Artificial Intelligence and Machine Learning-Based Prediction Models for Environmental Quality Assessment and Their Relevance to Daoist Ecology

Rashmi T V[1], Monika M[2], Rakesh V S[3*], Dr.Palanisamy.R[4], Dr. D Palanikkumar[5], Swathi K[6]

[1]*Assistant Professor, Department of Computer Science and Engineering, East Point College of Engineering and Technology, Bengaluru, Karnataka, India - 560049*
Email: rashmi.tv@eastpoint.ac.in

[2]Assistant Professor, Department of Master of Computer Applications, Cambridge Institute of Technology, K R Puram, Bengaluru, Karnataka, India
Email: monika.mohan.09@gmail.com

[3*]Assistant Professor, Department of Computer Science and Engineering, Cambridge Institute of Technology, K R Puram, Bengaluru, Karnataka, India
*Email: rakesh.tech102@gmail.com

[4]Assistant Professor, No 42, Veltech high-tech Dr rangarajan Dr Sakunthala Engineering college, vel nagar, Avadi, Chennai, India. Email: palanisamy.r@velhightech.com

[5]Professor, Department of Computer Science and Engineering, Dr NGP Institute of Technology, Coimbatore, Tamilnadu, Pin Code: 641048, India.
Email: palanikkumard@gmail.com

[6]Nitte (Deemed to be University) NMAM Institute of Technology (NMAMIT), Nitte, Department of Electrical and ElectronicsEngineering, Mangalore, India
Email: swathi.k@nitte.edu.in

Abstract

The assessment of environmental quality has become a very important and challenging task with the rapid increase of air pollution, water pollution, land use pollution, climate change and loss of biodiversity. Conventional methods for environmental monitoring have some constraints, such as high cost, time consuming, low spatial coverage, and difficulty to make real time predictions. However, the use of large datasets and environmental information by means of artificial intelligence and machine learning has made it possible to assess and predict the environmental quality with high accuracy. These methods include random forest, support vector machine, artificial neural networks, gradient boosting, long and short term memory networks, and also hybrid deep learning models. These models can be used to predict air quality, water quality, soil pollution, climate and ecological risk, and support early warning systems, pollution forecasting, resource management, and also evidence-based environmental management and decision making. However, the use of advanced environmental technologies should be based on an ethical and ecological framework. From this point of view, Daoist ecology can provide a suitable philosophy with nature harmonizing, balance, moderate, natural and non-forced actions. In this paper, the role of prediction models of environmental quality assessment by using artificial intelligence and machine learning will be discussed, and also their relevance to the principles of Daoist ecology will be examined. From the proposed perspective, environmental prediction models should not only focus on technical accuracy but also support sustainable, responsible and nature sensitive decisions.

Keywords: Artificial Intelligence; Machine Learning; Environmental Quality Assessment; Air Quality Prediction; Water Quality Index; Soil Pollution; Daoist Ecology; Sustainable Development; Ecological Ethics; Explainable AI.

1. Introduction

In the current environmental crisis, environmental quality assessment has become one of the major research areas in addition to pollution control and ecological restoration. The serious environmental pollution, greater climate variability, rapid urbanisation, more industrialisation, widespread deforestation, serious water pollution and land degradation as well as loss of biodiversity all have an effect on human health, food security, the eco-system and human development or well-being[1]. The assessment of environmental quality has traditionally relied on monitoring and models, with information collected by periodic field measurements and laboratory tests analyzed by conventional models to provide information on the environmental status[2]. However, these traditional methods of monitoring and assessment have several limitations, including requiring a lot of time and money, the inability to provide information over a wide area, and the delay in alerting people to environmental changes that occur in a short space of time[3]. Recently, there has been a growing body of research that makes use of information collected in real time from a wide range of sources and that employ a very wide range of models, in work on environmental monitoring and management using information technology[4]. Artificial intelligence (AI) and machine learning (ML), in particular, have recently been used in numerous environmental monitoring and management studies. For example, information from a wide variety of sources, including sensors, monitoring stations, weather systems, satellite imagery and GIS databases, can be analyzed using a very wide variety of models, including random forest, support vector machines, decision trees, gradient boosting, artificial neural networks, long and short term memory (LSTM) networks and deep learning models, among many others, in order to make predictions of environmental variables of interest to humans and to assess environmental quality[5]. The great majority of these models can be used for a very wide variety of environmental monitoring and management purposes, including predicting the air quality index, assessing the quality of both surface and and ground waters, identifying areas of soil contamination and forecasting rainfall and temperatures as well as detecting and assessing land degradation and ecological risk. Use of environmental assessment for environmental planning and management, however, should not be confined to technical prediction and assessment, since environmental problems also are problems of human values and behavior and so also need to be addressed through an understanding of the relationship between humans and nature, in other words, through an ecological perspective[6]. Viewed from this perspective, a number of environmental assessment approaches have been developed from ecological perspectives that are derived from religious and philosophical perspectives[7]. Among these, an eco-philosophical perspective derived from Daoist thought is especially worthy of mention, since it is widely practiced in China and viewed as having great practical value for promoting ecological harmony and human development or well-being[8]. From a Daoist perspective, the main goal of environmental assessment is not simply to measure the degrees of pollution and ecological destruction but also to use assessment information in order to discover imbalances in the environment and to bring about ecological restoration. In ecological terms, assessment of environmental quality using AI and ML models, therefore, should not be confined to simple technical prediction and assessment, and should instead be used to promote ecological harmony and human development or well-being through assessment that is followed by ecological restoration[9]. Furthermore, from a number of perspectives, including an eco-philosophical perspective derived from Daoist thought, the use of AI and ML models for environmental assessment should not involve simply using these models in order to monitor the environment from a outside perspective in a detached or impartial manner but also should involve using them in a manner that is in harmony with nature, in other words, in a sustainable manner[10]. The eco-philosophical perspective derived from Daoist thought viewed from this perspective stresses the importance of harmony between humans and nature, and so views the environmental crisis from the perspective of imbalances between humans and nature that have arisen as a result of human ecological values, activities and behavior that are not in harmony with nature. From this viewpoint, the use of assessment information derived from AI and ML models for environmental quality assessment should not be used in a manner that is in opposition to nature, in other words, in a manner that is not in harmony with or in opposition to nature, but instead should be used in a sustainable manner that is in harmony with or in accord with nature[11]. This study aims to provide information on a number of prediction models that are made use of in environmental quality assessment using AI and ML and to examine the relevance of these models from a perspective derived from an eco-philosophical perspective, i.e. a Daoist ecological perspective, with a view to identifying approaches to the use of such models that are in harmony with nature and which can be used in order to promote harmony between humans and nature and to achieve ecological restoration and sustainability.

2. Literature Review

AI and ML in Environmental Quality Assessment

Artificial intelligence (AI) and machine learning (ML) are increasingly used in the environmental assessment and monitoring. Large amounts of complex data are generated and analyzed in real time to enable the assessment of environmental quality. In many cases, supervised learning is applied, using labeled data such as concentrations of pollutants, water quality classes, soil contamination categories, or even a climate change risk score[12]. Various algorithms are used for this purpose, such as random forest, support vector machine, decision tree, gradient boosting, and artificial neural networks (ANN). Unsupervised learning, on the other hand, is applied for clustering of polluted areas, detection of anomalies in environmental data, and identification of hidden patterns in large datasets from environmental sensors and satellite images[13]. Deep learning models are powerful tools for time-series prediction, image classification, and other complex assessments. In addition, hybrid models that combine two or more different approaches are increasingly used to improve accuracy and robustness[14]. Explainable AI is another important aspect, since many environmental models are used for decision support and their results must be interpretable by the user. This means that the feature used for the assessment must be explained, and the relationships between the input data and the assessment results must be transparent. Finally, environmental AI should also be sustainable, i.e., energy efficient, and not create more environmental problems than they solve.

Air Quality Prediction

Air quality prediction is one of the most commonly studied topics by the use of artificial intelligence and machine learning models for the assessment of environmental quality. As it is common for other elements, the air quality index prediction is based on a set of variables related to air pollution as well as to a set of variables which are related to the weather as well as to human activities that affect to the air quality. The common set of variables for the prediction of air quality index are: PM2.5, PM10, NO2, SO2, CO, O3, among others[15]. Using a large set of environmental variables, supervised learning models can be used for the air quality index prediction in both a short-term or long-term framework. Using large set of variables, several types of learning models have been tested for this type of prediction. For instance, there are models based on decision trees (such as random forest and gradient boosting), regression models with kernel functions (such as support vector machines), artificial neural networks, as well as other models that can be hybridized. Deep learning models are being increasingly applied, especially for time-series data, with good results in the analysis of the information containing in the air pollution data which have a temporal and/or spatial component[16]. For example, in the assessment of air quality, models of long short-term memory networks (LSTM), gated recurrent units (GRU), as well as different architectures that are based on convolutional neural networks (CNN), and the use of CNN-LSTM in hybrid models, have been reported. Furthermore, there are recent reviews where the use of AI models are analyzed for the air pollution monitoring and prediction, where the best performance has been obtained with the use of machine learning models, while the use of deep learning models for the air quality assessment provides better results to model complex spatiotemporal patterns of air pollution.

Water Quality Prediction

Water quality prediction models are also of considerable value. By using a variety of environmental and hydrological parameters, including historical data from water quality monitoring stations, readings from online water quality sensors, and real-time data from a variety of sources, such models can determine water quality index scores (WQIs) for a wide range of parameters such as pH, turbidity, dissolved oxygen (DO), biochemical oxygen demand (BOD), chemical oxygen demand (COD), electrical conductivity (EC), temperature, and a variety of inorganic and organic pollutants, including nutrients (e.g., nitrate, phosphate) and inorganic and organic compounds that cause water to have high salinity or total dissolved solids (TDS) as well as to be contaminated with heavy metals and other hazardous substances[17]. Models are also developed for a variety of different water bodies including rivers, lakes, reservoirs, groundwater, and aquaculture water[18]. A wide variety of supervised learning models, including Random Forest (RF), Support Vector Machines (SVM), Support Vector Regression (SVR), and Gradient Boosting (GB), as well as a variety of types of artificial neural networks (ANNs) and a wide array of deep learning models, are used for water quality pollution prediction as well as for water quality classification into a number of different categories[19]. The non-linear relationships that exist between the parameters that are measured and determined for water can be detected by deep learning models, and, using past water quality data, such models can also be used to make a number of predictions about future water quality. However, as with all types of monitoring and assessment, there are a number of limitations and challenges associated with the use of models for water quality assessment, including the challenge of dealing with missing data as well as the challenge of the uneven distribution of water quality monitoring stations, the challenge of developing models that can be used to make good quality predictions using a limited amount of labeled data, as well as the challenges of model transferability, or the challenges of using models that were developed and calibrated using data from a particular location to make predictions at a very different location, as well as a number of other limitations.

Soil Pollution and Land Quality Assessment

To begin with, one has to remember that soil contains natural and anthropogenic components. It is constantly undergoing change, often unintentional, as a result of natural factors and human activities. In agricultural and other disciplines, land is utilized in a variety of different ways and is affected by a variety of human activities (industrial, waste disposal, mining, etc.). Many factors influence the quality of the top soil and subsoil, including the amount of irrigation water used, the type of waste disposed of, and the chemicals released into the atmosphere[20]. A multitude of factors influence the degree of soil pollution and the rates of processes taking place in the soil. Furthermore, in order to assess the degree of pollution by heavy metals, there are a number of other properties that need to be taken into account. For this reason, a geospatial approach to the assessment of pollution and the creation of maps showing the degree of risk of pollution to soil is often based on machine learning models. In order to assess the quality of the land and to forecast the bioavailability of individual heavy metals, a number of different models can be used including a variety of types of machine learning, including Random Forest, Support Vector Machine models, Feed Forward Neural Network models, Gradient Boosting models, as well as Graph-based models.

Remote Sensing, IoT, and Big Data

As monitoring equipment such as smart monitoring stations, UAVs (unmanned aerial vehicles) and sensors from the Internet of Things (IoT) have become increasingly common, it has become possible to collect a large volume of data from the environment. In addition to this, a variety of technologies are being used to process this data, including geographic information systems (GIS). Remote sensing can provide data on a variety of indicators, including information on the cover of vegetation, land surface temperature, and urban expansion over time. Indicators such as NDVI (Normalised Difference Vegetation Index), land surface temperature and a range of others can be used to assess a variety of environmental indicators, including signs of drought and ecosystem stress, as well as factors that may contribute to the creation of an urban heat island and other climate related change. Data collected by UAVs can be of high resolution and cover a large area in a short space of time, allowing for the creation of maps to show information on a variety of environments, including forests, rivers, farms and wetlands, as well as to help identify areas that may be at risk of natural disasters. The large amount of data that can be generated by IoT sensors can be used in real time to monitor a range of environmental indicators including air pollution, water quality, temperature, humidity, noise and soil condition. By combining data from remote sensing, the IoT, GIS and machine learning, it is possible to create a variety of models that can be used for tasks such as classifying and predicting information, detecting anomalies and providing early warning for environmental events, as well as to support a range of environmental monitoring and management, including flood risk mapping, drought prediction, biodiversity monitoring and assessment of degradation of ecosystems, and smart environmental governance.

Explainable and Green AI

In environmental quality assessment by means of AI, it is important to apply models that are explainable as well as green. Predictions made by models of environmental quality can have a serious impact on public health and the environment, in land use planning and management of natural resources and in preventing pollution and in measures to adapt to climate change. Therefore, the results of models of environmental quality should be clear to scientists, policy makers and other stakeholders. By using explainable AI methods such as SHAP values, LIME, feature importance, partial dependence plots and other methods of interpretable machine learning, the variables that most affect the predictions by a model of environmental quality can be determined. In order to apply AI in environmental assessment in a responsible manner, it is also important that the application of AI does not lead to an increase in environmental pollution. Therefore, models of environmental quality should be energy efficient, their computing requirements should be kept to a minimum and their features should be selected carefully. Furthermore, the data used by models of environmental quality can be collected by means of sensors that have a low power requirement. By using edge computing instead of cloudy computing, the computing requirements for transmission of data can be kept to a minimum. Finally, there is a need for carbon aware computing in order to avoid as much as possible unnecessary computing. In short, the use of AI in the environmental quality assessment should be accurate, interpretable, efficient and sustainable.

3. Daoist Ecology: Theoretical Foundation

Daoist ecology is a theory that studies environmental balance, human responsibility and harmony with nature. To a great extent, the ecological theory of Daoism stems from its core tenets. First and foremost, it posits that nature is by its very nature followable—i.e. it has its own laws, its own rhythm and its own internal patterns which, if left to themselves, are self-sustaining and in a state of constant regeneration. From an ecological perspective, the Dao, or Natural Order, is something to which human activity must learn to conform rather than attempt to disrupt it by some form of control, exploitation or interference. In terms of ecological study, the concept of ziran or naturalness/Self-so-being is central to this school of thought. Here, it is posited that all ecosystems have their own inherent momentum, their own balance and in themselves the seeds of their own regeneration, provided that no undue external interference is introduced. In short, as ecosystems, all aspects of the natural world—rivers, forests, soils, climate, plants and animals, human populations and so on—are seen as being part and parcel of some greater whole or even as a holistic entity in themselves. As such, all of these are inter-related and it is only by coming to some deep understanding of these inter-relationships that true wisdom may be obtained. An ecologically-oriented view of the world of human activity, accordingly, would seek to bring about wu wei or non-forced activity on the part of human beings. In practical terms, this does not necessarily mean a lack of action on the part of human beings. Rather, it involves all of those activities which are undertaken in a fashion that is in itself as natural and unforced as possible. The crux of the matter here is that all human activity has the potential to cause damage to the natural world. It follows that all such activity must be undertaken in a fashion that is itself as non-damaging as possible and which works in a manner that is in itself as supportive as possible of the natural world's own processes of ecological regeneration and recycling. As for the issue of human harmony with nature, here it is asserted that for human beings there is no conflict with nature and that, as part of the broader ecological system or systems, they must learn to live in harmony with nature. As it stands, the worst forms of destructive environmental behavior are a direct result of industrialization and of excessive technological advance which, when left to their own devices, place ever greater pressures upon the natural world and upon its finite resources. Harmony with nature, in short, is the key to all human ecological studies and it is something which can only be brought about by a return to simple methods of living in a fashion of greatest possible moderation in all respects. The study of environmental issues by means of Daoist ecological theory can, in turn, form the basis of a sound and workable environmental ethic which will promote the best possible of non-exploitative relationships with nature and which will, in addition, provide for the due protection of the totality of all natural balances which currently exist. It is this ecological perspective which forms the theoretical basis for a critical study of the potential of the use of artificial intelligence and machine learning for the assessment of the environmental quality. In summary, in order to effect a successful transition to a more sustainable mode of living it is necessary for human beings to adopt a more sustainable form of development and that, in turn, will only come about when there is given a complete shift in the focus of current interest from seeking out that which is in some fashion novel and which will in itself bring about some form of improvement to affairs as they currently are to that which will, in a most sustainable fashion, support human ecological studies in the best possible of ways. It is in this sense that the study of environmental issues by means of ecological theory of Daoism can be of greatest assistance. This is because, when viewed from a Chinese ecological perspective, all natural systems—whether they be in the form of trees, of rivers, of earth, of climate, of animals, of human populations and so on—are not viewed as separate, fragmented and isolated entities. Rather, they are studied as inter-connected parts of some greater whole or as whole ecological systems in themselves. It is from this perspective that the environmental quality assessment by means of artificial intelligence and machine learning can be interpreted in a most sustainable fashion so as to result in decisions that are to be the most beneficial to man and nature in the long run.

4. Proposed Conceptual Framework

This conceptual framework integrates environmental quality assessment using AI and ML with a Daoist ecology perspective on environmental issues in three layers. The first layer is the environmental data layer or input layer where all types of data regarding environmental quality are collected and analyzed. This include air quality data, water quality data, data regarding soil quality and land quality as well as climate, remote sensing, IoT, and other data. It contains physical, chemical, biological, spatial, or human-related data and indicates various aspects of the environmental system. Air quality data may include PM2.5, PM10, NO_2, SO_2, CO, O_3, temperature, humidity, wind speed, etc. while water quality data include pH, turbidity, dissolved oxygen, BOD, COD, etc. with parameters such as nitrate, phosphate, electrical conductivity, and various heavy metals. Similar parameters are collected for assessing soil and land quality. The second layer is the AI and ML prediction layer where the analyzed data are cleaned, structured, and analyzed by various models to predict different indicators regarding environmental quality. During this process, data preprocessing, missing value handling, outlier detection, feature selection, model training and testing, prediction by AI and ML models, explainable AI analysis, and model validation are conducted. Model types include random forest, support vector machine, decision tree, XGBoost, LightGBM, artificial neural network, LSTM, GRU, CNN-LSTM, autoencoder, graph neural network, and their hybrid models for assessing Air Quality Index, Water Quality Index, soil pollution, and ecological risk as well as climate-related and overall environmental degradation. Model performance is measured by accuracy, precision, recall, F1-score, RMSE, MAE, and R^2, etc. depending on the type of assessment. In addition, explainable AI indicates the most important input data and, therefore, makes environmental decision making transparent.

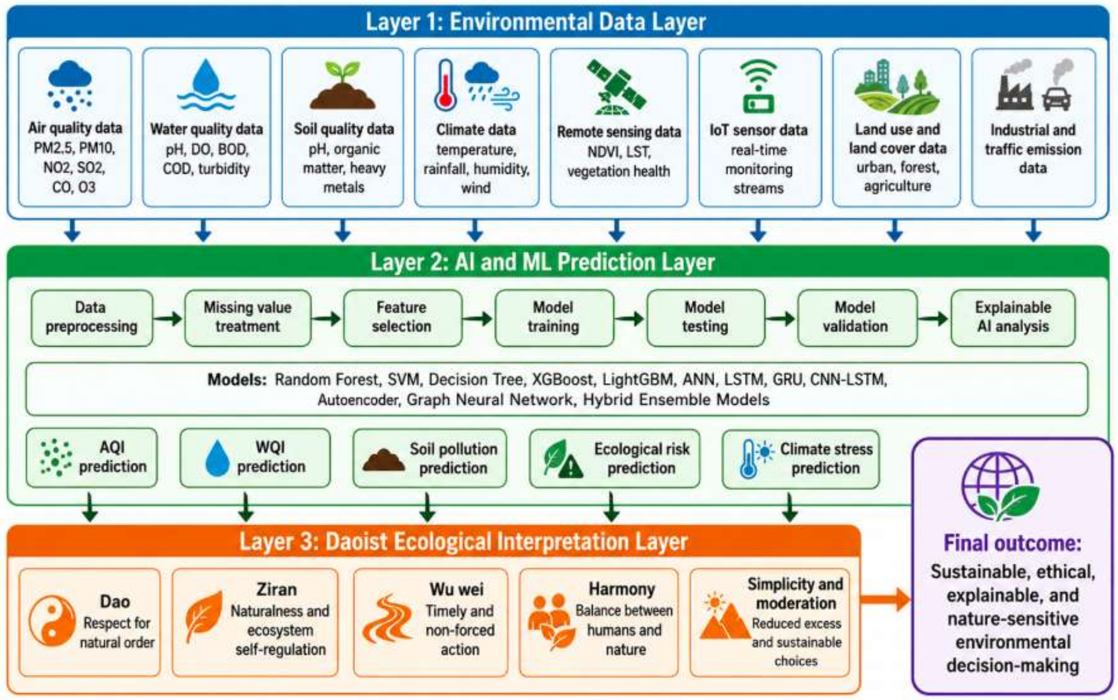

Figure 1. Proposed conceptual framework

Figure 1 shows the framework of three layers: environmental data, AI and ML prediction models and Daoist ecological interpretation. Environmental data layer (soil, water, air, climate, remote sensing, IoT, land use & emission data). The AI and ML layer pre-processes, selects relevant features, trains models, tests these models and finally uses explainable AI techniques on the environmental data layer. The prediction results that can be generated by AI and ML models include: the air and water quality index (AQI, WQI), assessment of the degree of pollution of soils and land ecological systems, assessment of the potential ecological risks, and climate stress assessment. The Daoist ecology layer (right side) interprets the results by referring to Dao, ziran, wu wei, environmental harmony and simplicity by avoiding excess.

The third layer is the Daoist ecology interpretation layer. Here, assessments and early warning systems generated by models are interpreted from an ecological perspective to identify disharmony between human activities and the natural environment and to indicate possible causes of pollution. Results show that over-industrialization indicates excessive activity and violate the principles of moderation and ecological restraint. In addition, the framework supports implementation of early warning systems in a careful, timely, and non-destructive manner, thus corresponding to the concept of wu wei. Green AI supports simplicity by reducing excessive computational requirements and technological excess. In addition, Explainable AI supports responsible decision making by making complex environmental information clear and accountable. Therefore, this framework enables accurate, ecologically wise environmental assessment that is technically sound and ethically responsible, and supports environmental assessment and monitoring in line with the Daoist principles of harmony, balance, naturalness, and moderation.

5. Methodology

This paper aims to design a conceptual framework for implementing environmental quality assessment using prediction models based on Artificial Intelligence (AI) and Machine Learning (ML), that are relevant to the ecological perspective of Daoism. In order to achieve the research objectives of the paper, a conceptual review-based methodology was employed. Recent articles, review papers, technical reports and philosophical texts were collected and analyzed in order to identify the main AI and ML approaches for the environmental quality assessment as well as the most commonly used models and their applications.

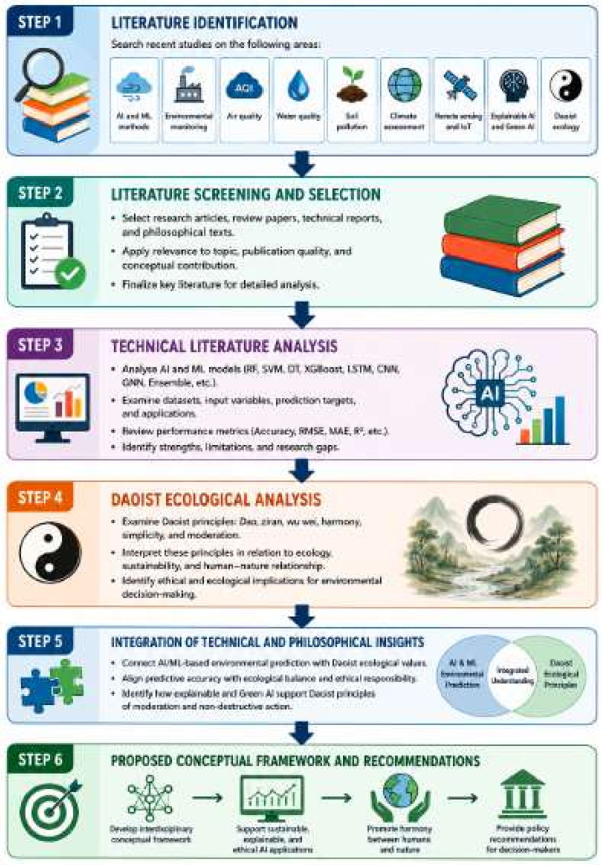

Figure 2: Methodological Flow of the Conceptual Review-Based Study

In addition, the types of datasets that are used for the environmental quality prediction, such as Air Quality Index, Water Quality Index, as well as datasets for soil pollution, remote sensing-based vegetation data, climate data and pollution monitoring data were identified and analyzed. Figure 2 shows the methodological approach of the conceptual review-based study. Start with literature identification on AI and/or ML and environmental monitoring and Daoist ecology. Literature of relevance to the study such as research articles, review papers, technical reports and philosophical texts on the topics of AI and/or ML and/or environmental monitoring and/or Daoist ecology were identified and selected for screening. The selected literature has been analyzed using a technical perspective as well as a Daoist ecological perspective. The outcomes of both viewpoints streams were then integrated to develop a conceptual framework as well as guidelines for the quality assessment of the sustainable environmental system.

The main aspects of the studies that were analyzed, included the input variables and prediction targets, the models' performance, the assessment metrics that were employed in order to evaluate the models, the potential applications of the models, the strengths and limitations of each study. In addition to the assessment of the studies that were analyzed in the technical part of the paper, the ecological perspective of Daoism was also analyzed in order to identify the main concepts of the philosophy, including Dao, ziran, wu wei, as well as the harmony between humans and nature, simplicity and moderation. The study's objective is to develop an integrated framework that links the data that are collected from the environmental sources with the models' prediction outcomes to the ecological principles of Daoism for the environmental quality assessment. The study aims at developing an integrated approach that combines the aspects of environmental assessment and sustainability, using technology, as well as ecological wisdom for responsible human actions and decisions.

6. Model Evaluation Metrics

Model evaluation metrics or measures are an essential component of any environmental quality prediction models, which are developed by using artificial intelligence and/or machine learning techniques. The choice of appropriate model evaluation or assessment measures depends on the nature of the prediction task(s) that a model is designed to perform, i.e., regression, classification, environmental index (Air Quality Index (AQI) and/or Water Quality Index (WQI)) prediction, spatial or geographic mapping, and/or temporal or time-series forecasting. For the regression-based environmental quality prediction models, several assessment measures, including Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R^2 (Coefficient of Determination) score, are commonly used. MAE provides an average difference between the predicted and observed values, and thus simple to interpret. MSE and/or RMSE assign a greater weight to larger errors, and thus are desired metrics when larger errors (in prediction) are considered more serious. RMSE expresses the prediction errors in the same unit of measurement as that of original variables, and thus are more interpretable for the practitioners or users who are working with environmental parameters, such as concentrations of pollutants, temperatures, rainfall, or water quality parameters. R^2 score measures the fraction of the variance in the observed data that is predictable from the model. For the classification-based assessment of environmental quality, several other metrics, including accuracy, precision, recall, F1-score, Receiver Operating Characteristic-Area Under the Curve (ROC-AUC), and confusion matrix, are typically used. Accuracy measures the proportion of correctly classified instances (predictions) by a model. Precision measures the proportion of predicted positive instances that are correctly classified. Recall also measures the proportion of correctly classified instances, but from the perspective of actually positive or at-risk instances, and thus is particularly important for purposes of pollution warning systems and/or for detecting and predicting ecological risks to environment and human health. F1-score is another commonly used metric, which provides a balanced measure of precision and recall, especially when classes in a dataset are imbalanced. ROC-AUC measures the model's ability to distinguish between different classes of risks or pollution levels. A confusion matrix reports the number of correct and incorrect classifications by a model for each class or category of pollution, including low, moderate, high, and severe pollution risks. In addition to the above-mentioned metrics, for the environmental quality index prediction models, measures of accuracy for Air Quality Index (AQI) and/or Water Quality Index (WQI) predictions, pollution risk class accuracy, spatial prediction accuracy, and/or temporal or time-series forecasting accuracy are typically considered important for environmental quality prediction models. The spatial prediction accuracy is also important for remote sensing and/or Geographic Information System (GIS)-based pollution mapping models and/or for modeling and/or mapping of land quality and/or other environmental resources. The temporal or time-series forecasting accuracy provides an assessment of a model's ability to forecast and/or predict time-dependent environmental variables and/or parameters at different time scales, including hourly, daily, seasonal, and/or long-term or permanent time-series data. Together, these assessment measures help to assess whether a model is not only technically sound and reliable to use, but also environmentally meaningful, and useful for making appropriate environmental management and/or decision-making.

7. Results and Discussion

The expected findings from the conceptual review indicate that artificial intelligence and machine learning models can significantly improve environmental quality assessment by handling complex, non-linear, multi-source, and high-dimensional environmental data. Such data could be from environmental monitoring, which in itself provides robust measurements of the environment. However, using these in prediction models enables early warning, trends and spatial assessments as well as real time decision support that would not be possible with traditional monitoring alone. Figures 3 and 4 show an example of a prediction of Air and Water Quality Index for a set of environmental data. The agreement between observed and predicted values for this simulated data set indicates that such a model would be able to pick up on seasonal cycles as well as pollution events and longer term environmental change. It also indicates the level of uncertainty in the predictions, which is critical for any environmental assessment and planning being used for decision making. Figure 4 shows the performance of a range of AI and ML models (Support Vector Machine, Random Forest, XGBoost, LightGBM, LSTM, CNN-LSTM and a number of hybrid models) on a range of different data sets. In general, tree-based models (such as Random Forest, XGBoost and LightGBM) performed best on the environmental data sets with primarily tabular data, due to their ability to handle a mixture of data types as well as non-linear relationships between variables and interactions between features. However, for time-series data (i.e. data where the past is used to predict future values, such as for air pollution, rainfall, temperature and water quality) deep learning models (such as LSTM, GRU and CNN-LSTM) performed best. The hybrid models also performed very well as they are able to leverage the strengths of a number of different models, thus lessening the impact of individual model weaknesses. Figure 5 provides an example of a spatial pollution risk map that could be created using remote sensing, GIS, IoT and big data, highlighting areas of high risk due to industrial and traffic pollution that would not be picked up by ground monitoring alone. Figure 6 integrates these predictions with Daoist ecological assessments, utilizing a Harmony Index and an explainable AI (XAI) -based driver contribution analysis to show how pollution, heavy metals, traffic and poor quality of water can all decrease ecological harmony whilst natural dilution factors as well as a balanced and healthy environment can increase it. Such assessments of environmental risk are being increasingly utilized for public health, biodiversity, land use, water resources and policy purposes. Importantly, the results of such assessments need to be interpreted in a transparent manner so that others can understand the underlying drivers.

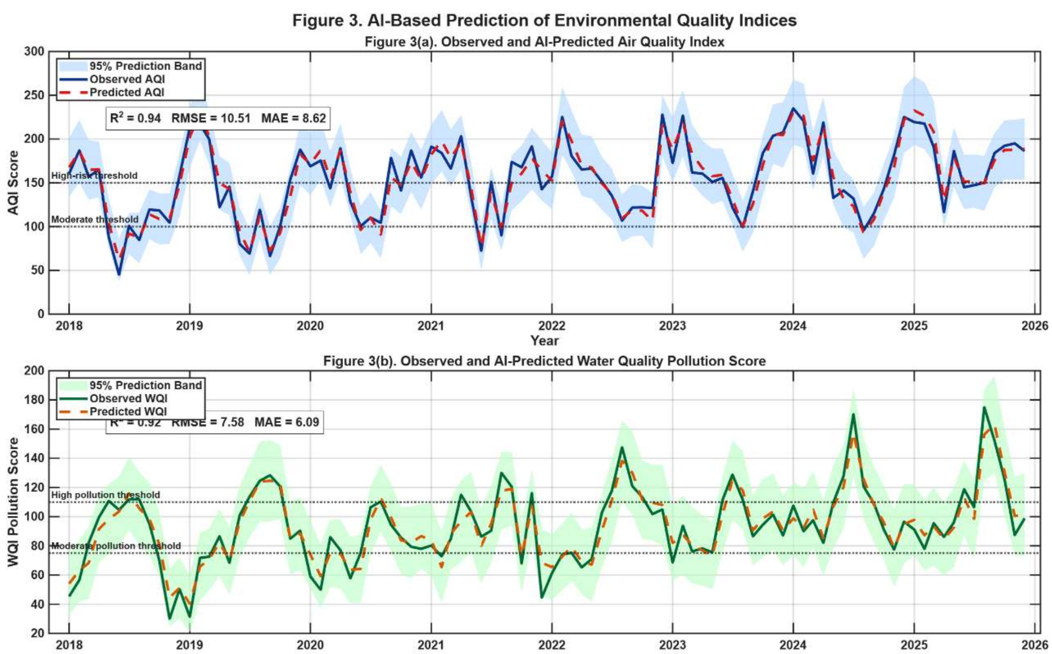

Figure 3: AI-Based Prediction of AQI and WQI

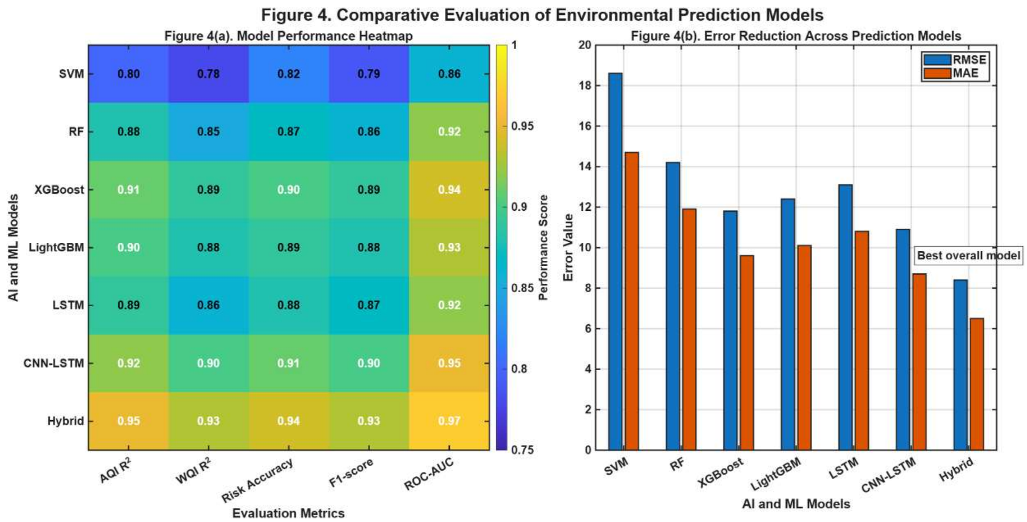

Figure 4: Comparative Evaluation of Environmental Prediction Models

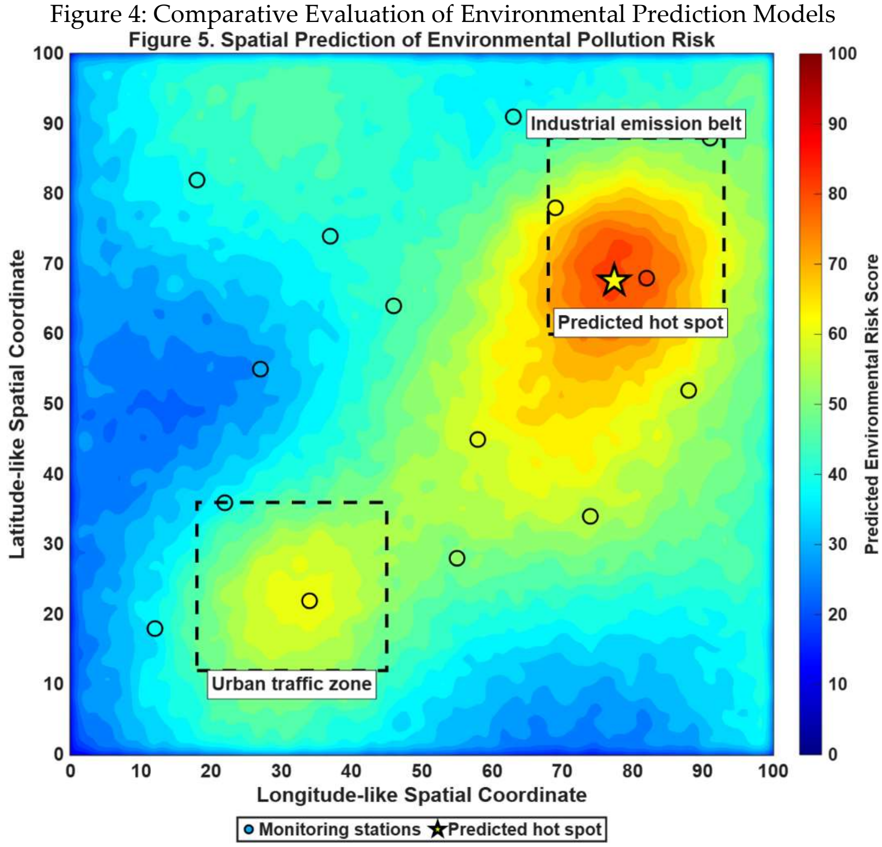

Figure 5: Spatial Prediction of Environmental Pollution Risk

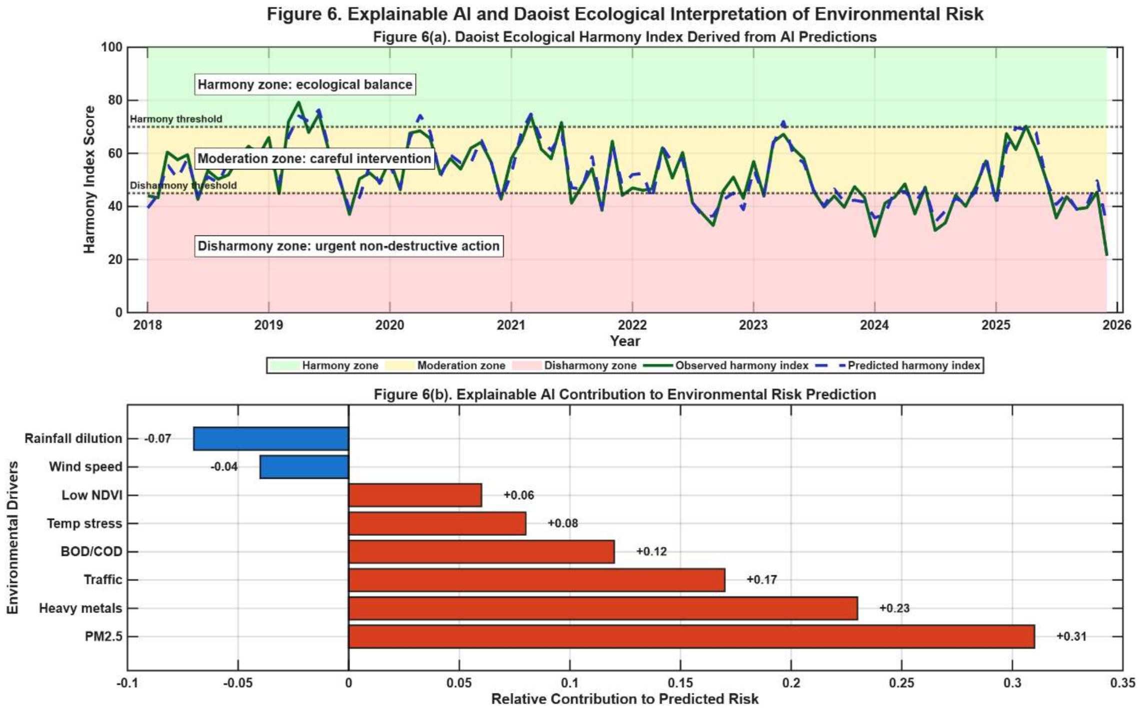

Figure 6: Explainable AI and Daoist Ecological Harmony Interpretation

This is particularly critical as nature is being constantly manipulated by humans and AI systems are increasingly being used to monitor and control the environment. From a Daoist ecological perspective, such monitoring and controlling of nature must be used in a manner that does not enable further technological domination of nature, rather it should be used to increase human awareness of environmental imbalances and enable timely intervention in order to promote moderation and reduce excess in human interaction with the environment in order to ensure that any human actions are non-destructive to the environment. Thus, the use of AI for environmental assessment is only meaningful when it is utilized in a manner that increases ecological harmony, or balance, and promotes wise stewardship of the environment.

8. Conclusion

Artificial intelligence and machine learning-based prediction models offer powerful support for environmental quality assessment by analysing complex environmental datasets, detecting pollution patterns, forecasting future risks, and improving decision-making. However, technical accuracy of the environmental AI systems is not sufficient for the sustainable environmental management. Therefore, the use of the environmental AI systems should be accompanied and guided by the ecological wisdom, ethics and nature. One of such philosophies is Daoist ecology that describes the ideal relationship between humans and nature by the concepts of harmony, balance, naturalness, moderation, simplicity and non-forced action. Thus, the future environmental AI systems should be designed as accurate, automated, explainable, energy efficient and ecologically sensitive models. The balanced combination of the technical environmental prediction models and the environmental AI system with the ideas of Daoist ecology can support the responsible environmental governance and promote the sustainable human activities in natural environment.

References

1. S. Chadalavada, O. Faust, M. Salvi, S. Seoni, N. Raj, U. Raghavendra, A. Gudigar, P. D. Barua, F. Molinari, and R. Acharya, "Application of artificial intelligence in air pollution monitoring and forecasting: A systematic review," *Environmental Modelling & Software*, 2024, doi: 10.1016/j.envsoft.2024.106312.

2. Z. Zhang, S. Zhang, C. Chen, and J. Yuan, "A systematic survey of air quality prediction based on deep learning," *Alexandria Engineering Journal*, vol. 93, pp. 128–141, 2024, doi: 10.1016/j.aej.2024.03.031.
3. I. Gryech, C. Assad, M. Ghogho, and A. Kobbane, "Applications of machine learning and IoT for outdoor air pollution monitoring and prediction: A systematic literature review," *arXiv preprint arXiv:2401.01788*, 2024, doi: 10.48550/arXiv.2401.01788.
4. W. Zhi, A. P. Appling, H. E. Golden, J. Podgorski, and L. Li, "Deep learning for water quality," *Nature Water*, vol. 2, pp. 228–241, 2024, doi: 10.1038/s44221-024-00202-z.
5. W. Li, Y. Zhao, Y. Zhu, Z. Dong, F. Wang, and F. Huang, "Research progress in water quality prediction based on deep learning technology: A review," *Environmental Science and Pollution Research*, vol. 31, pp. 26415–26431, 2024, doi: 10.1007/s11356-024-33058-7.
6. Y. Zhang, M. Lei, K. Li, and T. Ju, "Spatial prediction of soil contamination based on machine learning: A review," *Frontiers of Environmental Science & Engineering*, vol. 17, no. 93, 2023, doi: 10.1007/s11783-023-1693-1.
7. S. Zhong *et al.*, "Machine learning: New ideas and tools in environmental science and engineering," *Environmental Science & Technology*, vol. 55, no. 19, pp. 12741–12754, 2021.
8. M. Reichstein *et al.*, "Deep learning and process understanding for data-driven Earth system science," *Nature*, vol. 566, pp. 195–204, 2019, doi: 10.1038/s41586-019-0912-1.
9. C. Varadharajan *et al.*, "Can machine learning accelerate process understanding and decision-relevant predictions of river water quality?" *Hydrological Processes*, 2022, doi: 10.1002/hyp.14565.
10. V. Bolón-Canedo, L. Morán-Fernández, B. Cancela, and A. Alonso-Betanzos, "A review of green artificial intelligence: Towards a more sustainable future," *Neurocomputing*, 2024, doi: 10.1016/j.neucom.2024.128096.
11. S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proc. Advances in Neural Information Processing Systems*, 2017, pp. 4765–4774.
12. M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you? Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 1135–1144, doi: 10.1145/2939672.2939778.
13. L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
14. T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.
15. S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
16. K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *Proc. EMNLP*, 2014, pp. 1724–1734.
17. T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. International Conference on Learning Representations*, 2017.
18. N. J. Girardot, J. Miller, and X. Liu, Eds., *Daoism and Ecology: Ways Within a Cosmic Landscape*. Cambridge, MA, USA: Center for the Study of World Religions, Harvard Divinity School, 2001.
19. E. S. Nelson, *Daoism and Environmental Philosophy: Nourishing Life*. London, U.K.: Routledge, 2020, doi: 10.4324/9780429399145.
20. J. Miller, *China's Green Religion: Daoism and the Quest for a Sustainable Future*. New York, NY, USA: Columbia University Press, 2017.