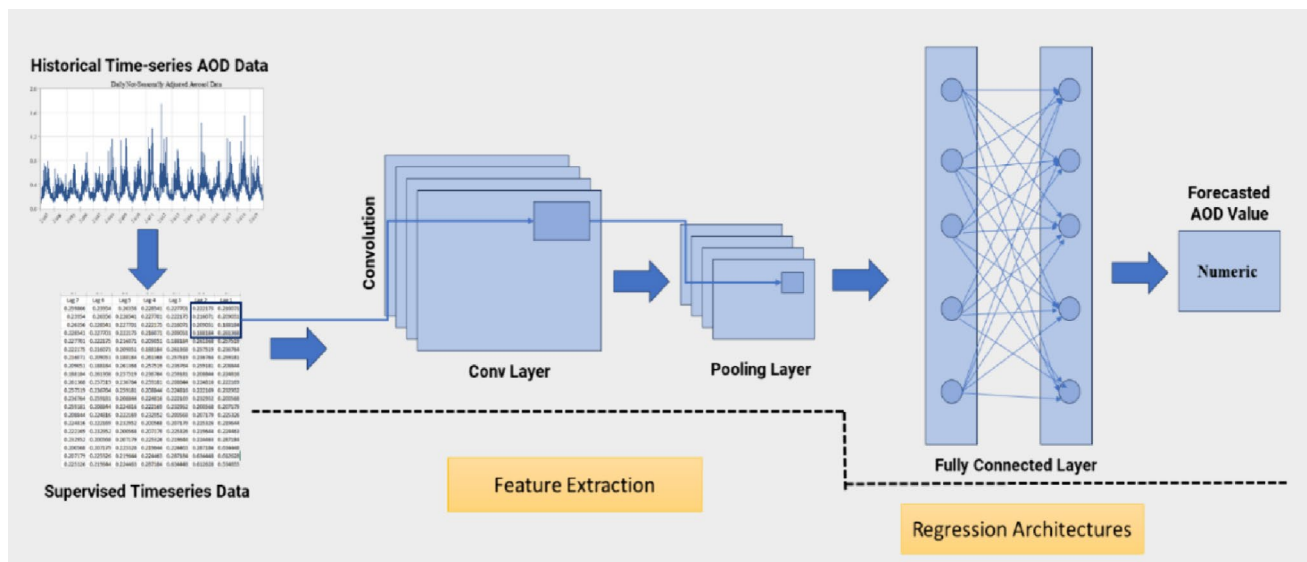| Hybrid models	Feature extraction method	Regression method	Key hyperparameters	Remarks/comments
CNN-LSTM	Convolutional Neural Network (CNN)	LSTM	CNN architecture, number of LSTM units, number of epochs, patch size…	Deploys CNN for feature extraction
ConvLSTM	Convolutional LSTM	Dense layer (fully connected layer)	Number of ConvLSTM layers, kernel architecture, number of epochs, patch size, loss function,	Integrates spatial–temporal features together
Encoder-decoder LSTM	Encoder-decoder LSTM	LSTM	Number of LSTM units for encoding and decoding, loss function, Batch size,	Encoder-decoder architecture for modeling
ConvLSTM-Bayesian Ridge	Convolutional LSTM	Bayesian Ridge	Kernel size, activation function, number of filters, Ridge alpha	Combines ConvLSTM with BayesianRidge

Fig. 4 The architecture of the Hybrid category models

includes CNN-LSTM, convolutional-LSTM (ConvLSTM), Encoder-Decoder LSTM, and ConvLSTM-BayesianRidge. The hybrid models incorporate feature extraction and regression into a single framework, following a two-step approach. First, the models learned the most informative features from the input time series. Second, the extracted features are passed as input into the regression part of the model to forecast the targeted output. Table 2 displays a succinct overview of the feature extraction and regression components, hyperparameters, and other pertinent details for each hybrid model. In the case of CNN-LSTM, a convolutional layer is employed to extract features from the input time series, which are then fed into the LSTM layer for regression. Another example is the encoder-decoder LSTM, which consists of an encoder to extract features and a decoder to

perform the regression. ConvLSTM combines the functionality of convolution and LSTM layers to simultaneously extract the features and perform the regression in a single layer. Lastly, ConvLSTM-BayesianRidge employs ConvLSTM layers for feature extraction, and Bayesian ridge for regression.

Figure 4 illustrates the core architecture of these hybrid models, incorporating both feature extraction and regression components, culminating in the final layer that produces the numerical AOD forecasts. The first phase, feature extraction, includes converting the input data X into a set of meaningful features Z. These extracted features can be thought of as learned patterns or representations from the input data, holding discriminative information such as statistical and temporal features, frequency domain features, nonlinear

dynamic patterns, etc. The process of feature extraction can be represented as a sequence of consecutive operations, involving convolutional layers, pooling layers, and non-linear transformations, which they are common techniques in many deep learning models such as CNN-LSTM, ConvLSTM, Encoder-Decoder LSTM, etc. The mathematical representation of the feature extraction component can be presented in the following equations:

$$
\begin{aligned}
Z_1 &= \text{Conv}_{\Theta_{\text{conv}_1}}(X) \\
A_1 &= \sigma(Z_1) \\
P_1 &= \text{Pool}_{\Theta_{\text{pool}_1}}(A_1) \\
Z_2 &= \text{Conv}_{\Theta_{\text{conv}_2}}(P_1) \\
A_2 &= \sigma(Z_2) \\
P_2 &= \text{Pool}_{\Theta_{\text{pool}_2}}(A_2) \\
&\vdots \\
Z_n &= \text{Conv}_{\Theta_{\text{conv}_n}}(P_{n-1}) \\
A_n &= \sigma(Z_n) \\
P_n &= \text{Pool}_{\Theta_{\text{pool}_n}}(A_n) \\
\\
Z &= \text{Flatten}(P_n)
\end{aligned}
\tag{7}
$$

where Z_i is the intermediate feature maps at layer i, A_i is obtained after applying a non-linear activation function (σ) to Z_i, P_i is the pooled features map obtained through pooling operations, $\text{Conv}_{\Theta_{\text{conv}_i}}$ is the convolution operation with learnable parameters Θ_{conv_i}, $\text{Pool}_{\Theta_{\text{pool}_i}}$ is the pooling operation with learnable parameters Θ_{pool_i}, and Z is the final flattened feature representation. The second component was, regression. In this stage, the relationship between the extracted features Z and the final target variable y were modeled using a specific regression structure such as a neural network, which can be a combination of recurrent layers followed by fully connected layers (i.e., LSTM), or different regression structures such as Bayesian ridge structures. The architecture of the LSTM regression model can be represented mathematically as follows:

$$
\begin{aligned}
h_1 &= \text{LSTM}_{\Theta_{\text{LSTM}_1}}(Z) \\
h_2 &= \text{LSTM}_{\Theta_{\text{LSTM}_2}}(h_1) \\
&\vdots \\
h_m &= \text{LSTM}_{\Theta_{\text{LSTM}_m}}(h_{m-1}) \\
\hat{y} &= \text{Dense}_{\Theta_{\text{dense}}}(h_m)
\end{aligned}
\tag{8}
$$

where h_i is the hidden state at LSTM layer i, $\text{LSTM}_{\Theta_{\text{LSTM}_i}}$ represents the LSTM layer with learnable parameters Θ_{LSTM_i}, Θ_{LSTM_i} represent the parameters of the LSTM layers, while $\hat{y}$ is the predicted output, which is obtained through a fully connected (dense) layer with parameters Θ_{dense}.

3.3.3 Combined Model

The utilization of combined forecasting models, which integrate the strengths of the sub models, has garnered substantial attention in recent research (Han et al. 2021; Zhang, et al. 2021). In the context of this study, we explore the amalgamation of multiple models selected from the basic and hybrid categories into a unified framework for AOD prediction. This approach capitalizes on the exceptional predictive abilities and distinct error distributions exhibited by the constituent models. Two distinct types of combined models were considered in this study to harness the collective predictive power of these sub-models:

Stacking generalization-based model: In this approach, the forecasting outputs from the constituent models (base learners) are used as input features, alongside the target AOD as output feature, to train a higher-level meta-learner model. Two models were selected as base learners due to their optimal performance in the Basic and Hybrid model categories: Bayesian ridge and ConvLSTM, The lagged values of the AOD time series was used as input features for the base learners. Experimentation of various lags and window sizes was performed (e.g. 30, 60, and 90) to find the optimal feature set for each model. Additionally, MaxAbsScaler was used to normalize the features before feeding them to the models. Various machine learning and deep learning models were tested as potential meta-learners in this study. The final meta-learner model chosen was an artificial neural network that constitutes of three layers: an input, hidden, and output layer. The choice of this meta-learner was based on its simplicity and robustness. The meta-learner is trained to autonomously learn how to effectively combine the predictions from the constituent models (base learners) to generate the final predictions. The optimal hyperparameters of the meta-learner were selected through a grid search methodology. The following is a mathematical representation of the stacked generalization model:

$$
P = \text{Meta Learner}\left(F_1, F_2, \ldots, F_M, Y\right).
\tag{9}
$$

The forecasting outputs from the constituent models (F_i, where $i = 1, 2, \ldots, M$), alongside the target AOD(Y) as output variable, are used to train a higher-level meta-learner model. The meta-learner combines these inputs F_i to produce the final prediction (P).

Weighted averaging-based model: in this model, the forecasting outputs from the constituent models are assigned weights. This process penalizes weaker-performing models while rewarding stronger ones. Subsequently, the weighted forecasts are averaged to produce the final predictions. The optimal weights for the constituent models were determined using a grid search methodology. The mathematical representation of the weighted averaging model is as follows:

$$P = \frac{\sum_{i=1}^{M} W_i \cdot F_i}{\sum_{i=1}^{M} W_i} \tag{10}$$

The forecasting outputs from the constituent models (F_i, where $i = 1, 2, \ldots, M$) are assigned weights (W_i) based on their performance. These weights are then used to find the weighted average of the forecasts, which produce the final predictions (P). The two base learners chosen for this technique were similar to the ones used in the stacking generalization-based model scheme. Namely, Bayesian ridge and ConvLSTM. Collectively, these combined models seek to improve the overall predictive accuracy by intelligently integrating the insights from heterogeneous sub-models within a unified framework.

3.3.4 Hyperparameter Optimization, Forecasting Specification and Evaluation

To ensure optimal performance of the adopted machine learning models on the AOD time series data, hyper parameter tuning is crucial during the model design and training phases (Fabianpedregosa et al. 2011). Hyperparameter optimization plays a pivotal role in this context, involving the fine-tuning of parameters such as the number of layers, layer sizes, activation functions, learning rates, regularization rates, and more. To systematically optimize these hyperparameters for each model, we employed the grid search methodology. This method enables the iterative exploration of diverse hyperparameter combinations while evaluating model performance on a dedicated validation set. Grid search is a simple and effective method that exhaustively searches over a pre-defined set of hyperparameter values to select the best combination based on a scoring function. Its advantage is its easy interpretability and implementation.

By refining our models based on feedback from the validation set, our goal was to ensure that these models effectively capture the intricate patterns and characteristics present in the AOD data without succumbing to overfitting. Overfitting, a phenomenon where a model becomes excessively tailored to the validation data, can lead to poor generalization on unseen data. To mitigate this risk, we set aside an entirely unseen dataset, referred to as the test data, to comprehensively evaluate the overall performance of the optimized models.

Our dataset was meticulously divided into three distinct subsets: the training data, validation data, and testing data. The training data encompassed the time series spanning from 1/1/2003 to 31/1/2018. During the hyperparameter optimization process, the final 5% of the training data points were reserved for validation purposes. Finally, the testing data comprised the entire time series for the year 2019. This rigorous dataset division approach allowed us to train, validate, and assess the models in an unbiased and systematic manner.

To evaluate the model's performance and avoid overfitting, we employed a form of cross-validation suitable for time-series data, known as a rolling prediction scheme. In this technique, the model is initially trained on a portion of the data in which the performance of the model is then evaluated on the subsequent "window" of data. The model is then updated to include this window of data in the training set, then the performance is evaluated on the next window of data. Specifically, the validation data was used as one year and then shifted one year forward with the accumulated training data of the previous years. This process is repeated until all data have been used. This method preserves the temporal order of the data and avoids using future data to predict the past unlike k-fold cross validation. This provides a robust assessment of the model's predictive performance. This technique also addresses the violation of independence among observations due to autocorrelation since it considers the temporal dependencies in the data by using a moving window for training and prediction.

In order to provide a rigorous evaluation of the forecasting performance exhibited by the various models, this study employs several widely recognized evaluation metrics, namely root mean squared error (RMSE), mean absolute error (MAE), coefficient of determination (R^2), and Pearson correlation coefficient (r). The MAE metric quantifies the average deviation between the predicted values and the corresponding values. On the other hand, RMSE assesses the overall magnitude of the prediction errors. The R^2 metric offers an assessment of the proximity between the predicted values and the actual values. The calculation formulas for RMSE, MAE, R^2, and r are presented as Eqs. (11–14) in the following form:

$$MAE = \frac{1}{m} \sum_{i=1}^{m} \left| (y_i - \hat{y}_i) \right| \tag{11}$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2} \tag{12}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{m} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{m} (\bar{y}_i - y_i)^2} \tag{13}$$

$$r = \frac{\sum_{i=1}^{m} (y_i - \bar{y}_i)(\hat{y}_i - \bar{\hat{y}}_i)}{\sqrt{\sum_{i=1}^{m} (y_i - \bar{y}_i)^2} \sqrt{\sum_{i=1}^{m} (\hat{y}_i - \bar{\hat{y}}_i)^2}} \tag{14}$$

Fig. 5 Daily Aerosol data for the Arabic Peninsula region—not adjusted series
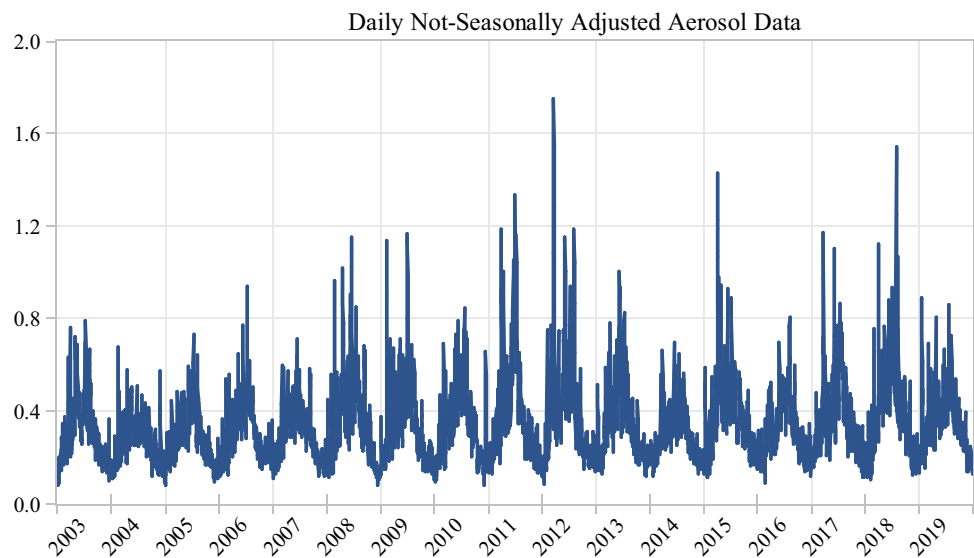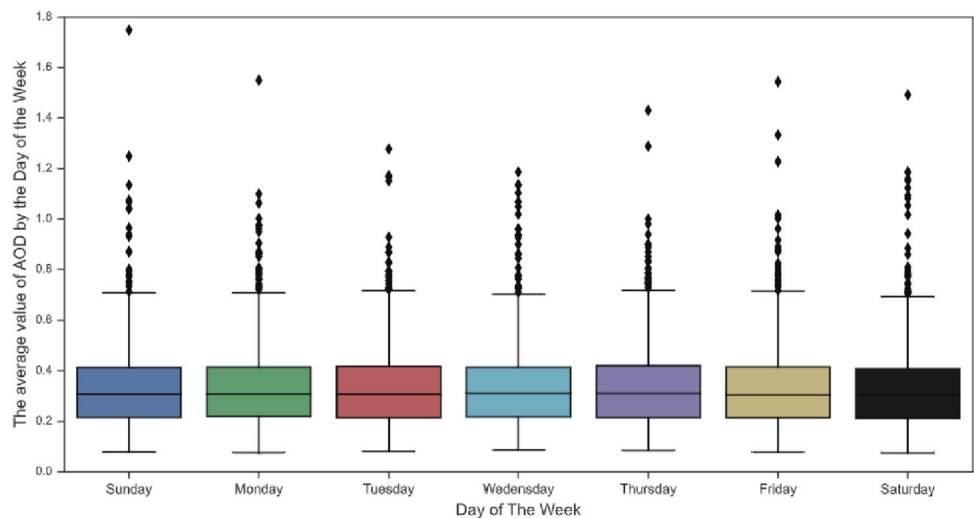

Daily Not-Seasonally Adjusted Aerosol Data

Fig. 6 The BoxPlot of AOD by the day of the week

where *m* is the number of data points in the test set; y_i is the actual value, the_i is the average value of the true values, *the* is the predicted value, and the_i is the average value of the predicted values.

The respective indexes are calculated for each day in the prediction horizon separately (1–7 days).

4 Results and Discussion

4.1 Aerosol Deseasonalization Results

The evolution of the daily aerosol data for the AP region is reported in Fig. 5. Regular patterns are observed in this daily time series, which clearly indicate evidence for seasonality. The figure shows that the period of October–March has the lowest values of AOD and the summer months show high

values of AOD. There is also strong evidence for the existence of outliers but no clear evidence for the existence of trend.[1]

An important step before deseasonalizing the aerosol time series is to analyze and determine the different types of seasonality that characterize the aerosol time series. To do this, we used boxplots to analyze day of the week, day of the month and day of the annual effects. The boxplots

---

[1] A results not reported here and can be obtained upon request from the corresponding author confirm the results from previous studies that the distribution of the AOD time series is not normal and that in general does not show evidence for the existence of a trend. The parametric (Jarque-Bera) and non-parametric tests (Kruskal-Wallis tests) are used to test the normality Hypothesis of the AOD series and the tests of Kruskal-Wallis are used to explore potential existence of a trend.
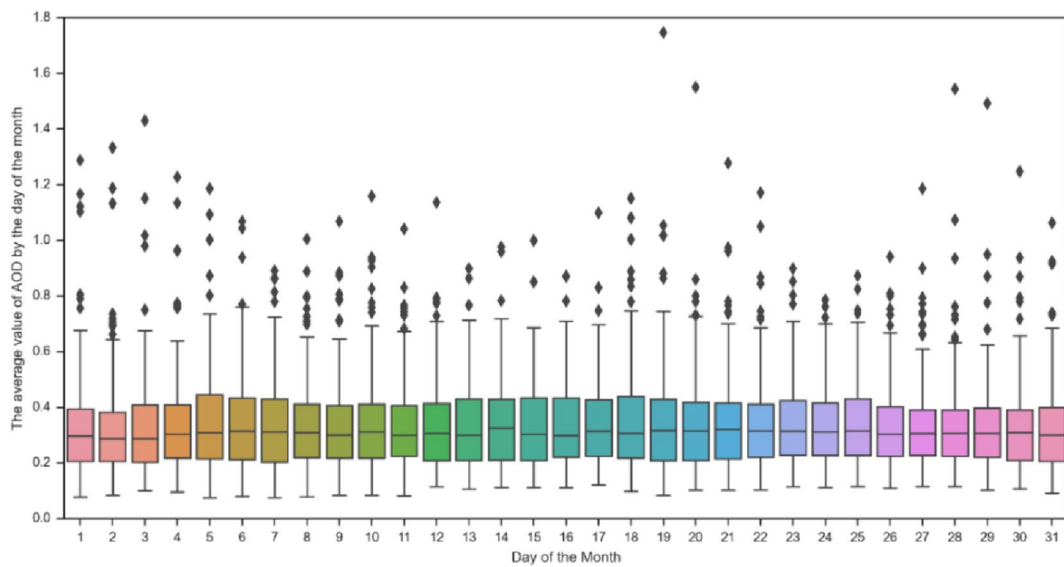
Fig. 7 The Boxplot of the AOD by the day of the month
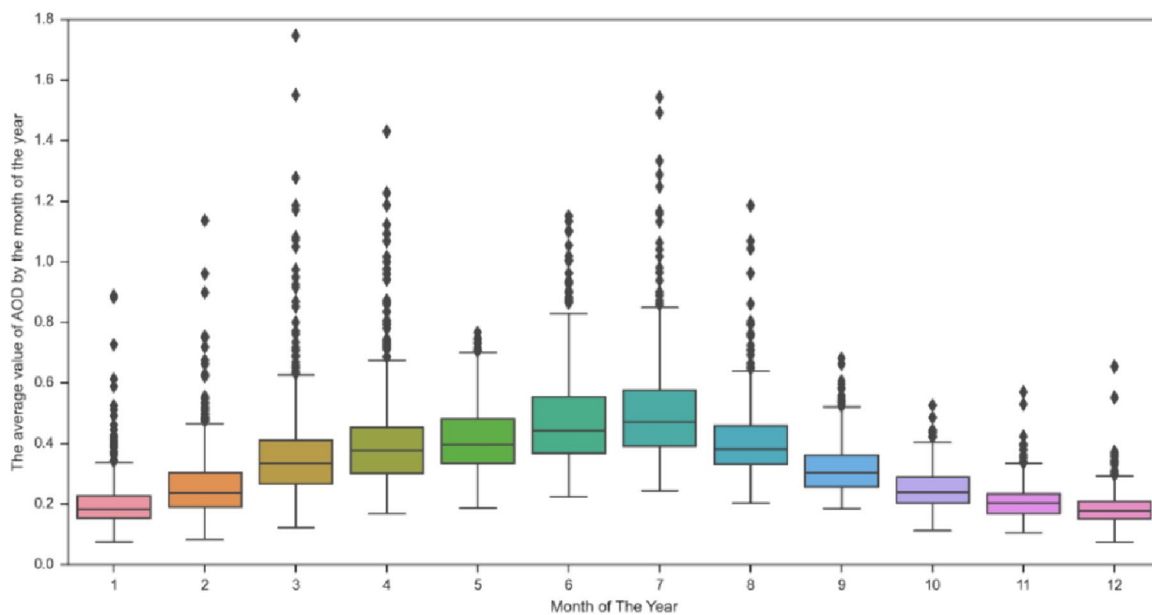

Fig. 8 The Boxplot of AOD by the month of the year

are reported in Figs. 6, 7 and 8, respectively. The day of the week effect shown in Fig. 6 does not show any evidence for changing effects over the days of the week as the median, first quantile, third quantile, maximum and minimum are similar for all 7 days. The consistent statistical metrics across all days suggest a lack of a "day-of-the-week" effect. This result indicates that aerosol levels do not substantially vary from one day to another within a week, which indicate absence of short-term and weekly cyclic activities in the AOD series. Consequently, in the DSA approach the day-of-the-week effect will be ignored.

To further deepen the analysis, we plotted the boxplot of AOD by the day of the month in Fig. 7. The results show slight fluctuation of the aerosol levels during the different days of the month. The results showed that the Q1, median and Q3 appeared to be different between all days of the month. The results also showed evidence of the existence of outliers since many observations exceed the Q3 level.

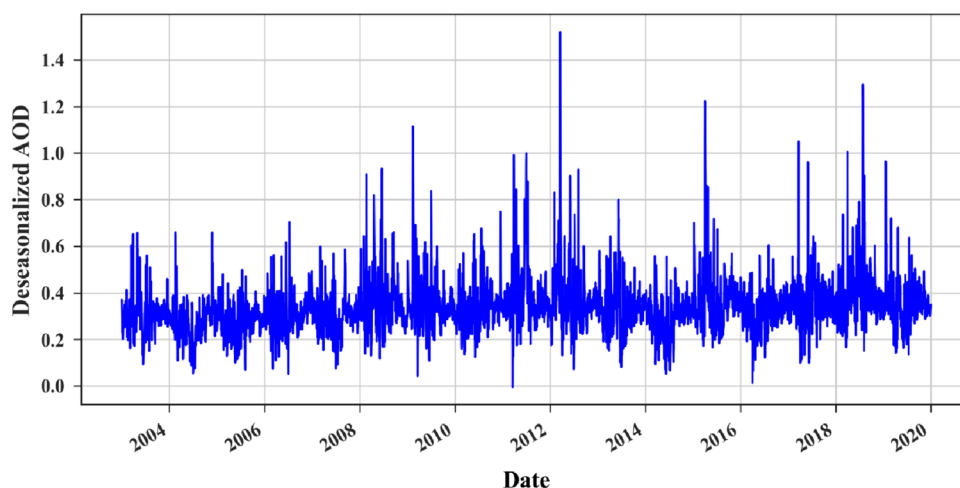**Fig. 9** Deseasonalized AOD data after applying DSA technique

Table 3 Optimized Hyperparameters of the various Deep learning Models in this study

Model	# Hidden layers	# of Neurons	Activation Function	Dropout Rate	Learning Rate	# of Epochs	Batch Size	MSE
BPNN	100	100	RELU	–	invscaling	–	–	**0.00162**
LSTM	1 LSTM layer	150 LSTM units	RELU	0.5	0.001	100	32	0.00203
BI-LSTM	2 PI-LSTM layer	140	RELU	0.0001	0.001	100	64	0.0018
Stacked-LSTM	2 Stacked -LSTM layer	100	RELU	0.0001	0.005	100	32	0.00198
CNN-LSTM	1 Conv1D + 1 LSTM	50 LSTM units	RELU	0.5	0.0001	100	32	0.00202
ConvLSTM	1 ConvLSTM	–	RELU	0.5	0.0001	50	32	**0.00167**
Encoder-Decoder LSTM	2 LSTM (Encoder, Decoder) layers	100	RELU	0.0001	0.001	100	32	0.00205

Note that for the case of CNN-LSTM and ConvLSTM the kernel size are 1 and 24 respectively and the number of filters are (1,2) and 32, respectively

Consequently, in the process of deseasonalization, we accounted for day of the month effects while it is not well pronounced.

Finally, by analyzing the boxplot for the month of the year effect (Fig. 8), the results showed strong and clear evidence that there exists some regularity where the 5th, 6th, 7th and 8th months in the AP region is characterized by high levels of AOD. The four months of February, March, April and September also show high levels of AOD but not as distinct as the four summer months. The higher AOD levels in the spring and summer months may be attributed to an increase in dust storms, industrial activity, and transportation emissions in the AP region. Furthermore, climatic conditions, such as temperature, precipitation, and wind speed also contribute to AOD levels. The rest of the four months showed evidence for small variability of AOD where the difference between Q1 and Q3 were minimal. Moreover, for all twelve months of the year, especially for the months of January, February, March, April, June, July and August there were substantial outliers in the data. Consequently, due to the clear evidence of seasonality in the data as observed in the month of the year boxplot, in the deseasonalization process, we took into consideration the existence of month of the year effects. The final time series deseasonalized is given below in Fig. 9.

4.2 Hyperparameters Optimization

Tables 3 and 4 present the optimized hyperparameters corresponding to the different models' structures used in this study. The hyperparameter configurations of the deep learning models were tuned on the dataset based on suggestions from previous studies (Arfin et al. 2023). Table 3 displays the tuned hyperparameters for the models of neural network-based structure along with the corresponding error measure. For these models, we optimized a range of hyperparameters including the number of hidden layers, the number of neurons within these layers, dropout rates, activation functions, learning rates, the number of training epochs, batch sizes, and further parameters dedicated to convolutional layers

Table 4 Optimized Hyperparameters of the various Machine learning Models in this study

Model	Hyperparameters Combinations	MSE
Bayesian Ridge	Max-Iteration = 300, alpha1 (shape parameter) = 1e-09, alpha 2 (inverse scale parameter) = 1e-09, lambda1 (shape parameter) = 1e-08, lambda 2 (inverse scale parameter) = 1e-09, tol (stop parameter) = 2.9 e-05	**0.00161**
SVR	Regularization parameter (C) = 1.0, epsilon = 0.1, kernel type (kernel) = rbf, kernel coefficient (gamma) = scale	0.00273
XGBoost	n_estimators (number of trees) = 200, max_depth (maximum depth of trees) = 3 eta (learning rate) = 0.03, subsample (Fraction of samples) = 0.9	0.00018

where applicable. A sensitivity analysis of hyperparameters was performed for each model by varying one hyperparameter at a time while keeping the others fixed at their optimal values and measured the change in the performance metrics. We found that some hyperparameters, such as the number of LSTM units, kernel sixe, and learning rate had a significant impact on model performance, while others, such as the batch size and dropout rate showed little impact.

The aim of this optimization is to strike a balance between the complexity of the models and the accuracy of their performance. For instance, the architecture of the BPNN model includes a relatively simple structure compared to other deep learning models with 100 hidden layers and 100 neurons per layer. The LSTM model, being a recurrent neural network, holds a single LSTM layer with over 150 units (neuros), enabling it to capture sequential dependencies efficiently. The bi-LSTM model with two bidirectional LSTM layers can handle both forward and backward temporal dependencies. The low dropout rate of 0.0001 indicates that it is crucial to keep information in both direction. The learning rate of 0.001 suggests a balanced learning approach. The batch size of 64 was used for more efficient computation. With this configuration, we aim to harness the bidirectional LSTM's capacity while preserving a low overfitting risk. Two stacked LSTM layers with 100 units each within the stacked-LSTM model can capture more complex patterns. Combining CNN and LSTM layers within the CNN-LSTM model with kernel size of one and 64 filters aims to capture spatial and temporal dependencies simultaneously. The selection of convolutional layer parameters (kernel size of (1,2) paired with 32 filters) demonstrates the importance of handling both spatial and short-term temporal patterns. The configuration of the encoder-decoder LSTM model was designed to capture complex dependencies for short-term forecasting.

Table 4 demonstrates the hyperparameter combinations corresponding to the machine learning models used in this study. More specifically, the structure of the Bayesian ridge suggests a relatively well-regularized model. The small values of alpha 1, alpha 2, lambda 1, and lambda 2, demonstrate the simplicity of the model. A maximum iteration of 300 indicates that the model converges efficiently. This configuration ascertains the balance between the model's complexity and reliability. The regularization parameter (C) of 1.0 in the SVR model indicates a moderate regularization, while the selection of radial basis function (rbf) kernel with a scaled gamma implies flexibility in handling complex dependencies. The defined configuration of the ensemble model (XGBoost) ascertains its simplicity. The moderate number of trees (n_estimators = 200) and a maximum tree depth of 3 helps in preventing overfitting and keeps individual trees shallow.

By systematically fine-tuning these hyperparameters, we verified that our models were functioning at their optimal performance, leading to the best possible predictive accuracy for atmospheric AOD predictions in the AP. This comprehensive optimization technique reinforces the effectiveness and reliability of our models in handling the temporal patterns of AOD. Hyperparameter optimization used to minimize the optimization loss can help prevent overfitting and improve the generalizability of the models on unseen data. The risk of overfitting the training data was further mitigated for all models by using regularization, dropout, early stopping, and cross-validation. Regularization techniques, such as drop out and weight decay, can reduce the complexity and variance of the models. Additionally, early stopping, used to monitor the validation loss, stops model training when loss starts to increase which enhances the models generalizability.

4.3 Model Evaluation

In the evaluation of the basic, hybrid and combined models, we focus our analysis in assessing the predictive accuracy of these models by using two predictive errors measures namely, RMSE and MAE. These metrics were utilized to gauge the models' performance over seven days prediction horizons. The conducted experiments run in windows 11 operating system, and the hardware device is Intel(R) Core I7-1065G7 CPU @ 1.30 GHz(8 CPUs), RAM: 16 GB, and the programming language is python 3.10.0. For higher performance, deep learning models were implemented on Google Colab platform with Keras API which is running on TensorFlow 2.12.0 (GPU version) as the backend (DiBattista et al. 2020). The rest of the models are implemented using Scikit-learn 1.0.2 (Djuric et al. 2016).
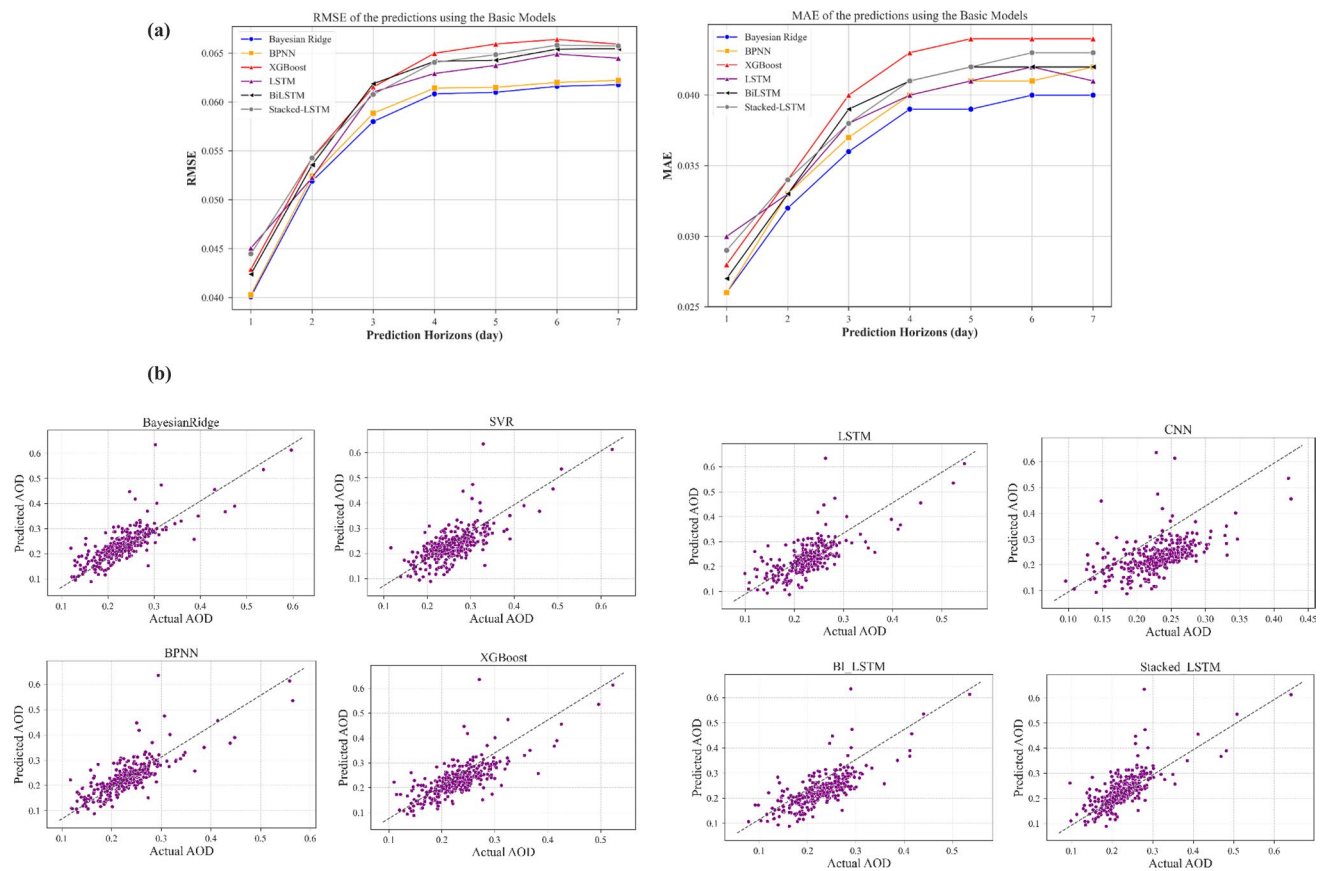
Fig. 10 **a** RMSE and MAE forecasting error measures for the best six basic models. **b** Actual vs predicted scatter plots of the basic models. The dashed line illustrates perfect predictions

4.3.1 Basic ML Models Results

The results of the RMSE and MAE for the six basic models are displayed in Fig. 10a. Focusing on the RMSE measure, reported in the top right of Fig. 10a, we find that the Bayesian ridge (blue color) exhibited the lowest RMSE values for all predicted horizons, with values gradually increasing over the seven days horizon. This indicates that Bayesian ridge provides the best forecasts of the actual AOD values and outperforms all other basic models. The results also showed that the BPNN model is the second-best choice after the Bayesian ridge as the forecast values are slightly higher than those of the Bayesian ridge. For the rest of the models, the results showed that their performance is weak when compared to Bayesian ridge with XGBoost having the worst forecasting performance. These findings are reinforced by the MAE results reported in Fig. 10a. The results showed that the Bayesian Ridge has the lowest MAE values, indicating its superior ability to forecast AOD accurately, followed by the BPNN model. Again, the forecasting performance of the rest of the models is weak and XGBoost has the lowest forecasting performance.

The high performance of the Bayesian ridge and the BPNN models compared to the XGBoost, LSTM, bi-LSTM, and stacked-LSTM models is expected, especially for applications requiring precise and reasonable short-term predictions. On the other hand, the lower variability of the XGBoost, LSTM, bi-LSTM, and stacked-LSTM models might be attributed to the inherent complexity of these models, leading to an increased sensitivity to data fluctuations. However, these models can still be valuable in scenarios requiring more complex patterns or longer-term predictions. Our results of the outperformance of the Bayesian ridge and the BPNN models are aligned with our expectations.

For a better analysis of the forecasting performance of the basic models, we report in Fig. 10b the scatter diagram of the actual versus predicted values of the AOD for one day ahead. The dashed line represents perfect predictions, where the actual and predicted values are equals. The closer the points are to the dashed line, the more accurate the predictions are. The points that deviate from the line indicate prediction errors. Visually, the two models, Bayesian ridge and the BPNN models, show better alignment with dashed
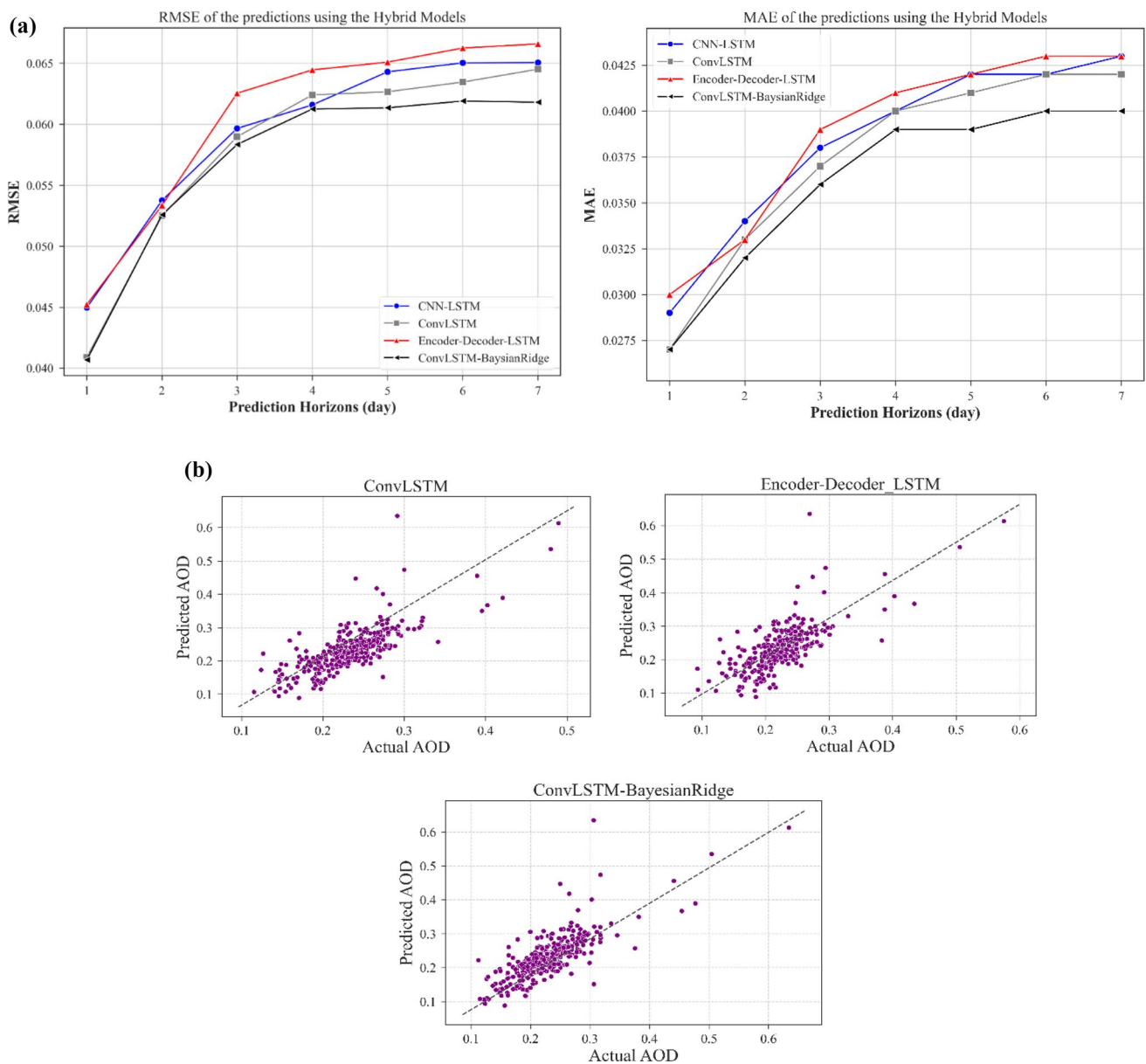
Fig. 11 **a** RMSE and MAE forecasting error measures for the best four hybrid models. **b** Actual vs predicted scatter plots of the hybrid models. The dashed line illustrates perfect predictions

line. The worst for the one day ahead are reported for the CNN, SVR and XGBOOST.

4.3.2 Hybrid ML Results

The forecasting errors, including RMSE and MAE, of the best four hybrid models are presented in Fig. 11a. First, the CNN-LSTM model displayed varying RMSE values across different predictions horizons (blue color), with values ranging from 0.0045 (day 1) to 0.065 (day 7). In the same vein, the MAE results for the CNN-LSTM model vary depending on the horizon time. In contrast, the RMSE values provided by the ConvLSTM model were lower, especially on days 1 to 4, where it outperforms the CNN-LSTM model. This proves that ConvLSTM, with its ability to capture spatial and temporal dependencies, outperforms the other models in short-term predictions. The encoder-decoder LSTM model has the highest RMSE and MAE values showing that this model has the lowest performance in terms of predictions. On the other hand, the ConvLSTM-BayesianRidge combined model shows intriguing results (black color), which consistently outperforms all the other hybrid models in terms of both RMSE and MAE error predictions measures through the different prediction horizons. The superior

predictive performance of the ConvLSTM-BayesianRidge model is due to its ability to minimize both the variance and bias of the forecasts.

Figure 11b shows the scatter plots of the actual versus predicted AOD values for the four hybrid models: ConvL-STM, Encoder-Decoder LSTM, and ConvLSTM-BayesianRidge. The figure reveals that the ConvLSTM-Bayesian-Ridge model has the highest concentration of points near the dashed line, indicating its superior performance in forecasting AOD. The other models have more scattered points, especially for higher AOD values, suggesting lower accuracy and higher errors. The figure also shows that the ConvLSTM model tends to underestimate the AOD values, while the Encoder-Decoder LSTM model tends to overestimate them. This implies that these models have some biases in their predictions, which can affect their reliability and usefulness.

The outperformance of the ConvLSTM-based models can be explained by its strong ability to capture spatio-temporal data. The integration of Bayesian ridge, the best-performing model among the basic models' category, for the regression task, combined with ConvLSTM for feature extraction, substantially enhances their forecasting capabilities. Moreover, by combining the RNN variant models with Bayesian methodologies to enhance forecasting has gained recognition from various studies in machine learning literature. Notably, Bayesian LSTM architectures have been deployed to improve forecasting accuracy, as introduced by Yarin, Ghahramani et al. (Gal and Z. 2016). Their findings suggest that grounding dropout in approximate Bayesian inference provides insights into its effective use with RNN models. These findings are consistent with other research studies that have examined the AOD time series or other air pollution indicators such as PM 2.5 among others (Daoud et al. 2021; Tian and Chen 2022; Han et al. 2021). Our results are also in line with prior findings in precipitation nowcasting, where ConvLSTM is found to have high performance (Levy et al. 2015). Furthermore, we found that the introduction of the CNN model in the hybrid group improved the fitting degree of the prediction models.

4.3.3 Combined ML Results

Figure 12a displays the RMSE and MAE measurements of the combined models. It is evident that both the stacked combination of ConvLSTM and Bayesian ridge (red color) and the weighted average models consistently outperform the individual basic and hybrid models, including Bayesian ridge and ConvLSTM models, across the majority of prediction horizons. These findings vividly demonstrate their outstanding capacity to produce remarkable precise forecasts for AOD in the AP. The stacked combining model exhibits RMSE values ranging from 0.0402 (Day 1) to 0.0618 (Day 7). Likewise, the weighted average model produces RMSE

values in the range of 0.0399 (Day 1) to 0.0617 (Day 7). This shows that the performance of the weighted averaging strategy is slightly better than stacking strategy for combining the prediction of ConvLSTM and Bayesian Ridge for the short-term AOD prediction task.

Figure 12b shows a scatter plot of the actual vs predicted AOD values for the two combined models: stacked combining and weighted average. The plot indicates that both models have high accuracy, as most of the points are clustered around the line. However, the weighted average model seems to have a slight edge over the stacked combining model, as it has fewer outliers and a higher correlation coefficient. This suggests that the weighted average model is more consistent and reliable in forecasting AOD values.

The superiority of the weighted average model in terms of forecast could be attributed to its ability to assign different weights to the predictions of the ConvLSTM and Bayesian ridge based on their individual performance. This flexible approach allows the model to capitalize on the strengths of both constituent models while mitigating their weaknesses, resulting in more accurate AOD forecasts. Furthermore, these combined models, especially the weighted average model, hold significant promise for practical applications in fields like environmental planning and climate research, where precise short-term AOD predictions are of great value. The superior performance of the combined models aligns with previous research indicating that ensemble methods, which combine the predictions of multiple models, can often outperform individual models. This is especially true in situations where the individual models have complementary strengths and weaknesses. The high performance of averaging techniques have been highlighted in several studies including modelling and forecasting several air pollution proxies and meteorological variables such as PM10, CO, NO_2, NO_x, SO_2, and O_3, and five meteorological variables, wind speed, wind direction, rainfall accumulation, temperature and solar radiation (see (Westerlund et al. 2014)). Our results are also aligned with other studies that found evidence of high performance of averaging techniques when analyzing the impact of fine particle on cause-specific mortality (see (Fang et al. 2016; Tran et al. 2018; Ramli, et al. 2023)).

4.3.4 Enhancing AOD Forecasting Precision Through Models Combination

Figure 13 provides a general insight into the Kernel Density Estimation (KDE) of the forecasting errors of the three classes of models. Notably, the curves exhibited by the combined models show a leptokurtic density with a mean almost equivalent to 0. This density pattern indicates that these models produce errors that are closely concentrated around 0, showing a robust and consistent performance. In
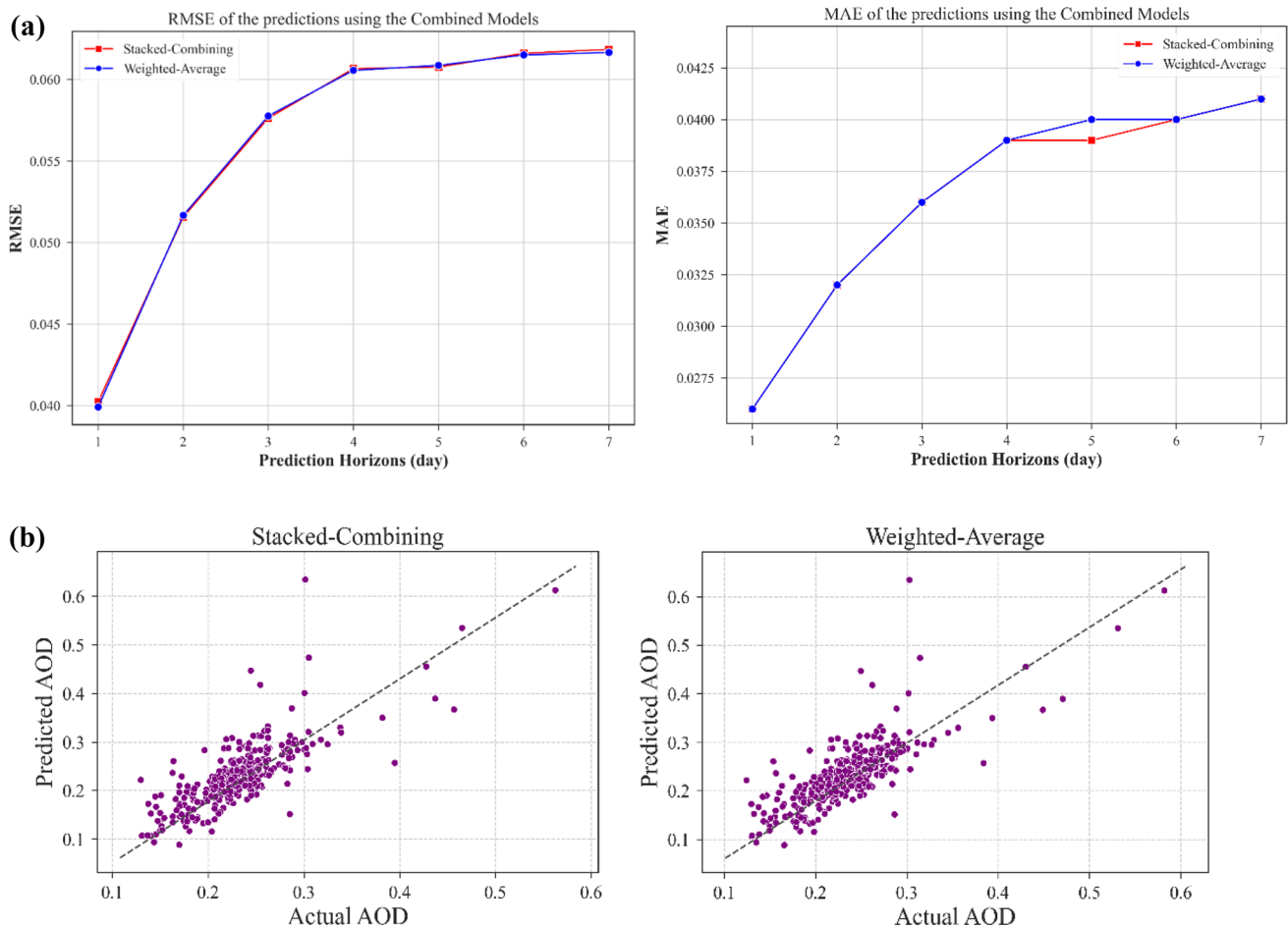
(a)

(b)

Fig. 12 **a** RMSE and MAE forecasting error measures for the two combined models. **b**: Actual vs predicted scatter plots of the combined models. The dashed line illustrates perfect predictions

contrast, the KDE curves for the majority of the other models are comparatively scattered, leading to a broader dispersion of errors.

The Pearson's correlation coefficient for all models across the seven-day forecasted horizon are shown in Fig. 14. The highest correlation was produced by the weighted-averaging combined model followed by the stacked combining and Bayesian ridge models. The performance of the models starts to vary as the prediction horizon gets closer to day seven. The most significant difference can be seen on the last day where the weighted-averaging model shows the highest correlation between actual and predicted AOD values compared to the rest of the models. All Pearson's correlation coefficient values of the combined models are statistically significant ($p < 0.01$) as shown in Fig. 15. This shows that combining the strength of the ConvLSTM model for capturing spatiotemporal dependencies along with the Bayesian ridge models ability to perform well for short-term forecasting results in more precise prediction of AOD values, especially for further ahead predictions where other models

tend to skew from the actual AOD values. This validates our hypothesis that combing the best-performing models can result in more optimal and accurate forecasting predictions.

4.4 Percentage Improvement Results

Table 5 shows the percentage improvement of RMSE for the weighted-average combined model relative to all other models. From this table, we can see that the RMSE scores were improved upon for almost all days in the prediction horizon due to the combination of ConvLSTM and Bayesian ridge through the use of the averaging strategy. The biggest improvement was a significant 51.15% for the seventh day prediction of the SVR model. In some cases, day one in the prediction horizon showed worse results for the weighted-average combined model compared to the rest of the models. However, a large improvement can be seen from day two and ahead. This shows that the combined weighted-averaging model perform better for further ahead predictions while other models tend to deteriorate. Furthermore, this
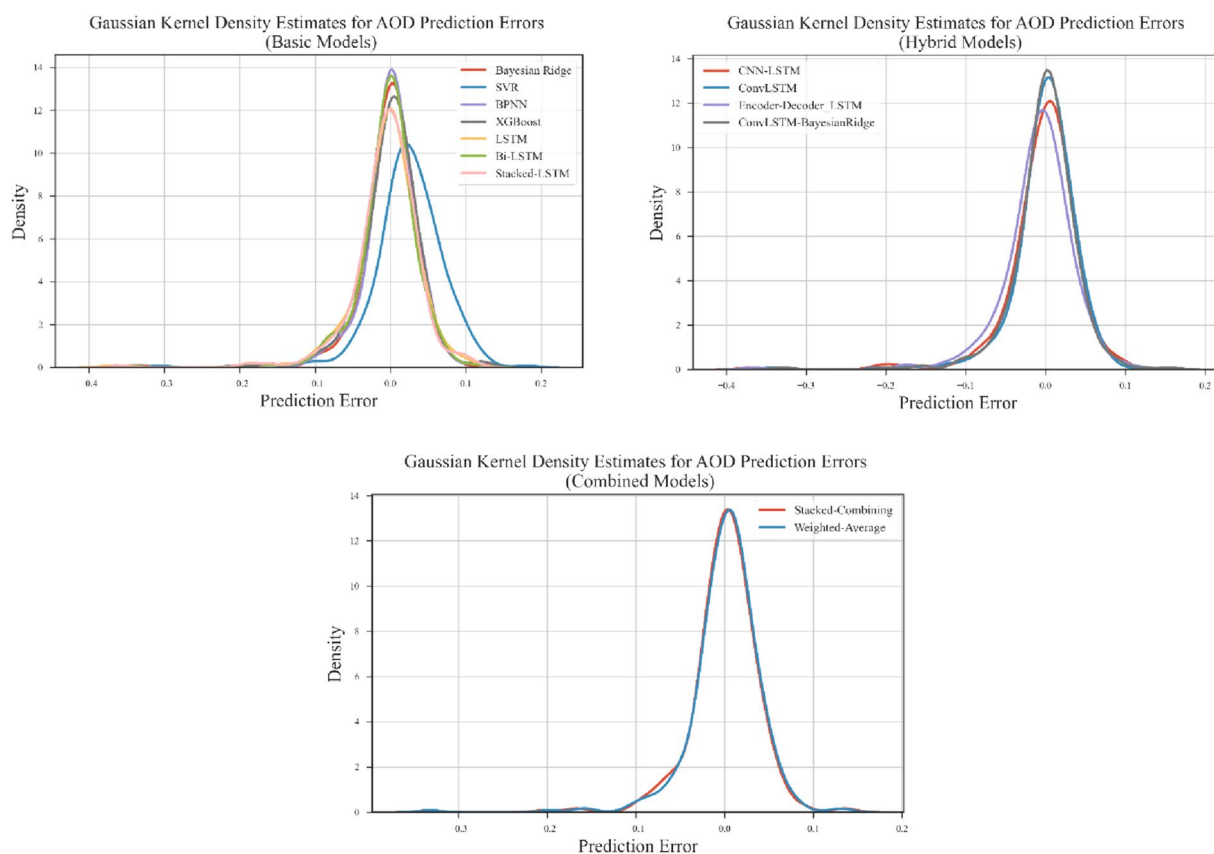
Fig. 13 Gaussian kernel density estimation of the prediction errors using the Basic, Hybrid, Combined models at all prediction horizons (1–7 days)

also signifies that the combined weighted-averaging model results in higher correlation between actual and predicted AOD values for future predictions, which can also be noted by the Pearson's correlation coefficient represented visually in Fig. 14.

Table 6 shows the percentage improvements of MAPE for the combine weighted-averaging model relative to all other models. Similar to the RMSE percentage improvement, MAPE had the highest improvement across all seven-day predictions when the models were combined compared to the performance of each one separately. The only deviation to this trend can be seen by the comparisons to the Bayesian ridge model in which improvement in MAPE scores can only be seen for day three and day four. However, the combined weighted-averaging model had higher performance in almost all other metrics showcasing that it was the top performing model overall for AOD predictions.

The goal of the combined models was to reduce prediction error by using the best performing basic models that are more suitable for small samples of univariant data with the best performing hybrid model that performs well with large samples of multivariant data. Thus, both combination schemes of the two models have the ability to take the best

overall prediction, which resulted in a more robust technique that has higher performance for both large and small sampled univariant and multivariant data. Overall, we found that the weighted-averaging technique to combine the two top-performing models significantly outperformed all other models on all metrics for predicting future AOD levels in the AP. Compared to a single model alone, the combination of the best-performing models makes up for certain shortcomings that either model alone would face. Compared to previously reported results for the task of AOD forecasting (Putaud et al. 2010; Arden Pope et al. 2011; Zhang et al. 2016; Song et al. 2018; Abuelgasim and Farahat 2019), we found that the weighted-averaging technique outperformed previous methods achieving higher accuracy rates and lower error than many of the existing methods. This shows the novelty and value of our proposed approach for AOD forecasting. Some possible factors that account for the superior performance of the models include the use of advanced combination techniques, the use of high-quality and high-resolution data, and the use of rigorous model evaluation and validation methods. Furthermore, our proposed models are not limited to AOD predictions only. They can be applied to other environmental time series forecasting problems. The
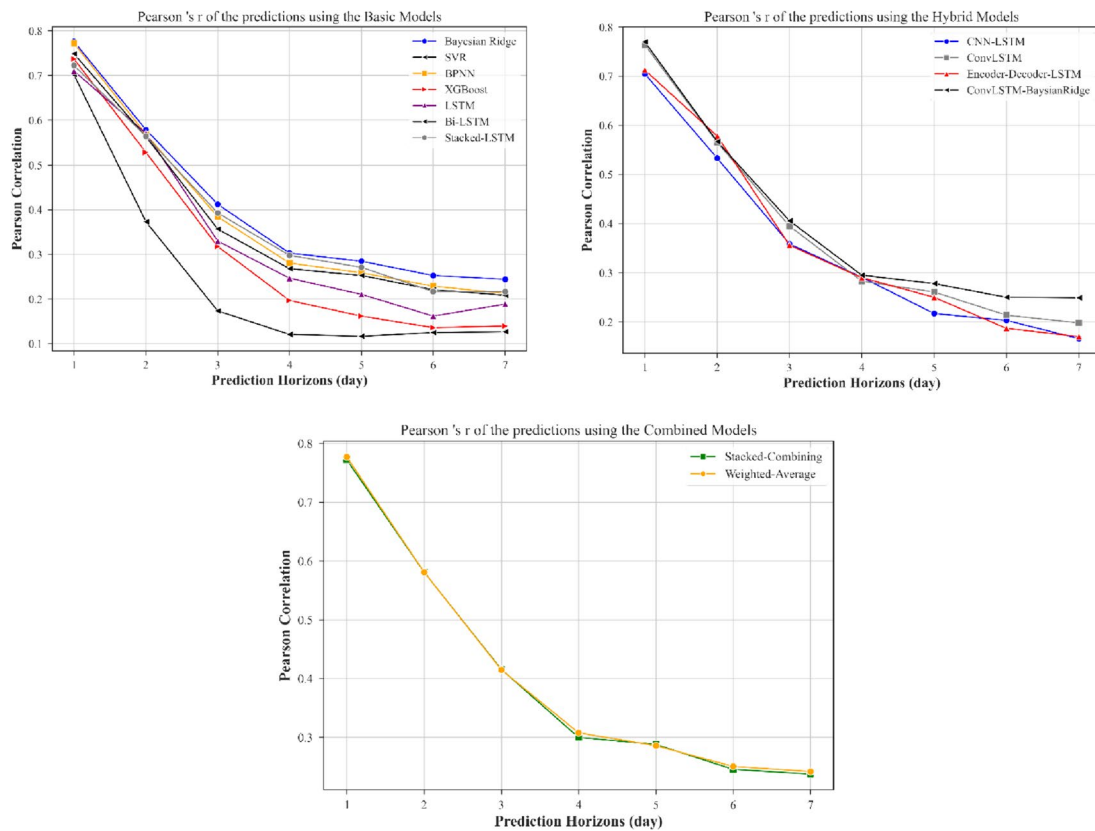
Fig. 14 Pearson's r of the predictions using the Basic, Hybrid, Combined models for different prediction horizons (1–7 days)

models are flexible and adaptable and can be easily modified to suit different data sources, features, and outputs. The models are scalable and robust in which they can handle different levels of data complexity, variability, and noise. Therefore, the models proposed in this paper have a wide range of applicability and usefulness for various environmental forecasting tasks and scenarios.

In our study, the decision to incorporate hybrid and combined models, despite their increased computational cost and complexity, was driven by the substantial enhancement in prediction accuracy they offered over basic models. While we acknowledge that the interpretability of combined models can be challenging due to their complexity, we would like to highlight that there are robust techniques available to address this concern without compromising performance.

For instance, permutation feature importance is a valuable tool we can employ to discern the impact of individual features on model predictions. By systematically permuting feature values and measuring the resulting decrease in model performance, we can identify the features that significantly influence the output. Additionally, visualization techniques such as saliency maps and attention maps provide a transparent view of the regions or patterns in the input data that the model prioritizes during prediction.

These methods serve as invaluable aids in unraveling the internal workings and logic of the models.

Overall, while the use of multiple models can indeed pose challenges to interpretation, our commitment to model transparency is evident in our planned utilization of these sophisticated interpretability techniques. We believe that by employing these methods, we can not only understand the nuanced contributions of individual models but also enhance the overall interpretability of our forecasting approach.

For long-term predictions, Fig. 16 displays the Aerosol Optical Depth (AOD) forecasting for six years ahead (blue line) from 2020 to 2026 with. The figure also shows the 95% confidence interval (shaded area) of the predictions, which indicates the uncertainty of the forecasts. The figure demonstrates that the proposed approach can capture the general trend and seasonal patterns of the AOD time series. The figure also reveals that the prediction errors tend to increase as the forecast horizon gets longer, which is expected for any time series forecasting problem. The figure illustrates the potential of the designed system for providing reliable and accurate AOD forecasts for the Arabian Peninsula.
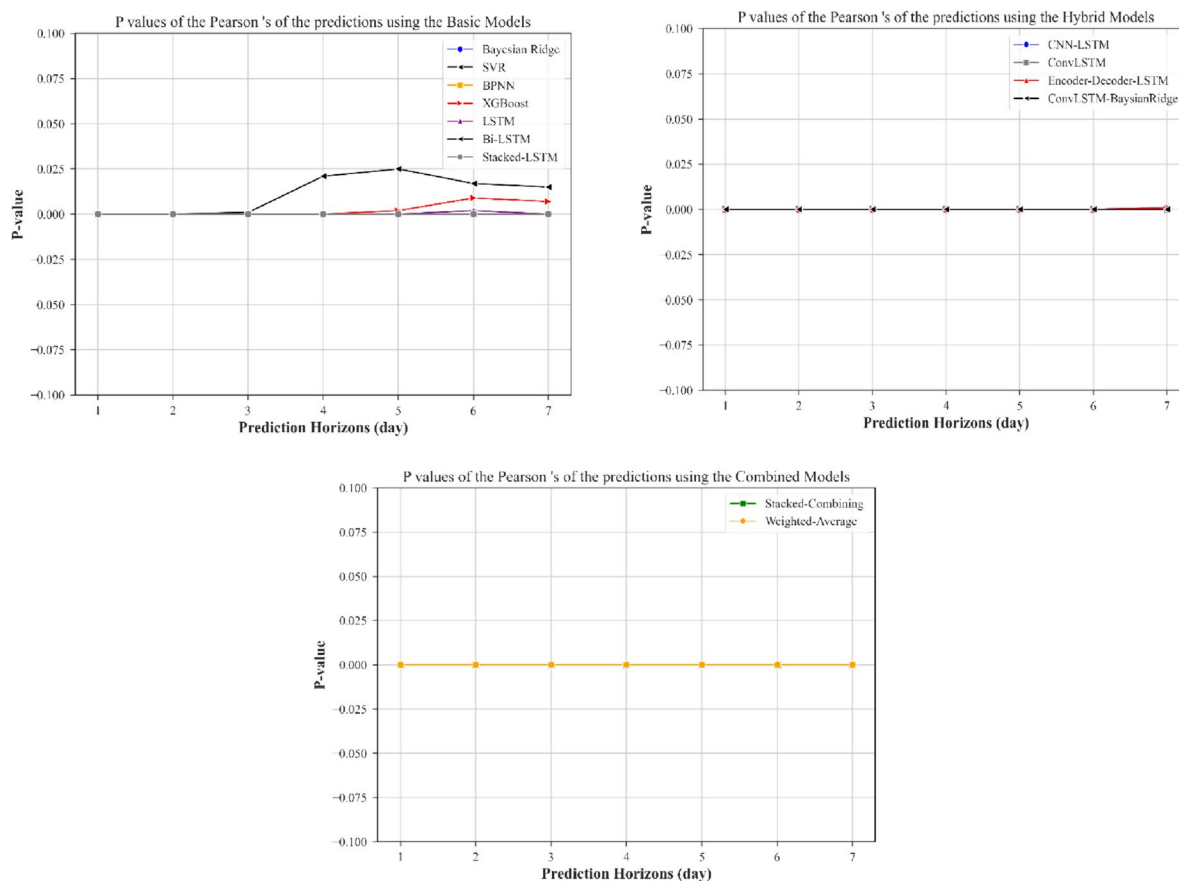
Fig. 15 P values for the Pearson's r of the predictions using the Basic, Hybrid, Combined models for different prediction horizons (1–7 days)

		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
Table 5 RMSE percentage improvements of the weighted-averaging model relative to all the other models	Basic models							
	Bayesian Ridge	0.436	0.455	0.416	0.464	0.211	0.139	0.200
	SVR	23.600	34.245	40.729	44.579	47.875	48.795	49.224
	BPNN	0.922	1.409	1.862	1.423	1.051	0.814	0.916
	XGBoost	6.965	4.835	6.130	6.816	7.682	7.358	6.475
	LSTM	11.414	1.150	5.354	3.763	4.530	5.241	4.374
	BI-LSTM	5.866	3.512	6.651	5.601	5.310	5.949	5.790
	Stacked-LSTM	10.263	4.772	4.971	5.471	6.135	6.528	6.217
	Hybrid models							
	CNN-LSTM	11.229	3.860	3.205	1.676	5.345	5.440	5.241
	ConvLSTM	2.435	1.584	2.092	2.979	2.867	3.074	4.453
	Encoder-Decoder_LSTM	11.738	3.158	7.671	6.058	6.520	7.176	7.418
	ConvLSTM-Bayesian Ridge	1.891	1.721	1.017	1.148	0.782	0.646	0.271
	Combined models							
	Stacked-combining	0.784	−0.223	−0.207	0.169	−0.171	0.157	0.301

Table 6 MAPE percentage improvements of the weighted-averaging model relative to all the other tested models

|	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
Basic models							
Bayesian Ridge	− 1.237	− 1.494	− 2.095	− 1.346	− 1.280	− 1.897	− 2.305
SVR	38.300	52.964	58.421	60.850	63.012	63.696	64.19202
BPNN	− 2.709	1.619	1.798	2.901	3.798	3.6308	3.483058
XGBoost	5.884	4.420	7.815	8.364	9.155	6.8168	7.891307
LSTM	10.713	− 0.167	1.363	− 2.511	− 0.93	− 2.069	− 3.56841
BI-LSTM	1.470	0.353	2.205	− 0.057	− 0.455	− 2.040	− 1.908
Stacked-LSTM	8.057	− 0.090	− 2.307	1.266	1.887	1.269	1.890
Hybrid models							
CNN-LSTM	8.107	4.528	2.521	0.285	2.108	− 2.103	1.854
ConvLSTM	1.909	1.058	− 1.156	− 2.187	− 0.596	0.033	− 1.268
Encoder-decoder_LSTM	12.478	0.139	2.199	-0.437	-0.52	1.140	− 0.657
ConvLSTM-Bayesian Ridge	− 0.330	− 0.890	− 1.471	− 1.475	− 1.880	− 1.607	− 2.620
Combined models							
Stacked-combining	− 0.372	− 0.272	− 0.705	0.017	− 0.309	0.318	− 0.011
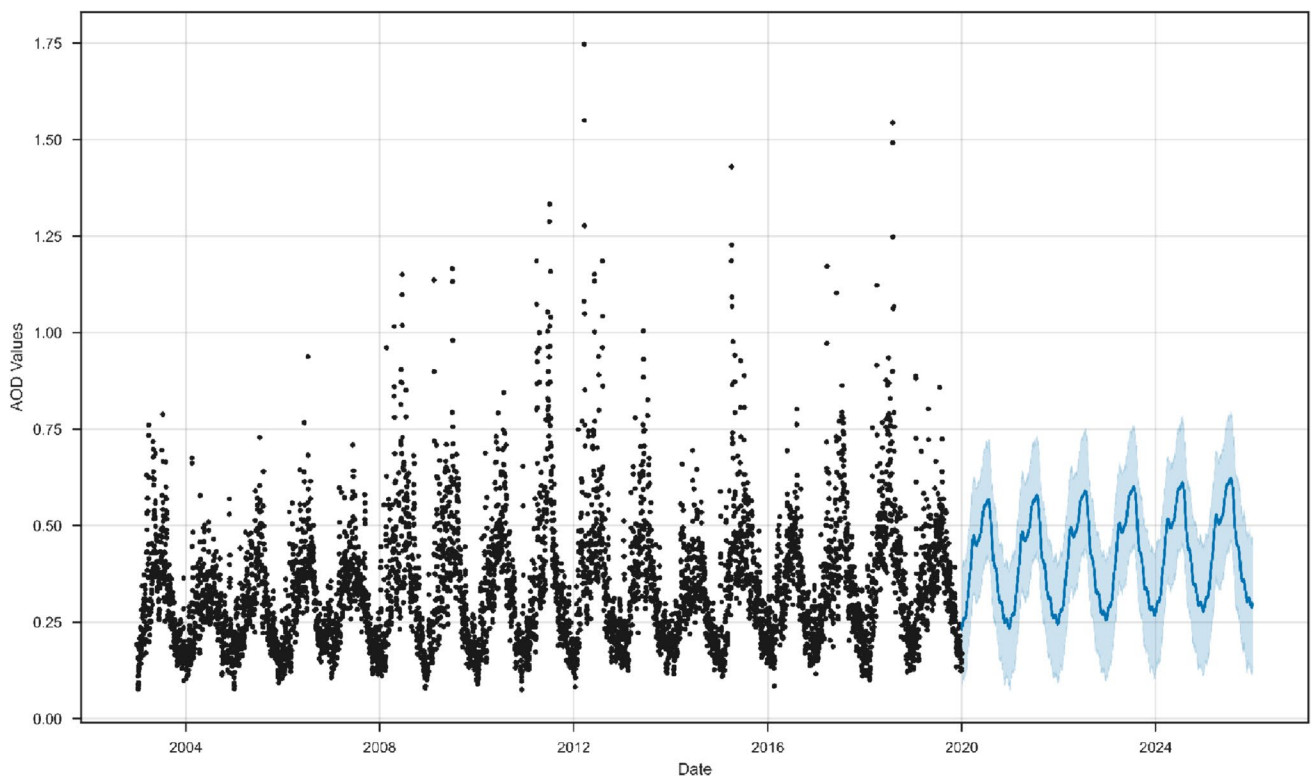
Fig. 16 AOD forecasting for 6 years ahead

5 Conclusion and Policy Implications

AOD, encompassing both absorption and scattering coefficients, plays a pivotal role in atmospheric interactions and impacts climate changes and human well-being. To enhance short- and mid-term daily AOD forecasting precision, we introduced three model categories: basic, hybrid, and combined models. These categories delve into the influence of feature extraction methods on predictive performance. Within the combined category, we explored stacked combining and weighted average architectures to amalgamate superior models from the basic and hybrid groups, namely Bayesian ridge

and ConvLSTM. This ensemble leverages their advanced time series prediction capabilities while addressing individual limitations. Our diverse array of machine and deep learning models underwent rigorous training, validation, and testing using a 17-year daily average AOD time series dataset, sourced from MODIS and MAIAC instruments, spanning the AP. This arid region, marked by expansive deserts and substantial atmospheric dust sources, presented an intricate forecasting challenge. Employing a range of regression error metrics, including RMSE, MAPE, MAE, Pearson's r, and KDE, we comprehensively evaluated the proposed approaches. Significantly, statistical analysis revealed that the weighted average model consistently outperformed other models across prediction horizons from 1 to 7 days.

This achievement holds promising implications, offering actionable insights for governmental bodies and authorities in diverse sectors such as transportation, tourism, and public health. By furnishing accurate daily AOD predictions, this research empowers stakeholders with advance notice of poor air quality events, consequently, they would be able to implement temporary restrictions on outdoor activities, industrial operations, and vehicular traffic to protect public health. The end-user can use the predictions of AOD to make informed decisions and actions, such as issuing alerts and advisories, adjusting emission standards and regulations, allocation resources and funds, and designing research projects and experiments. Furthermore, climate scientists will be able to track and monitor changes in aerosol distribution and concentrations and assess their impact on regional and global climate patterns, leading to developing more efficient climate change mitigation strategies. Lastly, accurate aerosol forecasting can inform policymakers and urban planners about potential air quality challenges associated with urban growth and development. This anticipated insight can lead to sustainable urban planning and designing of more green spaces to mitigate any potential air pollution. For future studies, the long term and spatiotemporal AOD predictions over a wide group of locations and stations can be investigated using high-performance computing techniques. In addition, a larger number of environmental features, such as temperature, humidity, precipitation and air pollutants can be included along the AOD to conduct a more comprehensive analysis.

In comparison with previous studies in the same region and similar climates, we observed seasonal patterns are consistent with those reported in these studies, which lends further credibility to our results (Abuelgasim and Farahat 2019; Bilal et al. 2013). For example, a study conducted over the United Arab Emirate, analyzed the seasonal variation of AOD over 16 years and found that the highest levels of particulate matter (PM) (which are shown to be correlated with AOD) were detected during the summer season, followed by spring, autumn, and finally winter (Abuelgasim and Farahat

2019). Another study conducted over the Arabian Peninsula from 2003 to 2017, analyzed the spatiotemporal variations of aerosols using MODIS-based Merged Dark Target (DT) and Deep Blue (DB) collection of six aerosol products. Furthermore, the study found that AOD was increasing annually at some stations like Kuwait University, Dhabi, and Hamim (Ali and Assiri 2019).

In contrast, there are some limitations in our work. First, our sample data did not include certain features to train the models that could potentially enhance performance results. Thus, future work can investigate the addition of other features such as temperature and humidity to the introduced weighted-average combined model from our work. This addition has the potential to enhance the accuracy of prediction for the ConvLSTM model in the combined technique as it has been found to work well with high density data. Therefore, more weight would be added to this model resulting in an improvement in AOD prediction. Second, we focused on daily AOD prediction, which is more challenging and relevant for practical applications. Exploration of other time scales (e.g., monthly, yearly) can be a possible direction for future work to assess how model performance varies across these scales. Lastly, further validation of model performance may be performed on an independent dataset to assess the practical applicability of the models in real-world scenarios. We plan to conduct such validation in our future work.

Declarations

Conflict of Interest All authors have no conflicts of interest.

References

Abuelgasim A, Farahat A (2019) Effect of dust loadings, meteorological conditions, and local emissions on aerosol mixing and loading variability over highly urbanized semiarid countries: United Arab

Emirates case study. J. Atmos. Solar-Terrestrial Phys. 199:105215. https://doi.org/10.1016/j.jastp.2020.105215

Abuelgasim A, Bilal M, Alfaki IA (2021) Spatiotemporal variations and long term trends analysis of aerosol optical depth over the United Arab Emirates. Remote Sens Appl Soc Environ 23:100532. https://doi.org/10.1016/j.rsase.2021.100532

Abulibdeh A (2021) Analysis of urban heat island characteristics and mitigation strategies for eight arid and semi-arid gulf region cities. Environ Earth Sci 80(7):259. https://doi.org/10.1007/s12665-021-09540-7

Abulibdeh A (2022) Time series analysis of environmental quality in the state of Qatar. Energy Policy 168:113089. https://doi.org/10.1016/J.ENPOL.2022.113089

Abulibdeh A, Zaidan E (2020) Managing the water-energy-food nexus on an integrated geographical scale. Environ Dev. https://doi.org/10.1016/j.envdev.2020.100498

Abulibdeh A, Al-Awadhi T, Al-Barwani M (2019a) Comparative analysis of the driving forces and spatiotemporal patterns of urbanisation in Muscat, Doha, and Dubai. Dev Pract 29(5):606–618. https://doi.org/10.1080/09614524.2019.1598335

Abulibdeh A, Zaidan E, Al-Saidi M (2019) "Development drivers of the water-energy-food nexus in the Gulf Cooperation Council region. Dev Pract. 29(5):582–593. https://doi.org/10.1080/09614524.2019.1602109

Abulibdeh A, Al-Awadhi T, Al Nasiri N, Al-Buloshi A, Abdelghani M (2021) Spatiotemporal mapping of groundwater salinity in Al-Batinah, Oman. Groundw. Sustain. Dev. 12:100551. https://doi.org/10.1016/j.gsd.2021.100551

Ali MA, Assiri M (2019) Analysis of AOD from MODIS-merged DT–DB products over the Arabian Peninsula. Earth Syst Environ 3(3):625–636. https://doi.org/10.1007/s41748-019-00108-x

Ali MA, Assiri M, Dambul R (2017) Seasonal aerosol optical depth (AOD) variability using satellite data and its comparison over Saudi Arabia for the period 2002–2013. Aerosol Air Qual Res 17(5):1267–1280. https://doi.org/10.4209/aaqr.2016.11.0492

Ali MA, Islam MM, Islam MN, Almazroui M (2019) Investigations of MODIS AOD and cloud properties with CERES sensor based net cloud radiative effect and a NOAA HYSPLIT Model over Bangladesh for the period 2001–2016. Atmos. Res. 215:268–283. https://doi.org/10.1016/j.atmosres.2018.09.001

Almazroui M (2019) A comparison study between AOD data from MODIS deep blue collections 51 and 06 and from AERONET over Saudi Arabia. Atmos Res 225:88–95. https://doi.org/10.1016/j.atmosres.2019.03.040

Arden Pope C et al (2011) Lung cancer and cardiovascular disease mortality associated with ambient air pollution and cigarette smoke: shape of the exposure-response relationships. Environ. Health Perspect. 119(11):1616–1621. https://doi.org/10.1289/ehp.1103639

Arfin T, Pillai AM, Mathew N, Tirpude A, Bang R, Mondal P (2023) An overview of atmospheric aerosol and their effects on human health". Environ. Sci. Pollut. Res. 30(60):125347–125369. https://doi.org/10.1007/S11356-023-29652-W

Bandara K, Hyndman RJ, Bergmeir C (2021) MSTL: a seasonal-trend decomposition algorithm for time series with multiple seasonal patterns. Int J Oper Res 1(1):1. https://doi.org/10.1504/ijor.2022.10048281

Bellouin N et al (2020) Bounding global aerosol radiative forcing of climate change. Rev Geophys 58(1):1–45. https://doi.org/10.1029/2019RG000660

Bilal M, Nichol JE, Bleiweiss MP, Dubois D (2013) A Simplified high resolution MODIS aerosol retrieval algorithm (SARA) for use over mixed surfaces. Remote Sens Environ 136:135–145. https://doi.org/10.1016/j.rse.2013.04.014

Bilal M et al (2019) Evaluation of terra-MODIS C6 and C6.1 aerosol products against Beijing, XiangHe, and Xinglong AERONET

sites in China during 2004–2014. Remote Sens. https://doi.org/10.3390/rs11050486

Butt MJ, Assiri ME, Ali MA (2017) Assessment of AOD variability over Saudi Arabia using MODIS deep blue products. Environ Pollut 231:143–153. https://doi.org/10.1016/j.envpol.2017.07.104

Chabane F, Arif A, Moummi N, A. B.-I. (Iranica) J. of Energy, and undefined 2020, "Prediction of solar radiation according to aerosol optical depth," ijee.netF Chabane, A Arif, N Moummi, A BrimaIranian J. Energy Environ. 2020 ijee.net, Accessed: Oct. 05, 2023. [Online]. Available: https://www.ijee.net/article_64627.htmlhttp://www.ijee.net/article_119984.html.

Charfeddine L, Zaidan E, Alban AQ, Bennasr H, Abulibdeh A (2023) Modeling and forecasting electricity consumption amid the COVID-19 pandemic: machine learning vs. nonlinear econometric time series models. Sustain Cities Soc 98:104860. https://doi.org/10.1016/J.SCS.2023.104860

Chen BB et al (2013) Empirical relationship between particulate matter and aerosol optical depth over Northern Tien-Shan, Central Asia. Air Qual Atmos Heal 6(2):385–396. https://doi.org/10.1007/s11869-012-0192-5

Chen X et al (2021) Validation and comparison of high-resolution MAIAC aerosol products over Central Asia. Atmos Environ 251:118273. https://doi.org/10.1016/J.ATMOSENV.2021.118273

Chowdhury S et al (2022) Global health burden of ambient PM2.5 and the contribution of anthropogenic black carbon and organic aerosols. Environ Int 159:107020. https://doi.org/10.1016/J.ENVINT.2021.107020

Cleveland R, Cleveland W, McRae J, I. T.-J. O. Stat, and undefined 1990, "STL: A seasonal-trend decomposition," nniiem.ru, Accessed: Oct. 09, 2023. [Online]. Available: http://www.nniiem.ru/file/news/2016/stl-statistical-model.pdf.

Daoud N, Eltahan M, Elhennawi A (2021) Aerosol optical depth forecast over global dust belt based on LSTM, CNN-LSTM, CONV-LSTM and FFT algorithms," EUROCON 2021—19th IEEE Int. Conf. Smart Technol. Proc. p 186–191. doi: https://doi.org/10.1109/EUROCON52738.2021.9535571.

DiBattista JD et al (2020) Population genomic response to geographic gradients by widespread and endemic fishes of the Arabian Peninsula. Ecol Evol 10(10):4314–4330. https://doi.org/10.1002/ece3.6199

Djuric N, Kansakar L, S. V.-A. Intelligence, and undefined 2016, "Semi-supervised combination of experts for aerosol optical depth estimation," Elsevier, Accessed: Oct. 05, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0004370215001460.

Eibedingil IG, Gill TE, Van Pelt RS, Tong DQ (2021) Comparison of aerosol optical depth from MODIS product collection 6.1 and AERONET in the Western United States". Remote Sens. 13(12):2316. https://doi.org/10.3390/RS13122316

Engelstaedter S, Tegen I, Washington R (2006) North African dust emissions and transport. Earth-Sci Rev 79(1–2):73–100. https://doi.org/10.1016/j.earscirev.2006.06.004

Esmaeil N, Gharagozloo M, Rezaei A, Grunig G (2014) Dust events, pulmonary diseases and immune system. Am. J. Clin. Exp. Immunol. 3(1):20–9

Fabianpedregosa P, et al. (2011) "Scikit-learn: Machine Learning in Python Gaël Varoquaux Bertrand Thirion Vincent Dubourg Alexandre Passos PEDREGOSA, VAROQUAUX, GRAMFORT ET AL. Matthieu Perrot," J. Mach. Learn. Res. 12: 2825–2830. 2011, Accessed: Oct. 05, 2023. [Online]. Available: http://scikit-learn.sourceforge.net.

Fang X, Li R, Kan H, Bottai M, Fang F, Cao Y (2016) Bayesian model averaging method for evaluating associations between air pollution and respiratory mortality: a time-series study. BMJ Open 6(8):e011487. https://doi.org/10.1136/BMJOPEN-2016-011487

Farahat A, El-Askary H, Al-Shaibani A (2015) Study of aerosols' characteristics and dynamics over the Kingdom of Saudi Arabia using a multisensor approach combined with ground observations. Adv Meteorol. https://doi.org/10.1155/2015/247531

Gal Y, Z. G.-A. In neural information, and undefined 2016, "A theoretically grounded application of dropout in recurrent neural networks," proceedings.neurips.ccY Gal, Z GhahramaniAdvances neural Inf. Process. Syst. 2016●proceedings.neurips.cc, Accessed: Oct. 05, 2023. [Online]. Available: https://proceedings.neurips.cc/paper/2016/hash/076a0c97d09cf1a0ec3e19c7f2529f2b-Abstract.html.

Golbraikh A, Shen M, Xiao Z, De Xiao Y, Lee KH, Tropsha A (2003) Rational selection of training and test sets for the development of validated QSAR models. J Comput Aided Mol Des 17(2–4):241–253. https://doi.org/10.1023/A:1025386326946

Han Y, Zhang Q, Li VOK, Lam JCK (2023) "Deep-AIR: A Hybrid CNN-LSTM Framework for Air Quality Modeling in Metropolitan Cities," Mar. 2021, Accessed: Oct. 11, 2023. [Online]. Available: http://arxiv.org/abs/2103.14587.

Hansen J, Sato M, Kharecha P, Von Schuckmann K (2011) Earth's energy imbalance and implications. Atmos Chem Phys 11(24):13421–13449. https://doi.org/10.5194/ACP-11-13421-2011

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780. https://doi.org/10.1162/NECO.1997.9.8.1735

Hoyne C, Mukkavilli SK, Meger D (2023) "Deep learning for Aerosol Forecasting," Oct. 2019, Accessed: Oct. 05, 2023. [Online]. Available: https://arxiv.org/abs/1910.06789v1.

Hsu NC et al (2013) Enhanced Deep Blue aerosol retrieval algorithm: the second generation. J Geophys Res Atmos 118(16):9296–9315. https://doi.org/10.1002/jgrd.50712

Huang R, Liu M, Ding Y (2020) Spatial-temporal distribution of COVID-19 in China and its prediction: a data-driven modeling analysis. J Infect Dev Ctries 14(03):246–253. https://doi.org/10.3855/jidc.12585

Jerrett M et al (2005) A review and evaluation of intraurban air pollution exposure models. J. Expo. Sci. Environ. Epidemiol. 15(2):185–204. https://doi.org/10.1038/sj.jea.7500388

Jing L, Feng X, Jintuo L (2017) 36th C and undefined 2017, "Contrastive research of SVM and BP neural network in AOD prediction," ieeexplore.ieee.orgL Jing, X Feng, L Jintuo, M Rui, L Weiliang, L Yongjun2017 36th Chinese Control Conf. (CCC), 2017 ieeexplore.ieee.org, Accessed: Oct. 05, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8027996/.

Klingmüller K, Pozzer A, Metzger S, Stenchikov GL, Lelieveld J (2016) Aerosol optical depth trend over the Middle East. Atmos Chem Phys 16(8):5063–5073. https://doi.org/10.5194/acp-16-5063-2016

Kumar KR et al (2018) Aerosol optical depth variability over the Arabian peninsula as inferred from satellite measurements. Atmos Environ 187:346–357. https://doi.org/10.1016/j.atmosenv.2018.06.011

Levy RC, Remer LA, Dubovik O (2007) Global aerosol optical properties and application to Moderate Resolution Imaging Spectroradiometer aerosol retrieval over land. J Geophys Res Atmos 112(13):1–15. https://doi.org/10.1029/2006JD007815

Levy RC et al (2013) The Collection 6 MODIS aerosol products over land and ocean. Atmos Meas Tech 6(11):2989–3034. https://doi.org/10.5194/amt-6-2989-2013

Levy RC, Munchak LA, Mattoo S, Patadia F, Remer LA, Holz RE (2015) Towards a long-term global aerosol optical depth record: Applying a consistent aerosol retrieval algorithm to MODIS and VIIRS-observed reflectance. Atmos Meas Tech 8(10):4083–4110. https://doi.org/10.5194/amt-8-4083-2015

Li C, Li J, Dubovik O, Zeng ZC, Yung YL (2020) Impact of aerosol vertical distribution on aerosol optical depth retrieval from passive satellite sensors. Remote Sens. https://doi.org/10.3390/RS12091524

Li J et al (2022) Scattering and absorbing aerosols in the climate system. Nat. Rev. Earth Environ. 3(6):363–379. https://doi.org/10.1038/s43017-022-00296-7

Lyapustin A et al (2011) Multiangle implementation of atmospheric correction (MAIAC): 2. aerosol algorithm. J. Geophys. Res. Atmos. 116(3):3211. https://doi.org/10.1029/2010JD014986

Meo SA, Al-Kheraiji MFA, AlFaraj ZF, Alwehaibi NA, Aldereihim AA (2013) Respiratory and general health complaints in subjects exposed to sandstorm at Riyadh, Saudi Arabia. Pakistan J Med Sci 29(2):642–646. https://doi.org/10.12669/pjms.292.3065

Mhawish A, Banerjee T, Sorek-Hamer M, Lyapustin A, Broday DM, Chatfield R (2019) Comparison and evaluation of MODIS Multi-angle Implementation of Atmospheric Correction (MAIAC) aerosol product over South Asia. Remote Sens. Environ. 224:12–28. https://doi.org/10.1016/j.rse.2019.01.033

Mushtaq Z, Sharma M, Bangotra P, Gautam AS, Gautam S (2022) Atmospheric aerosols: some highlights and highlighters, past to recent years. Aerosol Sci Eng 6(2):135–145. https://doi.org/10.1007/S41810-022-00133-W/FIGURES/4

Nabavi S, Haimberger L, Abbasi R, C. S.-A. research, and undefined 2018, "Prediction of aerosol optical depth in West Asia using deterministic models and machine learning algorithms," Elsevier, Accessed: Oct. 05, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1875963718300260.

Nath P, Roy B, Saha P, Middya AI, Roy S (2022) Hybrid learning model for spatio-temporal forecasting of PM 2.5 using aerosol optical depth. Neural Comput Appl 34(23):21367–21386. https://doi.org/10.1007/s00521-022-07616-4

Nelson M, Hill T, W. R.-J. of forecasting, and undefined 1999, "Time series forecasting using neural networks: Should the data be deseasonalized first?," Wiley Online Libr. Nelson, T Hill, W Remus, M O'ConnorJournal Forecast. 1999 Wiley Online Libr., Accessed: Oct. 05, 2023. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/https://doi.org/10.1002/(SICI)1099-131X(199909)18:5%3C359::AID-FOR746%3E3.0.CO;2-P.

Nichol JE, Bilal M (2016) Validation of MODIS 3 km resolution aerosol optical depth retrievals over Asia. Remote Sens. https://doi.org/10.3390/rs8040328

Oh HJ, Ma Y, Kim J (2020) Human inhalation exposure to aerosol and health effect: aerosol monitoring and modelling regional deposited doses. Int. J. Environ. Res. Public Heal. 17(6):1923. https://doi.org/10.3390/IJERPH17061923

Ollech D (2021) Seasonal adjustment of daily time series. J Time Ser Econom 13(2):235–264. https://doi.org/10.1515/JTSE-2020-0028/MACHINEREADABLECITATION/RIS

Patlakas P, Stathopoulos C, Flocas H, Kalogeri C, Kallos G (2019) Regional climatic features of the Arabian Peninsula. Atmosphere (Basel) 10(4):1–32. https://doi.org/10.3390/ATMOS10040220

Putaud JP et al (2010) A European aerosol phenomenology—3: physical and chemical characteristics of particulate matter from 60 rural, urban, and kerbside sites across Europe. Atmos Environ 44(10):1308–1320. https://doi.org/10.1016/J.ATMOSENV.2009.12.011

Qin W, Wang L, Lin A, Zhang M, Bilal M (2018) Improving the estimation of daily aerosol optical depth and aerosol radiative effect using an optimized artificial neural network. Remote Sens. 10(7):1022. https://doi.org/10.3390/rs10071022

Qin W et al (2021) MODIS high-resolution MAIAC aerosol product: global validation and analysis. Atmos. Environ. 264:118684. https://doi.org/10.1016/j.atmosenv.2021.118684

Ramli N et al (2023) Performance of Bayesian Model Averaging (BMA) for short-term prediction of PM10 concentration in the

Peninsular Malaysia. Atmos. 14(2):311. https://doi.org/10.3390/ATMOS14020311

Ranjan AK, Patra AK, Gorai AK (2021) A review on estimation of particulate matter from satellite-based aerosol optical depth: data, methods, and challenges. Asia-Pacific J Atmos Sci 57(3):679–699. https://doi.org/10.1007/s13143-020-00215-0

Remer LA et al (2008) Global aerosol climatology from the MODIS satellite sensors. J Geophys Res Atmos 113(14):1–18. https://doi.org/10.1029/2007JD009661

Remer LA et al (2020) The dark target algorithm for observing the global aerosol system: Past, present, and future. Remote Sens. https://doi.org/10.3390/RS12182900

Sahu SK, Chen L, Liu S, Ding D, Xing J (2020) The impact of aerosol direct radiative effects on PM2.5-related health risk in Northern Hemisphere during 2013–2017. Chemosphere 254:126832. https://doi.org/10.1016/J.CHEMOSPHERE.2020.126832

Seddon N, Chausson A, Berry P, Girardin CAJ, Smith A, Turner B (2020) Understanding the value and limits of nature-based solutions to climate change and other global challenges. Philos. Trans. R. Soc. B 375:1794. https://doi.org/10.1098/RSTB.2019.0120

Song J, Xia X, Zhang X, Che H, Li X (2018) Weekday AOD smaller than weekend AOD in eastern China on the basis of the MODIS AOD product. Theor Appl Climatol 132(3–4):1019–1027. https://doi.org/10.1007/s00704-017-2142-5

Subba T, Gogoi MM, Pathak B, Bhuyan PK, Babu SS (2020) Recent trend in the global distribution of aerosol direct radiative forcing from satellite measurements. Atmos Sci Lett 21(11):1–12. https://doi.org/10.1002/asl.975

Taneja K, Ahmad S, Ahmad K, Attri SD (2016) Time series analysis of aerosol optical depth over New Delhi using Box-Jenkins ARIMA modeling approach. Atmos Pollut Res 7(4):585–596. https://doi.org/10.1016/j.apr.2016.02.004

Tian H, Chen J (2022) Deep learning with spatial attention-based CONV-LSTM for SOC estimation of lithium-ion batteries. Process. 10(11):2185. https://doi.org/10.3390/PR10112185

Tran H, Kim J, Kim D, Choi M, Choi M (2018) Impact of air pollution on cause-specific mortality in Korea: results from Bayesian model averaging and principle component regression approaches. Sci Total Environ 636:1020–1031. https://doi.org/10.1016/J.SCITOTENV.2018.04.273

Watson-Parris D et al (2019) In situ constraints on the vertical distribution of global aerosol. Atmos Chem Phys 19(18):11765–11790. https://doi.org/10.5194/acp-19-11765-2019

Webel K, Ollech D, Bundesbank D, DG Statistics (2023) An overall seasonality test based on recursive feature elimination in conditional random forests," itise.ugr.es, Accessed: Oct. 09, 2023. [Online]. Available: https://itise.ugr.es/ITISE2018_Webel_Ollech.pdf.

Wei J, Peng Y, Mahmood R, Sun L, Guo J (2019) Intercomparison in spatial distributions and temporal trends derived from multisource satellite aerosol products. Atmos Chem Phys 19(10):7183–7207. https://doi.org/10.5194/acp-19-7183-2019

Wei X, Bin Chang N, Bai K, Gao W (2020) Satellite remote sensing of aerosol optical depth: advances, challenges, and perspectives". Crit. Rev. Environ. Sci. Technol. 50(16):1640–1725. https://doi.org/10.1080/10643389.2019.1665944

Westerlund J, Urbain JP, Bonilla J (2014) Application of air quality combination forecasting to Bogota. Atmos Environ 89:22–28. https://doi.org/10.1016/J.ATMOSENV.2014.02.015

Yeom J, Jeong S, Ha J, K. L.-I. T. on, and undefined 2021, "Estimation of the hourly aerosol optical depth from GOCI geostationary satellite data: deep neural network, machine learning, and physical models," ieeexplore.ieee.orgJM Yeom, S Jeong, JS Ha, KH Lee, CS Lee, S Park. Trans. Geosci. Remote Sensing, 2021●ieeexplore.ieee.org, Accessed: Oct. 05, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9533183/.

Yuan S, Bao F, Zhang X, Li Y (2022) Severe biomass-burning aerosol pollution during the 2019 Amazon wildfire and its direct radiative-forcing impact: a space perspective from MODIS retrievals". Remote Sens. https://doi.org/10.3390/rs14092080

Zaheer K, Saeed S, Tariq S (2023) Prediction of aerosol optical depth over Pakistan using novel hybrid machine learning model. Acta Geophys 71(4):2009–2029. https://doi.org/10.1007/S11600-023-01072-X

Zbizika R, Pakszys P, Zielinski T (2022) Deep neural networks for aerosol optical depth retrieval. Atmos. 13(1):101. https://doi.org/10.3390/ATMOS13010101

Zhang S et al (2016) On the characteristics of aerosol indirect effect based on dynamic regimes in global climate models. Atmos Chem Phys 16(5):2765–2783. https://doi.org/10.5194/acp-16-2765-2016

Zhang S, Wu J, Fan W, Yang Q, Zhao D (2020) Review of aerosol optical depth retrieval using visibility data. Earth-Science Rev 200:102986. https://doi.org/10.1016/J.EARSCIREV.2019.102986

Zhang L et al (2021) Improved 1-km-resolution hourly estimates of aerosol optical depth using conditional generative adversarial networks. Remote Sens. 19:3834. https://doi.org/10.3390/RS13193834

Zhao H et al (2021) Climatological variations in aerosol optical depth and aerosol type identification in Liaoning of Northeast China based on MODIS data from 2002 to 2019. Sci Total Environ 781:146810. https://doi.org/10.1016/j.scitotenv.2021.146810

Zhen Y, Shi G (2023) Evaluation of MACC total aerosol optical depth and its correction model based on the random forest regression. Theor Appl Climatol 152(3–4):1243–1258. https://doi.org/10.1007/S00704-023-04455-8

Zhou H, Yue X, Lei Y, Tian C, Ma Y, Cao Y (2021) Aerosol radiative and climatic effects on ecosystem productivity and evapotranspiration. Curr Opin Environ Sci Heal 19:100218. https://doi.org/10.1016/J.COESH.2020.10.006

Ahmad Qadeib Alban finished his master degree in computer science from Qatar University. His research interests include cybersecurity, cloud computing, machine learning, and deep leaning as well as distributed systems. He is also interested in blockchain and edge data structures. Recently, he has been pursuing a PhD in Computer Science at Université Paris Dauphine-PSL.

Dr. Ammar Abulibdeh is an associate professor of geography and urban planning in the Applied Geography and GIS program, Department of Humanities, Qatar University. He received his PhD degree in Geography and Environmental Management from University of Waterloo in Canada. Dr. Abulibdeh has advanced skills in spatial analysis and in using geographic information system (GIS) and remote sensing (RS) in his teaching and research. Dr. Abulibdeh research focuses on smart urban planning and design, spatiotemporal dynamics of land cover-land use, sustainable built environment, and energy efficiency in urban settings. He published a wide range of articles (more than 50) in top-tier ISI/Scopus journals. Dr. Abulibdeh is the lead principal investigator of research projects funded by Qatar University and Qatar National Research Fund (QNRF).

Dr. Charfeddine Lanouar is a full Professor of Economics and head of the research and policy unit in the Center of Entrepreneurship and Organizational Excellence at the College of Business and Economics, Qatar University. Dr Charfedine's research interests include green finance, environmental sustainability, energy economics, international macroeconomics, and nonlinear econometric models. He has published more than 55 articles on these topics in internationally respected peer-reviewed journals with high-impact factors such as Energy Economics, Energy, Energy Policy, Economic Modelling, Emerging Market

Review, Renewable Energy, Cleaner Energy Production, and renewable and sustainable energy review among many others.

Rawan Abulibdeh is a Ph.D. candidate in the Department of Electrical and Computer Engineering at the University of Toronto, Toronto, Ontario, Canada. Her research interests are on the integration of machine learning technology for the analysis of real-world data with a special focus on the applications of deep neural networks and natural language processing. Rawan completed her M.Sc. in Computer Science at the University of Guelph, Guelph, Ontario, Canada in 2021, where she worked on AI and security. She completed her B.Sc. in Computer Engineering at Qatar University, Doha, Qatar in 2019.

Dr. Abdelgadir Abuelgasim received the master's and Ph.D. degrees from Boston University, Boston, MA, USA, in 1992 and 1996, respectively. He worked for one year as an Assistant Professor of remote sensing and surveying with Sultan Qabous University, Oman. He later joined NASA Goddard Space Flight Center, Maryland, USA, in 1998 as a Visiting Scientist. In 2000, he relocated to Canada where he has been with the Canada Centre for Remote Sensing and the Canadian Space Agency as a Research Scientist. He is currently an Associate Professor of Geography, Rabdan Academy, Abu Dhabi, UAE. His research and teaching interests include the analysis of geospatial data for information products generation, physical and human geography, hydrology, and on the effects of human induced environmental and ecological changes.