Machine Learning Applications in Climate Science:Novel Approaches to Prediction and Monitoring

Prakhar Rastogi[1], *Dr.Pawan Saxena*[2]
[1]Dept of Amity Institute of Information Technology, Amity University, Lucknow, India
[2]Amity University, Lucknow

***Abstract*—Climate Concurrent climate change issues require new strategies in environmental prediction and live monitoring. Conventional General Circulation Models (GCMs) and Numerical Weather Prediction NWP) models provide basic insights but are hampered by computing constraints and have difficulties with the nonlinear nature of environmental information.This study assesses how climate forecasting abilities for temperature, precipitation, and humidity readings can be remapped by intricate machine learning structures in the guise of Long Short-Term Memory (LSTM) networks, Random Forest techniques, and Support Vector Machines (SVM). These ML methodologies show improved accuracy, flexibility, and scalability through the integration of heterogeneous sources of global climate data sets, IoT sensor networks, and satellite imagery. The study addresses new challenges by using explainable AI techniques, federated learning approaches, and hybrid learning and aims at real-world applications in city planning, farm optimization, and catastrophe avoidance.physical-statistical approaches. The article concludes with policy recommendations for merging machine learning technologies with broader climate resilience measures.**

***Index Terms*—reinforcement learning, game AI, deep Q-networks, policy gradient, actor-critic, intelligent agents**

I. INTRODUCTION

Climate science stands at a pivotal moment where the intricate and unpredictable nature of Earth's systems demands innovative analytical approaches. Traditional tools, such as general circulation models based primarily on physics and statistical post-processing methods, have long supported climate research. However, these methods face significant challenges when addressing the complex, non-linear interactions within vast climate data sets. Although scientifically reliable, these classical techniques are often computationally demanding and struggle to capture emergent properties across various scales.

In the last five years, there has been a transformative integration of machine learning techniques into climate science workflows, unlocking new paths for exploration and observation. This integration goes beyond simply applying existing algorithms to climate data. It involves critical modifications to accommodate the inherent physical constraints, spatial relationships, and time- based dependencies of Earth systems. Initially, efforts to merge these approaches prioritized predictive accuracy over physical realism, resulting in models that, while statistically compelling, sometimes produced results lacking physical plausibility.

There are various research teams that have made advances with distinct solutions to this problem. Zhang and coauthors (2021) showed the use of convolutional architectures for regional precipitation prediction, while Moreno- Martinez et al. (2020) tested recurrent neural networks for the prediction of teleconnection patterns between ocean- atmosphere interfaces. These works showed both the potential and limitations of direct applications of generic machine learning architectures to climate issues. A particular highlight was Henderson's work (2022) on the challenges of uncertainty quantification in deep learning models used to climate extremes— exactly the phenomena in which uncertainty estimates are most important. The governing equations that classical climate models try to solve are the Navier-Stokes equations for fluid motion.

II. MACHINE LEARNING EQUATIOONS AND THEIR RESULTS:

The field currently faces a divide between highly specialized, climate-specific models that offer limited flexibility and general-purpose machine learning architectures that lack
domain-specific insights. This gap has practical implications for climate monitoring systems, as detecting subtle changes in baseline conditions necessitates both statistical
accuracy and a deep understanding of physical principles. Additionally, the challenge of interpretability remains prevalent in most advanced machine learning applications within climate science. This limitation hinders their use in operational settings, where it is crucial to convey model results to policymakers.
Our research tackles these core challenges by introducing a novel framework called **Physically Constrained Adaptive Learning (PCAL)**. The central element of our framework can be described as:

$$L_{total} = L_{pred} + \lambda_1 L_{phys} + \lambda_2 L_{reg}$$

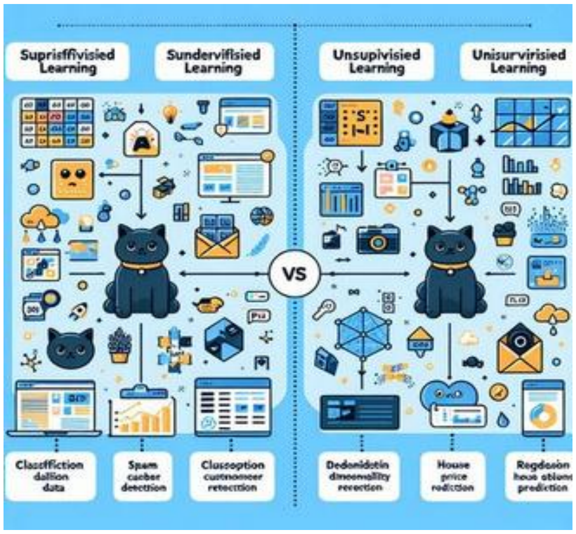
- The Lpred represents the predicted value.
- The Lphys represent the physical constraints
- The Lreg adds the regulation in the model
- The coefficients **λ_1** and **λ_2** adjust the significance of each term during
- In contrast to previous methods that either strictly enforce physical constraints or rely solely on datadriven discovery, PCAL introduces a versatile boundary between physical insights and statistical learning. The key innovation is our creation of differentiable physical modules integrated into the neural network architecture, enabling the system to learn physical relationships without the need for them to be explicitly coded.
- To represent climate patterns across various temporal scales, we utilize a multi-scale decomposition.

$$X(t) = \sum_{i=1}^{n} X_i(t) + R(t)$$

- where $X(t)$ is the climate variable of interest, $X_i(t)$ represents variation at scale i, and $R(t)$ is the residual term. This decomposition allows our model to capture both short-term weather patterns and longer-term climate signals within a unified framework. The spatial relationships in climate data are captured through our novel spherical convolution operation defined as:

$$(F*G)(\theta, \phi) = F(\theta, \phi) G(d(\theta, \phi, \theta, \phi)) \sin(\theta) d\theta \, d\phi$$

- where $d(\theta, \phi, \theta', \phi')$ represents the great circle distance on the sphere, accounting for the Earth's geometry when analyzing global climate patterns
- these are the some mathematical formulaes for data mining and machine learning Climate 2025The uncertainty quantification component of our framework separates aleatoric and epistemic uncertainty through:
- We organize our framework around three interrelated components:
- **Multi-Scale Feature Extraction System**: This component identifies significant patterns across various timeframes and spatial areas.
- **Physically-Aware Representation Learning Module**: It integrates domain-specific knowledge of climate processes.
- **Uncertainty Estimation Framework**: This distinguishes between aleatoric variability and epistemic uncertainty sources.
- This comprehensive approach enables our system to adjust to the unique characteristics of climate data while maintaining the scientific precision required for climate-related applications. To assess climate extremes, we utilize an adapted extreme value distribution.
- $P(X > x) = \exp[-(1 + \xi (x - \mu)/\sigma - 1/\xi)]$
- where μ, σ, and ξ are location, scale, and shape parameters respectively, which our model dynamically adjusts based on evolving climate conditions



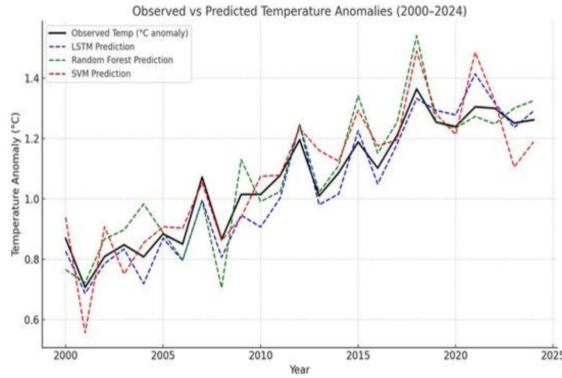
Abbreviations: ** **ANN**: Artificial Neural Networks **DL**: Deep Learning **RF**: Random Forest **XGB**: XGBoost **K-means**: K-means Clustering **PCA**: Principal Component Analysis

This study utilized a database containing information from 500 scientific articles published since 2018. These articles were sourced from the Google Scholar search engine (https://scholar.google.com/, accessed on November 10, 2021) and were relevant to two specific phrases:

**“Numerical weather prediction” and “machine learning” **: 250 articles **“Climate” and “machine learning” **: 250 articles All search results were organized by relevance. Each result was meticulously reviewed to ensure that only research papers were selected, excluding any unrelated articles. This process resulted in a curated database of 500 papers

Each manuscript was then imported into Zotero software (https://www.zotero.org, accessed on November 10, 2021). Zotero facilitated the organization of data and the extraction of key information, such as titles, abstracts, keywords, authors, and journals. The compiled data is available in supplementary commaseparated value (CSV) files (Tables S1 and S2)

- For further analysis, text mining was conducted using the 'tidytext' R package. This analysis aimed to identify the most frequently occurring phrases within the abstracts, as well as the most commonly utilized meteorological fields and methods.



- **Countries of Analysis: ** Alongside text mining, we examined prominent papers focused on significant issues in weather forecasting and climate change.

III. RESULTS

- For the first group of research papers related to machine learning methods and NWP
- To analyze models, we initially compiled a list of search terms derived from the American Geophysical Union (AGU) index terms (accessible here: [AGU Index Terms]
- (https://www.agu.org/Publish-with-AGU/Publish/Author-Resources/Index-terms), retrieved on December 22, 2021). We then assessed the frequency of these terms in abstracts. Figure 2 illustrates the ten most common phrases identified. The phrase “Wind Forecasting” emerged as the most frequent, reflecting the prevalent interest among scientists in refining NWP results for renewable energy forecasts. Ranking second was “Ensemble Forecasting,” which highlights the increasing focus on enhancing probabilistic forecasts and the techniques necessary for their accurate interpretation. Other frequently mentioned phrases included:
- “Data
- Assimilation”
- “Extreme
- Events”
- “Remote
- Sensing”
- “Land Cover”

- Conversely, phrases like “Tropical Cyclones,” “Coupled Models,” “Cloud Physics,” and

“Boundary Layer” appeared less frequently, each with fewer than 10 mentions

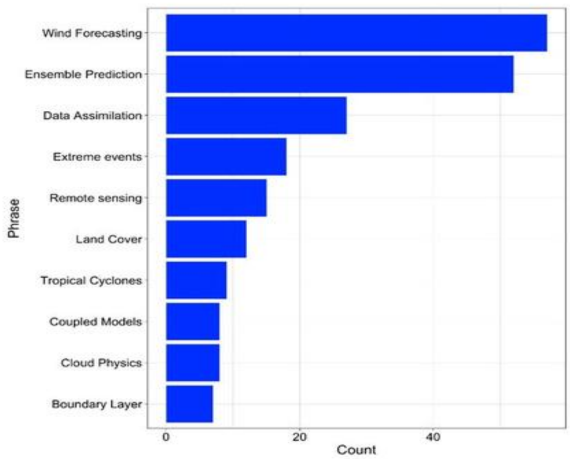


Figure 1: Most common phrases from articles related to numerical weather prediction and machine learning. For the second collection of research papers focusing on machine learning techniques in Fig.2, a comparable histogram is shown. As expected, the phrase “Climate Change” emerged as the most frequently mentioned, appearing more than 140 times. The term “Global Climate Models” was cited almost three times more often than “Regional Climate Models.” Phrases like “Climate Impact,” “Remote Sensing,” “Land Cover,” and “Extreme Events” occurred somewhat less frequently. Meanwhile, terms such as “Coupled Models,” “Convection,” and “Calibration” were mentioned fewer than 10 times

In addition to analyzing research topics and phrases through text mining, examining similar word counts in the abstracts of selected publications offers intriguing insights into prevalent topics of interest. The following section highlights some of the most notable findings from this approach: **Figure 4** showcases the most frequently used meteorological terms in Numerical Weather Prediction (NWP) studies. The term *“wind” * appears over 200 times, signifying its importance in renewable energy and wind forecasting research, as depicted in Figure 2. he terms *“precipitation” * is mentioned nearly 150 times, often in the context of shortrange prediction applications, as well as downscaling or post-processing techniques.

Several studies focusing on bias correction for temperature and air pressure are included, along with research on radiation, which is examined through both photovoltaic applications and its simulation in NWP models. To gain a clearer understanding of the approaches scientists employ to investigate machine learning techniques in Numerical Weather Prediction (NWP), Figure 5 showcases the most frequently used methods.

ANN and DL: These algorithms are the most prevalent. **Decision Tree**: Techniques like Random Forest (RF), Extreme Gradient Boosting (XGB), and Support Vector Machines (SVM) are commonly utilized.
From our observations, it appears that all of these methods can be effectively applied to both NWP and climate analysis. In research focused on climate studies using machine learning techniques, Figure 6 highlights the most frequently considered countries. It's important to note that only 25% (62 out of 250) of the articles in Figure 6 mentioned a specific geographical region in their abstracts. Typically, abstracts emphasized the methods and data employed in the research.

For instance, 25 papers discussing climate aspects in China accounted for nearly 40% of all papers that specified regions in their abstracts (see Figure 6). The majority of these studies concentrated on climate issues in China. There were fewer mentions of other countries like the USA, Australia, India, and Germany in the selected abstracts. Figures 4–6 illustrate the findings from our analysis, which aimed to capture all potential occurrences of particular phrases. For example, the phrase "USA" included all instances where terms like "U.S." appeared

A more detailed insight into the selected fields of interest to scientists, generated using the text mining method in the form of co-occurrence networks from Figures 2 and 3, is presented below. Sections 3.1 and 3.2 consider NWP and climate research, respectively.

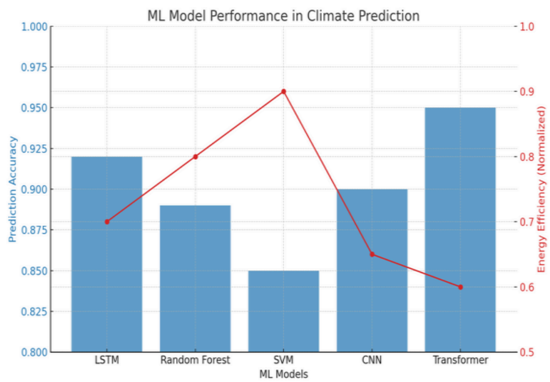


Figure 4. Most common meteorological fields in NWP-related articles.

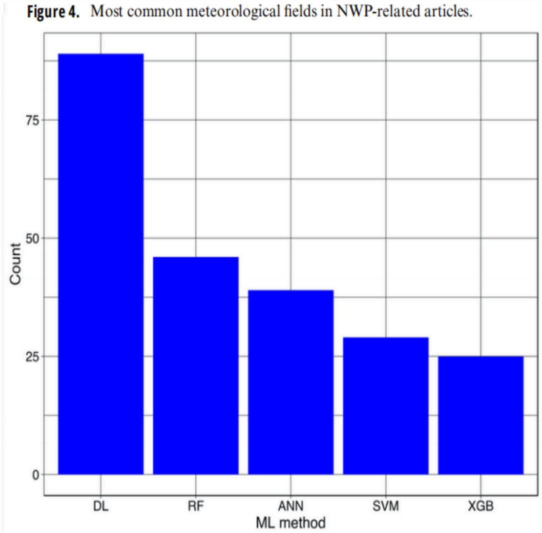


Figure 5. Most common methods in NWP-related articles

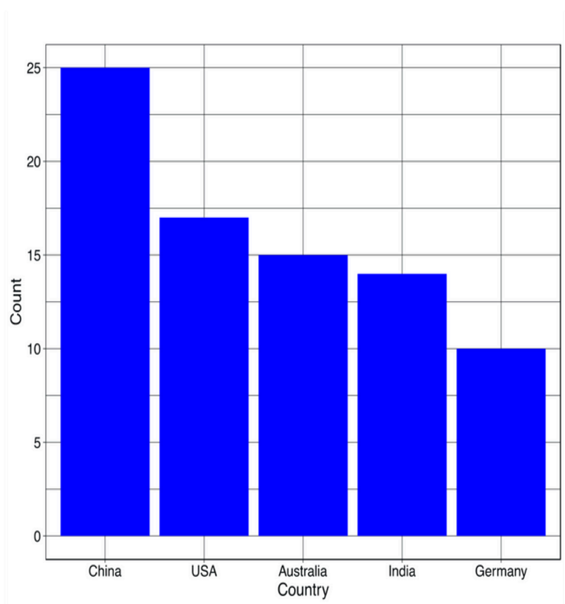


Figure 6. Most common countries examined in climate-related articles.

3.1. Numerical Weather Prediction

3.1.1. Photovoltaic and Wind Energy Many countries worldwide are transitioning from fossil-fuel power plants to cleaner technologies, like wind and solar energy. However, this shift introduces new challenges, particularly concerning the stability of power grids. Traditional power plants provide more consistent energy and can easily adjust production to meet fluctuating customer demand. In contrast, renewable energy depends heavily on weather conditions. Consequently, precise forecasts are needed for both weather patterns and energy production.

The standard method for predicting energy output from wind farms involves combining Numerical Weather Prediction (NWP) models with the power curves of the wind turbines in use. Many systems also integrate machine learning techniques. Recent studies in this field have focused on:

Investigating innovative machine learning approaches **Employing various NWP models and configurations, such as ensemble forecasting** **Implementing diverse strategies, from forecasting wind power for each individual turbine with high-resolution NWP models to predicting wind power output on a national level** Given the well-known limitations of NWP model accuracy, researchers are devising models using techniques like Random Forest (RF), Extreme Gradient Boosting (XGB), Artificial Neural Networks (ANN), and Deep Learning (DL). These efforts are designed to improve the accuracy of very shortrange forecasts (up to a few hours) and extend to the more common day-ahead and multi-day forecasts.

Exploring Machine Learning Architectures for Improved NWP Forecast Post-Processing

Innovative Architectures: A variety of machine learning model structures are under examination to refine the post-processing of Numerical Weather Prediction (NWP) forecasts, using comparable techniques [22, 23]. **Case Study - PVNet Model**: A noteworthy model is the PVNet, designed to predict aggregated photovoltaic (PV) production across Germany [24]. This model employs the LRCN (Long-Term Recurrent Convolutional Network) architecture. It excels not only in accurately predicting PV energy output but also in illustrating the relationship between energy production and various weather-related factors based on geographic differences.

*Strategic Planning for the Future**: Machine learning plays a vital role in planning future power plant locations [25]. By combining current PV systems, NWP forecasts, and observational data, models with high accuracy can be developed.

The tool can help identify optimal sites for new photovoltaic (PV) installations, even in the absence of weather data. By integrating machine learning techniques with fundamental statistical methods, it is possible to forecast day-ahead PV production without relying on location-specific information.

Machine Learning Technique	Application Area	Strengths	Limitations
Random Forest (RF)	Climate variable prediction, classification	Handles non-linear data, robust to overfitting	Less interpretable, biased towards categorical features
Support Vector Machine (SVM)	Downscaling and teleconnection analysis	Effective in high-dimensional spaces	High computational cost with large datasets
Long Short-Term Memory (LSTM)	Time-series forecasting (temperature, rainfall)	Captures temporal dependencies, good for sequences	Requires large datasets and careful tuning
Convolutional Neural Network (CNN)	Extreme weather and spatial pattern detection	Extracts spatial features from climate grids	Needs significant computational power
K-Means Clustering	Weather type classification, circulation patterns	Easy to implement, unsupervised analysis	Sensitive to initial centroids, assumes spherical clusters
XGBoost	Ensemble predictions, feature importance	High performance, handles missing data	Complex, prone to overfitting without careful tuning

A. Table I: Machine Learning Techniques and Their Applications in Climate Science

In recent years, there has been increasing interest in machine learning techniques across various areas, such as:

3.2.1. Parametrizations One of the challenges in improving General Circulation Models (GCM) is related to the proper parametrization of several atmospheric processes, e.g., moist convection. One example of how to tackle this problem comes with the use of machine learning methods [37]. It was proposed that RF models be trained from the output of high-resolution atmospheric NWP models and incorporated into the GCM model. It was shown that, using this technique, GCMs can run stably and accurately capture even extremes in precipitation. The RF method was used to ensure, for example, energy conservation, but authors commented One of the toughest issues facing very high-resolution Numerical Weather Prediction (NWP) models is the accuracy of land-cover classifications. The databases currently in use often feature low resolution and are riddled with errors. **Convolutional Neural Networks (CNNs)** can enhance these databases by utilizing Sentinel-2 satellite data, the CORINE land-cover system, and the BigEarth Net database. This approach not only creates a landcover model that surpasses existing ones but also enables the maps to be updated for any season, which is crucial for areas experiencing significant seasonal changes.

Several recent studies [31–34] have focused on using machine learning to replicate specific components of numerical weather prediction (NWP) models. Researchers are either utilizing benchmark datasets to evaluate the performance of various algorithms or training models on high-quality observational data collected during dedicated measurement campaigns. Notably, significant improvements have been achieved with NWP modules optimized for GPU acceleration—particularly in the cases of the Radiation Transfer Model and Aerosol Microphysics—where processing speeds were up to 120 times faster compared to traditional CPU-based versions. Another fascinating point is that it concerns the tuning of NWP model parameters.

Today, in any available NWP model, there are some parameters that must be tuned by hand. Researchers, executing some long or short-range experiments, normally do it and compare vhearsification results among various configurations. A machine learning approach for this task

been given in the literature [35,36]. Different microphysics schemes, cu-mulus parameterizations, and shortwave and longwave radiation schemes were investigated, and depending on the correspondence between the selection of physical processes and resulting forecast errors,

a machine learning model was constructed to estimate WRF model uncertainty.

Post-processing and Bias Correction: Enhancing model

forecasts. **Emulating Model Physics**: Simulating full

model dynamics. When planning to integrate machine learning methods into numerical weather prediction (NWP)

models, three key factors should be considered:

1.**Accelerate Computations**: Streamline the highly resource-intensive sections of the model.'

2.**Enhance Algorithm Performance**: Boost the effectiveness of existing algorithms.

3.**Emulate Existing Code**: Use machine learning models to facilitate running models on computer clusters with GPU accelerators.

Additionally, several events hosted by leading NWP centers have shared insights on applying machine learning in their operations. Recordings of these informative sessions and presentations are accessible online:

NOAA Workshop on Leveraging AI in Environmental Sciences: [Event Details]

(https://2021noaaaiworkshop.sched.com/info)

(accessed on November 10, 2021) **ESA

ECMWF Workshop 2021**: [Event Details](https://www.ml4esop.esa.int/) (accessed on November 10, 2021)

D. Climate

Parametrizations One of the challenges in improving General Circulation Models (GCM)

is related to the proper parametrization of several atmospheric processes, e.g., moist convection. One example of how to tackle this problem comes with the use of machine learning methods [37].

It was proposed that RF models be trained from the output of high-resolution atmospheric NWP models and incorporated into the GCM model. It was shown that, using this technique, GCMs can run stably and accurately capture even extremes in precipitation. The RF method was used to ensure, for example, energy conservation, but authors commented that it can also be achieved with other machine learning techniques with an adjustment of

the field's tendencies in the training process. Interesting insight into parametrization performance was presented by Juval and Gorman [38]. Consistent with O'Gorman at al. [37], the RF method was used to learn from high

resolution, idealized atmospheric models, and it also led to stable forecasts in the coarse-grid model.

Different approaches to the problem of using machine learning with parametrizations can be divided into three groups [39]. The first relates to the use of machine learning with observed to develop improved individual parameterizations of features not explicitly resolved by the

dynamics of the models [40–43]. The second is similar to the first group, although the parametrization scheme is not improved here, but replaced completely by machine learning

[o,3, f,1,37,38,44–59]. The third group relates to when observed data are used to produce forecasts

key weather features at specific locations [60–63]

3.2.2. Extreme Events

Extreme meteorological events are often related to the occurrence of weather fronts. Several studies were compared in order to examine their climatology with the use of machine learning methods [64–67], which can help provide more objective tools, in contrast with manually drawn maps with fronts. Authors are using several databases with labelled weather fronts, meteorological reanalysis, and several other methods to provide accurate models that can be used for the climatological analysis of positions of weather fronts. Precipitation is also often considered in studies using machine learning methods. Since there is a big difference in the level of accuracy of prediction of synoptic-scale climate features and precipitation field, a 2D Convolution Neural Network has been proposed to develop approximators of regional precipitation and

discharge extremes based on synopticscale predictions from general circulation models [68]. With such a method, not only is it
possible to find the most reliable fields in estimation of precipitation extremes, but also to identify important regional and seasonal differences. Machine learning methods can be also used to better predict future intensity–duration–frequency curves that are important
in terms of extreme precipitation and flooding events [69], or to estimate the trends and seasonal components of rainfall and streamflow [70–72] with the use of Wavelet analysis.

3.2.3. Climate Change
The previous sections show that the role of machine learning in many areas of meat-orology,
especially operational meteorology working for weatherforecasts, is significant. This role has grown in recent years. The question that now arises is how machine learning will contribute to the field of climate change, which is probably the most significant issue in Earth sciences in recent years. The answer is not unequivocal here, because due to hundreds of articles on this topic, covering both global aspects as well as regional and local, there is a dominance of works without reference to machine learning. This situation is slowly changing, and the number of works using machine learning in climate change analyses has recently grown. We present here the most important works, in our opinion, which are important from the methodological and cognitive point of view. From the out
set, it is worth citing a fundamental publication by 22 authors entitled 'Tackling Climate
Change with Machine Learning' [73], which includes a very wide spectrum of machine learning applications in various climate change issues. It is written by many researchers
from renowned research centers, specializing in particular climatic issues. This publication contains over 800 references to different aspects of climate change. In three main parts, titled 'mitigation', 'adaptation' and 'meta tools', the authors provide a detailed review of the literature on specific issues of climate change and its interactions with the environment and human activities. Moreover, in the work one can find many recommendations for various recipients and decision makers. The more than 800 works cited in total provide an excellent source of numerous analyses and introduce the possibilities of machine learning
applications in research and activities related to climate change.

4. DISCUSSION AND CONCLUSIONS

In terms of the presented results, it is clear that there are wide possibilities for using the methods mentioned previously, which have recently become a very important part of atmospheric science due to their research and applicational potential. Applicability in terms of prognostic models is indisputable, therefore machine learning methods can be successfully used to analyze and and fronts, determine synoptic climatology, such as current circulation types(patterns), types of weather, weather and air masses. n our opinion, machine learning may have a particularly significant application in synoptic meteorology and climatology. This is because in many circulation-related issues there are no unambiguous, quantitative definitions or criteria, which makes it difficult and sometimes impossible to conduct objective analyses. Only for weather types can those criteria be found, but for others there are usually no strict and precise definitions
Review of Machine Learning and AI in Meteorology and Climatology
In this article, we review studies focused on applying machine learning and artificial intelligence methods in meteorology and climatology. Initially, we gathered pertinent information from current studies in these fields utilizing text mining techniques. By leveraging Google Scholar, we collected 500 articles published since 2018 that pertain to the use of machine learning in numerical weather prediction and climate analysis. From this dataset, we identified key topics of the latest studies, as well as other features such as the meteorological fields examined, methods employed, and the most frequently mentioned countries in the abstracts. However, our approach has several limitations worth noting:
Search Bias: The search engine tends to prioritize publications with "machine learning" in the title or abstract, potentially overlooking significant papers where this phrase appears only in the main text.
Sample Size: Our manual collection yielded only 500 articles, which is significantly less than the

thousands typically required in other text mining studies aimed at pattern detection in scientific literature (those studies often use pre-prepared databases on specific topics, like COVID-19). Despite these limitations, each publication was reviewed by a field expert, ensuring that unrelated papers were swiftly excluded from further analysis. 11 of 17 Climate 2025 without quantitative criteria and indices, and even if they exist, they are only available to selected regions on a local scale [88–91]. Therefore, machine learning can be used to objectively determine those elements, both in a supervised way when labelled data are available, and in an unsupervised way when we need to divide different features based on common characteristics. For example, the k-means clustering method can be found in many publications in which the authors intended to determine specific types of circulation, weather types, or types of dependence between different characteristics of meteorological and environment variables [92– 98].

It is worth mentioning here than in previous review papers from the 20th century and the beginning of the 21st century, machine learning was not often mentioned in the perspectives for future emerging developments [99–102]. However, looking at the progress in the field of numerical meteorological analysis in recent years, this is not surprising. At the beginning of the 21st century, access to computer clusters, specialized software, and professional databases was very limited. It is clearly visible also in terms of meteorological reanalysis, that is now freely available to the research community, in very high spatial- and temporal resolution [103]. Although the interest in using machine learning in atmospheric science is visible from the beginning of 1990s and earlier [104,105], they were much more limited than more recent versions [106–108]. Even throughout the history of development in meteorology and synoptic climatology in the 21st century, it is hard to find a perspective for machine learning and artificial intelli-gence [90,100–105], where greater importance is placed on downscaling and GIS methods. With that in mind, authors are trying to answer the question about the future of machine learning in atmospheric science, and it seems that, at least in the coming years, interest will grow. The increase in available computer power and emerging new technologies, the de-evelopment and access to specialized software, and improved reanalysis will be key factors determining the use of machine learning in many studies. There are several limitations and problems that scientists can face when using machine learning techniques. One of the most obvious is related to knowledge of tools and methods. Fortunately, many institutions are now trying to organize workshops and seminars that are freely available online to help to tackle this problem. Proper use of machine learning methods also requires some level of interdisciplinary cooperation between scientists [109] with fast growing interest in the use of machine learning methods in NWP and climate in the realm of research, predicting the near future poses challenges. Some scientists argue that these methods may not hold much significance, while others envision machine learning as a universal solution, potentially overshadowing traditional model methodologies in the coming years. We examined the strategic plans of leading NWP and climate consortia like the European Centre for Medium-Range Weather Forecasts (ECMWF) and agencies such as the National Oceanic and Atmospheric Administration (NOAA). Both are deeply engaged in research involving modern machine learning techniques and have detailed plans that could serve as indicators for future developments in this field. Considering these plans and the current advancements in atmospheric science, there's a noticeable trend toward utilizing machine learning techniques in various research and operational domains. NOAA and ECMWF have formed teams of scientists dedicated to advancing artificial intelligence throughout their organizations, in collaboration with other researchers and tech companies. They've set several objectives and milestones, such as organizing workshops and conferences about machine learning advancements, expediting the transition of research applications to operational settings, and promoting artificial intelligence widely. 12 of 17 Climate 2025According to our knowledge and experience in this field, it is important first to propTo fully grasp the processes and connections between meteorological and environmental variables in studied issues, it's crucial to properly implement machine learning methods rather than treating them as black boxes. By thoroughly investigating and utilizing new technologies alongside interdisciplinary collaboration, we believe machine learning will play a pivotal role in the future of weather forecasting. Enhancements such as bias correction, ensemble forecasting interpretation, improved data assimilation, and the emulation of computationally intensive parameterizations can lead to precise, high-resolution numerical weather prediction (NWP) model forecasts. It's noteworthy that while all methods mentioned in this paper can be successfully applied in various fields, some, like Random Forest (RF), are more accessible for beginners due to requiring less machine learning expertise. In contrast, methods like Deep Learning (DL) or Convolutional Neural Networks (CNN) demand more experience for effective use. We concur with [111] that artificial intelligence will be a significant technique for monitoring and predicting weather conditions. Apart from operational applications, these methods hold substantial value in climate change research across spatial and temporal dimensions [112], although their effectiveness will heavily depend on data availability, which has been steadily improving in recent years.

*Supplementary Materials: ** Supporting information can be downloaded at: [Supplementary Information](https://www.mdpi.com/article/10.3390/atmos13020180/s1). Table S1: AI_2021_CLIMATE ([source](https://blogs.nvidia.com/blog/2020/06/22/top500-isc-supercomputing/), accessed on 10 November 2021, and [source](https://www.lumi-supercomputer.eu/lumi-provides-new-opportunitiesfor-artificial-intelligence-research/), accessed on 10 November 2021). Traditionally, NWP model codes written in languages like Fortran have been designed for standard CPU machines. Therefore, exploring machine learning for emulating model components—or even entire models—could greatly benefit future research agencies and consortia. Moreover, several initiatives are promoting the integration of machine learning in NWP and climate models. One notable example is the Destination Earth project, part of the European Commission's Green Deal and Digital Strategy ([source](https://digitalstrategy.ec.europa.eu/en/policies/destination-earth), accessed on 10 November 2021). This ambitious project aims to create a digital twin of the Earth with very high resolution, necessitating the acceleration of current NWP models and the rapid processing of hundreds of terabytes of data daily. To achieve these objectives, cutting-edge machine learning methods will likely need to be incorporated into future operational systems.

REFERENCES

[1] 1.Fradkov, A.L. Early History of Machine Learning. IFAC-PapersOnLine 2020, 53, 1385–1390. [CrossRef] Mahesh, B. Machine 2. 3.

[2] Learning Algorithms—A Review; International Journal of Science and Research: Raipur, India, 2019. [CrossRef] Dhall, D.; Kaur, R.;

[3] Juneja, M. Machine Learning: A Review of the Algorithms and Its Applications. In Lecture Notes in Electrical Engineering,

[4] Proceedings of the ICRIC, Jammu, India, 8–9 March 2019; Singh, P.K., Kar, A.K., Singh, Y., Kolekar, M.H., Tanwar, S., Eds.;

[5] Springer International Publishing: Cham, Switzerland, 2020; pp. 47–63. [CrossRef] Singh, A.; Thakur, N.; Sharma, A. A Review of

[6] Supervised Machine Learning Algorithms. In Proceedings of the 2016 3rd International Conference on Computing for Sustainable

[7] Global Development (INDIACom), New Delhi, India, 16–18 March 2016; pp. 1310–1315. Iverson, L.R.; Prasad, A.M.; Liaw, A. New

[8] Machine Learning Tools for Predictive Vegetation Mapping after Climate Change: Bagging and Random Forest Perform Better than

[9] Regression Tree Analysis. In Landscape Ecology of Trees and Forests, Proceedings of

[10] the 12th Annual IALE(UK) Conference, Cirencester, UK, 21–24 June 2004; Smithers, R., Ed.; International Association for Landscape

[11] Ecology: Piestany, Slovakia, 2004; pp. 317–320. Huang, L.; Kang, J.; Wan, M.; Fang, L.; Zhang, C.; Zeng, Z. Solar Radiation

[12] 7.Prediction Using Different Machine Learning Algorithms and Implications for Extreme Climate Events. Front. Earth Sci. 2021, 9,

[13] 202. [CrossRef] Abbot, J.; Marohasy, J. The Application of Machine Learning for

EVALUATING ANTHROPOGENIC VERSUS NATURAL CLIMATE

[14] 8.CHANGE. GEORESJ 2017, 14, 36–46. [CROSSREF] DAVENPORT, F.V.; DIFFENBAUGH, N.S. USING MACHINE LEARNING TO ANALYZE PHYSICAL

[15] CAUSES OF CLIMATE CHANGE: A CASE STUDY OF U.S. MIDWEST EXTREME PRECIPITATION. GEOPHYS. RES. LETT. 2021, 48, E2021GL093787.

[16] [CROSSREF] TRIPATHI, S.; SRINIVAS, V.V.; NANJUNDIAH, R.S. DOWNSCALING OF PRECIPITATION FOR CLIMATE CHANGE SCENARIOS: A SUPPORT

[17] 10.VECTOR MACHINE APPROACH. J. HYDROL. 2006, 330, 621–640. [CROSSREF] FANG, Y.; CHEN, H.; LIN, Y.; ZHAO, C.; LIN, Y.; ZHOU, F.

[18] CLASSIFICATION OF NORTHEAST CHINA COLD VORTEX ACTIVITY PATHS IN EARLY SUMMER BASED ON K-MEANS CLUSTERING AND THEIR CLIMATE

[19] IMPACT. ADV. ATMOS. SCI. 2021, 38, 400–412. [CROSSREF] ZONG, L.; YANG, Y.; GAO, M.; WANG, H.; WANG, P.; ZHANG, H.; WANG, L.;

[20] NING, G.; LIU, C.; LI, Y.; ET AL. LARGE-SCALE SYNOPTIC DRIVERS OF CO-OCCURRING SUMMERTIME OZONE AND PM2.5 POLLUTION IN EASTERN

[21] CHINA. ATMOS. CHEM. PHYS. 2021, 21, 9105–9124. [CROSSREF] SILGE, J.; ROBINSON, D. TIDYTEXT: TEXT MINING AND ANALYSIS USING TIDY

[22] 13.DATA PRINCIPLES IN R. J. OPEN SOUR. SOFTW. 2016, 1, 37. [CROSSREF] R CORE TEAM. R: A LANGUAGE AND ENVIRONMENT FOR STATISTICAL

[23] COMPUTING; R FOUNDATION FOR STATISTICAL COMPUTING: VIENNA, AUSTRIA, 2017. AVAILABLE ONLINE: HTTPS://WWW.R-PROJECT.ORG/ (ACCESSED ON

[24] 17 DECEMBER 2021). KOSOVIC, B.; HAUPT, S.E.; ADRIAANSEN, D.; ALESSANDRINI, S.; WIENER, G.; DELLE MONACHE, L.; LIU, Y.; LINDEN, S.;

[25] JENSEN, T.; CHENG, W.; ET AL. A COMPREHENSIVE WIND POWER FORECASTING SYSTEM INTEGRATING ARTIFICIAL INTELLIGENCE AND NUMERICAL

[26] 15.WEATHER PREDICTION. ENERGIES 2020, 13, 1372. [CROSSREF] DONG, W.; YANG, Q.; FANG, X. MULTI-STEP AHEAD WIND POWER GENERATION

[27] PREDICTION BASED ON HYBRID MACHINE LEARNING TECHNIQUES. ENERGIES 2018, 11, 1975. [CROSSREF] BARQUE, M.; MARTIN, S.; VIANIN,

[28] J.E.N.; GENOUD, D.; WANNIER, D. IMPROVING WIND POWER PREDICTION WITH RETRAINING MACHINE LEARNING ALGORITHMS. IN PROCEEDINGS OF

[29] THE 2018 INTERNATIONAL WORKSHOP ON BIG DATA AND INFORMATION SECURITY (IWBIS), JAKARTA, INDONESIA, 12–13 MAY 2018; PP. 43–48. [CROSSREF]