European Commission

Horizon Projects using

ENVIRONMENTAL OBSERVATIONS AND ARTIFICIAL INTELLIGENCE

for the benefit of science and society

Research and Innovation

Horizon Projects using environmental observations and artificial intelligence for the benefit of science and society

European Commission
Directorate-General for Research and Innovation
Directorate B — Healthy Planet
Unit B3 — Climate & Planetary Boundaries

Contact Franz IMMLER
Email RTD-EARTH-OBSERVATION@ec.europa.eu
Franz.IMMLER@ec.europa.eu
RTD-PUBLICATIONS@ec.europa.eu
European Commission
B-1049 Brussels

Manuscript completed in July 2024
1st edition.

PDF ISBN 978-92-68-18659-6 doi:10.2777/167982 KI-02-24-699-EN-N

Luxembourg: Publications Office of the European Union, 2024

EUROPEAN COMMISSION

Horizon Projects using

ENVIRONMENTAL OBSERVATIONS AND ARTIFICIAL INTELLIGENCE

for the benefit of science and society

edited by
Franz IMMLER, Paschalis TZIASTAS

2024 Directorate-General for Research and Innovation

TABLE OF CONTENTS

INTRODUCTION

Over the past decade a revolution in the development and use of Artificial Intelligence (AI) technologies in various domains was enabled by the availability and evolution of modern computational capabilities and the development of new algorithms. These developments have prompted the European Union to create a comprehensive and forward-looking AI policy framework, which aims to ensure that AI technologies are developed and used responsibly, ethically, and in a manner that benefits society as a whole.

Some of the most consequential applications of AI will be achieved in the scientific discovery. Different experts claim that investment in the responsible adoption of AI in Science could be one of the best societally beneficial investments[1], as it can accelerate research and time to access markets, thereby helping to tackle the most pressing challenges like climate change, biodiversity loss or in healthcare. And AI in Science is becoming quickly a reality with a fast adoption pace: yearly growths of 30% in scientific publications using AI as a tool[2]. For this reason, the European Commission is exploring[3] policy options to accelerate the uptake of AI by the scientific community in the EU. As part of these efforts, the EU Scientific Advice Mechanism (SAM) has recently published a report on the uptake of Artificial Intelligence in research and innovation[4], suggesting to set up a European institute for AI in science. A recent CORDIS results pack showcases the use of AI in various scientific domains[5]. The European Research Council published a foresight report which examines the use and impact of Artificial Intelligence in the scientific process[6].

In order to get an overview and better understanding on what, why, and how AI is used in current EO research activities, we have asked a set of questions to selected projects funded under Horizon 2020 and Horizon Europe framework programmes that you can find in the next section.

1 OECD (2023), Artificial Intelligence in Science: Challenges, Opportunities and the Future of Research, OECD Publishing, Paris, https://doi.org/10.1787/a8d820bd-en

2 European Commission, DG RTD (2023), Trends in the use of AI in science, https://research-and-innovation.ec.europa.eu/knowledge-publications-tools-and-data/publications/all-publications/trends-use-ai-science_en

3 European Commission, DG RTD (2023), AI in Science: Harnessing the power of AI to accelerate discovery and foster innovation, DG RTD Policy Brief, 2023, https://research-and-innovation.ec.europa.eu/document/download/1e2a4c9c-d3f1-43e9-9488-c8152aabf25f_en

4 https://scientificadvice.eu/advice/artificial-intelligence-in-science/

5 https://op.europa.eu/en/publication-detail/-/publication/c0e52bea-5bb0-11ee-9220-01aa75ed71a1

6 https://erc.europa.eu/sites/default/files/2023-12/AI_in_science.pdf

The responses from the 19 surveyed projects are presented in section 4 and showcase a wide range of benefits of using different AI technologies in climate and environmental research. The new opportunities are currently driving the scientific progress substantially and are crucial for effective exploitation of EO data and the delivery of value-added products to end-users. As projects were also asked to express their views on research needs in AI, this publication may also be useful to inform future research and innovation strategies and programmes.

This survey may also be of interest for researchers to learn more about the state-of-the-art of AI and inspire the integration of AI in their own research activities. This publication also indicates different experts or groups who use AI, which may be contacted for further information.

As the information comes directly from researchers who work with AI, it provides an actual picture of what its use means in practice, demonstrates its benefits, but also its limitations and related risks and may therefore be useful for an informed debate on the responsible use of AI in Science.

For a wider range of readers, this brochure may be of interest as it indicates what the society can expect from AI in terms of its contribution to accelerating research and innovation, to support decision and policy making, and for producing results that impact citizen's everyday life.

WHAT HAVE WE LEARNED ABOUT THE USE OF AI?

The responses reveal that research groups working with earth observations in EU-funded projects already took advantage of the new capabilities of AI for many years to reach their research objectives in different domains, ranging from biodiversity monitoring to prediction and attribution of climate extremes. The use of AI in this area can therefore be considered comparatively advanced.

A wide range of different AI methodologies are used in EO research activities: Machine Learning (ML), Deep Learning (DL), Neural Network (NN), Neural Language Processing (NLP), Generative AI (genAI), Explainable AI (XAI), and others. In many cases, AI methodologies are used to analyse EO data, assess data quality and uncertainty, or fill data gaps. AI also helps to interpret the data by pattern recognition or type classifications. Many EO projects also use those tools to create or improve models and predictions. These products are then typically provided to end-users to support decision-making.

WHICH ARE THE ADVANTAGES OF USING AI?

Earth observations (EO), from satellite, other remote sensing, or in-situ instruments generate a huge volume of data. AI has the capacity to process and synthesise large data volumes at significantly higher speed and potentially with less computing resources than traditional methods. EO data can therefore be quickly analysed and linked with other data. The capability of AI methods to detect patterns and inter-linkages that otherwise would be hard or even impossible to reveal, allows to deduct deeper and more relevant insights from EO data than ever. These methods provide opportunities for the development of user-friendly applications based on earth observations in various sectors.

WHAT CAN BE IMPROVED?

Challenges encountered in using AI include the complexity of tools and data for non-technical users. Many projects mention the need to further develop the available AI technologies, including methods to ensure data privacy and security, as well as data visualisation and interactivity tools. Another challenge is understanding the causality of results, their robustness, reproducibility, as well as possible uncertainty and biases. For the optimal use of the potential of AI a new generation of scientist needs to be trained, who are not only IT specialist but cover a broad range of skills. For example, a meaningful adoption of advanced AI methods to the climate science domain requires deep knowledge in different disciplines like data science and Earth sciences.

HOW TO READ THIS DOCUMENT

Selected projects, funded from Horizon 2020 and Cluster 5 or 6 in Pillar II of the Horizon Europe programme, that deal with environmental observations and/or climate models, were asked to report on their use of Artificial Intelligence (AI) to achieve their research objectives, following these guiding questions:

Which elements of the project includes AI and what tools and methods are used? What are the advantages compared to more traditional methods?

Why is AI useful or needed to reach the objectives of the project? How does it aid in moving beyond the state-of-the-art?

What are beneficial outcomes for end users, including policy makers, businesses, or citizens? What are their perceived risks and limitations (if any)?

What are research needs related to the use and development of Artificial Intelligence/ Machine Learning technologies? (Optional question)

Answers to these questions, are provided in the **same sequence** and indicated using the **same colour** as above. The name of the researcher who provided the answers, on behalf of the project, is provided below the title of the project, with the respective affiliation and the position in the project.

In the end of each section, you can find some information for the project, (hyper-) links to CORDIS and the project's website (where available), where a lot more details on the project can be found.

Projects replies are presented in alphabetical order of its Acronym.

AKNOWLEDGMENTS

We would like to acknowledge the contribution for the preparation of this publication from our colleagues in RTD B.3, Katarzyna DRABICKA, Mathilde MOUSSON and David ARRANZ in RTD E.4. Also, we would like to thank for their cooperation, the project advisers Lara CONGIOU, Antonio SCARAFINO, Georgios CHARALAMPOUS, Erwin GOOR, Javier MARTIN-MEMBIELA from the Research Executive Agency (REA), and Richard TAVARES and Marko ADAMOVIC from the European Climate, Infrastructure and Environment Executive Agency (CINEA).

4.1 AD4GD

All Data 4 Green Deal - An Integrated, FAIR Approach for the Common European Data Space

Answers provided by Joan Masó Pau (CREAF), project coordinator

AD4GD has prepared a cloud infrastructure to run AI/ML. A first aim is to improve data quality and uncertainty in heterogeneous multidimensional data by automatically detecting and fixing anomalies, such as missing or incoherent data. Then there is the use of Neural Networks (NN), such as Long Short-Term Memory (LSTM) to prtedict future values for specific variables, or conditional Generative Adversarial Networks (cGAN) to translate input maps to useful environmental estimations. Physics Informed ML (PIML) is also considered to obtain customised models for climate analysis purposes.

In the habitat connectivity analysis, current methodologies based on analysis of neighbours and instances are computationally heavy, which translates into slow performances, and we expect that ML is capable to provide similar results and faster responses. In other words, the current analysis consists of obtaining the analytical solution for a highly complex equation, whereas ML could provide an approximation in much less time.
In the water parameters evolution, ML is used to interpolate data and complete time series with gaps or inherences without having to know the physical models behind the variables, as well as to forecast selected variables.

The use of AI is beneficial, because it can interpolate data that was not captured in the past. It can also extrapolate what can happen in the future. In the management of lakes in Berlin, this can help public authorities to anticipate corrective measures to improve water quality and recreational services. In the habitat connectivity case, it can help the regional government to detect hot spots where habitat connectivity issues exist and articulate better mitigation actions.
The main limitations of this methodologies are the lack of knowledge about the causalities of the different factors influencing the system, and the possible adaptation needed as new kind of data wants to be added.

For the case of habitat connectivity analysis, a way to adapt the available data (both for input and output) to state-of-the art image-to-image deep learning algorithms (cGAN) needs to be explored. Then, regarding the water quality analysis, a way to use satellite data to determine chemical characteristics of a water surface using AI would represent an important step forward.

AD4GD All Data 4 Green Deal - An Integrated, FAIR Approach for the Common European Data Space

COORDINATING ENTITY:
CENTRO DE INVESTIGACION ECOLOGICA Y APLICACIONES FORESTALES (SPAIN)

PARTICIPATING COUNTRIES:
Spain, Belgium, Germany, Poland, UK, Luxembourg, Hungary, Switzerland

PROJECT WEBSITE:
https://ad4gd.eu/

START/END DATE:
1 September 2022 / 31 August 2025

EU contribution (€):
4 136 964,75

4.2 AI4PEX

Artificial Intelligence and Machine Learning for Enhanced Representation of Processes and Extremes in Earth System Models

Answers provided by Nuno Carvalhais (MPG - Max-Planck Institute for Biogeochemistry, Jena), project coordinator

AI4PEX uses machine learning (ML) and AI approaches in the Earth system model (ESM) development cycle. The goal is to connect ESMs, large Earth Observation (EO) datasets, and numerical simulations to reduce uncertainties in processes underpinning global feedbacks, particularly those connected to clouds, land and ocean carbon uptake, and ocean heat uptake. The project will focus on three key aspects: observations, modelling, and evaluation. These efforts will be accompanied by analysis of extreme climate events and efforts to develop confidence in ESM based projections. The advantage of the approach is a large emphasis on observations, reduction in errors and interpretable uncertainties in model behaviour.

ML/AI approaches show a substantially larger potential to represent and interpret observational datasets when compared to traditional methods. The use of such methods in our models and in the model development frameworks further improves quantification of uncertainties. By leveraging state-of-the-art causality and explainable AI methods for process-oriented data exploration and model evaluation, we achieve insights into natural processes and the reasons behind ESM uncertainties and steps towards improving them. So far, experiments bridging ML/AI approaches and physical models have only focused on the Earth system in its parts. AI4PEX plans to step beyond the state of the art.

AI4PEX Artificial Intelligence and Machine Learning for Enhanced Representation of Processes and Extremes in Earth System Models

COORDINATING ENTITY:
MAX-PLANCK-GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV (GERMANY)

PARTICIPATING COUNTRIES:
Germany, France, Finland, Denmark, Sweden, Spain, Belgium, Switzerland, UK

PROJECT WEBSITE:
AI4PEX.eu (not yet active)

START/END DATE:
1 April 2024 / 31 March 2028

EU contribution (€):
6 638 882,50

Uncertainty reduction in ESMs would provide improved and more confident projections of Earth's climate. More reliable projections would support IPCC, EU's Green Deal implementation, as well as provide information for different stakeholders and businesses that could be utilized in their decision making bodies and or businesses. Sharing of the tools would support further applications and clarification of use. Furthermore, the demonstration of the use of ML/AI methods for such overarching and socially relevant challenges would also highlight the potential of ML/AI as a versatile tool in the environmental and Earth system science domains. Limitations in the available datasets, as well as the use of models beyond the training domains could lead to interpretation biases and limits in the generalizability of models.

I believe there are five aspects that can catalyse the critical and mature use of ML/AI approaches: 1) further development of causality methods that are both valid on the characteristics of observations as well as on modelling approaches – ensuring that models are generating the right answers for the right reasons; 2) development of uncertainty quantification approaches that can consider both physical and ML/AI uncertainties in modelling outputs – supporting the provision of robust uncertainties in model projections; 3) investigating the extrapolation / out of distribution problems in Earth system science – understanding to what extent can we trust, and how much can we trust, the estimations of model projections; 4) development of computational tools that allow a deep integration between physical modelling and ML/AI codes – to maximize information transfer from observations into ESMs; 5) data visualization and interactivity tools that allow researchers and users to easily inquire models, understand their behaviour and identify errors.

4.3 CITIOBS

Enhancing Citizen Observatories for healthy, sustainable, resilient, and inclusive cities

Answers provided by Amirhossein Hassani (Researcher); Nuria Castell (Project Coordinator), STIFTELSEN NILU

Our project utilizes low-cost air quality sensors (LCSs) in Citizen Observatories (COs) in various European cities for monitoring air quality, raise awareness and for policy making. The core use of AI, particularly Machine Learning (ML) techniques and algorithms, is to increase sensor data quality e.g. by flagging suspicious data, correct them, and make them more consistent with official reference data.
Our approach using AI is superior to traditional methods in several ways.
AI can further assist in integrating trustworthy LCS data with other available environmental data resources, such as meteorological or remote sensing products, to generate continuous, high resolution maps of PM2.5, for example.
Another potential usage, are alert systems for early detection of pollution from agricultural or wildfires. AI algorithms can analyse real-time PM2.5 data from LCSs and identify patterns indicative of increased pollution levels associated with wildfires.

It is possible to use classic statistical techniques to perform quality control and data correction on sensor data. However, AI-based quality control and correction methods are adaptive and can learn from new data.
AI allows us to enhance the quality and reliability of air quality data collected by LCSs. This is relevant as the use of non-validated data from LCSs, can lead to inaccurate or erroneous assessments of air quality. It is important that communities and policymakers base their decisions on validated data of known quality.

CITIOBS Enhancing Citizen Observatories for healthy, sustainable, resilient, and inclusive cities

COORDINATING ENTITY:
STIFTELSEN NILU (NORWAY)

PROJECT WEBSITE:
https://citiobs.eu/

EU contribution (€):
4 991 205,00

PARTICIPATING COUNTRIES:
Norway, Sweden, Ukraine, Netherlands, Ireland, Greece, Belgium, Spain, Poland, Germany, France

START/END DATE:
1 January 2023 / 31 December 2026

Validated LCS data can benefit policymakers, businesses, and citizens in many ways:
Policymakers:
- Formulating evidence-based air quality regulations and policies.
- Implementing targeted interventions to improve air quality in specific areas.
- Monitoring the effectiveness of air quality improvement initiatives.
- Identifying environmental hotspots and prioritizing resources for mitigation efforts.

Businesses:
- Developing and deploying innovative environmental technologies and solutions.
- Optimizing operations to reduce emissions and improve environmental sustainability.
- Enhancing corporate social responsibility efforts and public image.
- Identifying market opportunities for clean energy and environmental products.

Citizens:
- Accessing real-time air quality information to make informed decisions about outdoor activities.
- Protecting personal health by avoiding exposure to high pollution areas.
- Advocating for environmental improvements and holding policymakers and businesses accountable.
- Participating in citizen science projects and contributing to collective efforts to monitor and address air quality issues.

Such data can also be useful for researchers to improve the quality of existing models and datasets for air quality, for example, through data assimilation and data fusion techniques.
Some limitations might be:
AI models developed for LCS data may have limited generalization capabilities, particularly if they are trained on data from specific locations or time periods. Another limitation is the fact that AI models act as a black box, making it challenging to understand and interpret the calibration process. Also, AI models used for LCS data may be overfitted to the training data, capturing noises that are not representative of the true relationship between sensor measurements and actual pollutant concentrations.

Research needs related to the use and development of Artificial Intelligence/ Machine Learning (AI/ML) technologies in the domain of air quality LCSs include:
- Developing AI/ML algorithms designed specifically to address LCS challenges, such as algorithms dedicated to sensor calibration or drift/outlier detection.
- Enhancing the interpretability and transparency of AI/ML models to understand how they make predictions and decisions, especially for policymakers and stakeholders. Typically, commercial LCSs use proprietary AI models and software, required to be more transparent.
- Scaling AI/ML methods to handle large-scale LCS datasets and ensuring their generalization across different geographical regions and environmental conditions.
- Investigating the ethical and societal implications of AI/ML technologies in air quality monitoring using LCSs, mainly regarding privacy concerns.

4.4 CLINT

CLImate INTelligence: Extreme events detection, attribution and adaptation design using machine learning

Answers provided by Andrea Castelletti (Politecnico di Milano), Project Coordinator

In CLINT, we use AI to study the detection, attribution, and causation of extreme climate events such as tropical cyclones, droughts, heatwaves, and compound events. To cover such a wide range of applications, we use several different AI tools, from neural networks to causal feature selection algorithms. Such tools are much better suited than “traditional” ones to processing large climate datasets, which are at the core of our research. They are also able to extract more nuanced patterns, which are key to studying and predicting complex climate events.

When studying extreme events, on one hand you have massive datasets, but on the other the events are so rare that few of the samples you have are informative. AI can help address both problems: it can process data much faster than traditional methods and can also be used to generate new data on these rare events. Furthermore, we still do not fully understand the physics behind some of these events, so whereas traditional methods struggle because they require you to specify the equations governing these phenomena, AI can find patterns directly from the data.

The CLINT project's outcomes offer significant benefits for various end users. For policymakers, it provides actionable insights for adaptation and mitigation policies based on accurate climate data and predictions. Businesses can make informed decisions to minimize risks and optimize operations in the face of extreme weather events. Citizens benefit from improved disaster preparedness and response measures. However, potential risks include reliance on AI models, which may introduce the need for careful interpretation of results to avoid misinforming decisions.

Higher model accuracy and computational efficiency are clearly two targets to strive for, but a third, perhaps more important one is explainability. Especially because we do not fully understand the physics behind these extreme climate events, it's crucial to be able to extract from AI models information that is interpretable, so that we can improve our understanding of these phenomena.

CLINT CLImate INTelligence: Extreme events detection, attribution and adaptation design using machine learning

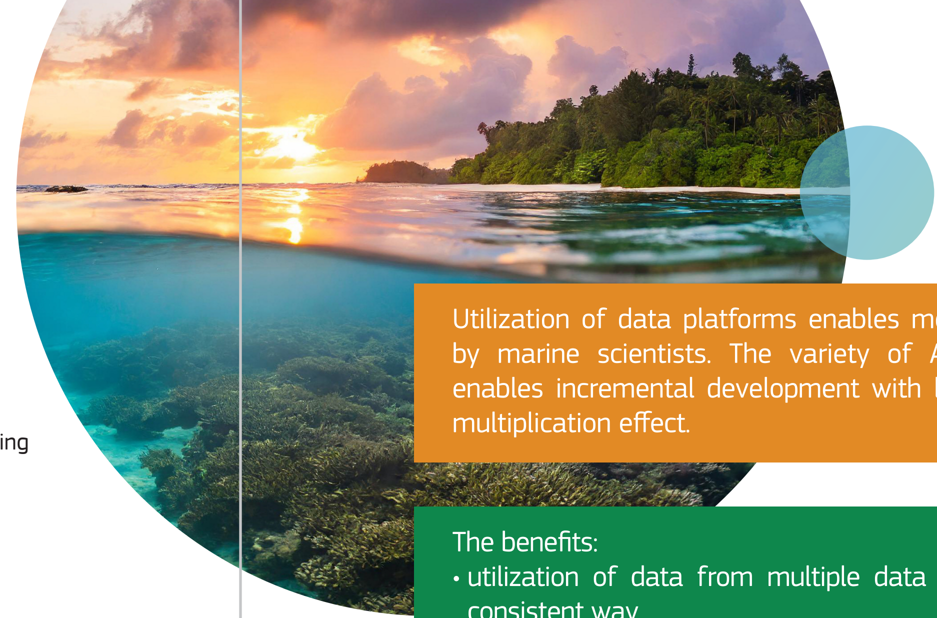
COORDINATING ENTITY:
POLITECNICO DI MILANO (ITALY)

PROJECT WEBSITE:
https://climateintelligence.eu/

EU contribution (€):
6 067 719,98

PARTICIPATING COUNTRIES:
Italy, Germany, Spain, Sweden, Netherlands, Greece, France, United Kingdom, Belgium

START/END DATE:
1 July 2021 / 30 June 2025



4.5 DIVERSEA

Integrated observation, monitoring and prediction architecture for functional biodiversity of coastal seas

Answers provided by Vassil Vassilev (GATE Institute of Sofia University and School of Computing and Digital Media of London Metropolitan University), Research Lead of project WP 2, Hai Nguyen (Chiron Predictive Technologies AS), project WP 3, Robert Brian O'Hara (Department of Mathematical Sciences, NTNU), Leading project WP3 , Murat V. Ardelan (Department of Chemistry, NTNU) Project Coordinator, Pinar Øzturk (Norwegian University of Science and Technology), project WP 3

DiverSea aims in creating a holistic picture of the marine biodiversity in the seas surrounding Europe. The analysis of marine biodiversity dynamics involves multiple tasks – stressors identification, trends recognition, dynamic prediction, factor analysis, etc. They can be solved using AI/ML methods supported by multiple tools. Developing policies for monitoring, controlling and protecting marine biodiversity requires justified decisions. This needs Explainable AI to account the expert knowledge about causal dependencies.
We will apply several unsupervised machine learning methods, such as DBSCAN, KNN, One-class SVM, and Association Mining, to analyze huge amount of data. One emphasis will be e-DNA data, improving current pipelines for its analysis. Our unique combination of hierarchical statistical and AI/ML techniques will let us discover the key relationships between molecular data and other types of data.

Integration of data and meta-data by embedding ontological models in the data space guarantees semantic interoperability and enables federated and distributed data analysis of marine biodiversity data.
The use of system dynamics models requires determining of the key indicators of biodiversity data. Machine learning can automate this by training on historical data which creates hybridization of decision making and data analytics technologies.

DIVERSEA Integrated observation, monitoring and prediction architecture for functional biodiversity of coastal seas

COORDINATING ENTITY:
NORGES TEKNISK-NATURVITENSKAPELIGE UNIVERSITET NTNU (NORWAY)

PROJECT WEBSITE:
https://www.ntnu.edu/diversea/diversea

EU contribution (€):
9 688 013,92

PARTICIPATING COUNTRIES:
Norway, Denmark, Belgium, Italy, Portugal, Spain, Turkey, Ukraine, Bulgaria, Greece, Sweden, Ghana, USA

START/END DATE:
1 September 2023 / 31 August 2027

Utilization of data platforms enables modern DataOps practices directly by marine scientists. The variety of AI tools available on platforms enables incremental development with little investments and significant multiplication effect.

The benefits:
- utilization of data from multiple data sources for different needs in a consistent way
- holistic picture of the marine biodiversity on local, regional and pan-European level
- comprehensive insight into marine biodiversity system by utilization of domain expertise
- support for effective control, conservation and protection policies by impact assessment of environmental and economic factors

The risks of AI/ML adoption is very low since the data space is guarded by combining data processing with domain expertise. Some limitations come from the need for expert model validation. The use of ML is also sensitive to the data quality which may be a particular problem with biodiversity data, as the methods may struggle to distinguish between signals of ecological processes and signals of the way the data were sampled and processed.
Artificial Intelligence and Machine Learning (AIML) will be pivotal in identifying the connections and variations among biogeochemical drivers and stressors that contribute to biodiversity decline in coastal waters. By analysing data across various spatial and temporal conditions, AI can uncover patterns and changes that traditional methods might overlook. This advanced capability is essential for formulating targeted conservation strategies and mitigating the adverse impacts on marine ecosystems.

The spatial and temporal data partitioning of the biodiversity data requires further research into the techniques for replication, synchronization and indexing to enable federated and distributed analytics.
The current explosion in the use of large linguistic models (LLMs) creates a new opportunity to use transformer-based systems such as ChatGPT for providing explanation of the results and justification of the recommendations. Utilizing the full potential of AI/ML in marine biodiversity research can be intensified by adopting adequate and easy to use AI/ML tools and modern engineering practices such as DataOps/MLOps directly from non-technical staff.

4.6 E4WARNING

Eco-Epidemiological Intelligence for early Warning and response to mosquito-borne disease risk in Endemic and Emergence settings

Answers provided by Alex Richter-Boix (CSIC) with the inputs from Joan Garriga (CSIC) and Joao Encarnaçao (IRIDEON), Coordination team

In E4Warning, AI powers Mosquito Alert (citizen science) and IRIDEON's smart-traps for advanced mosquito monitoring. MA uses AI deep learning models to classify mosquito images from a smartphone app, alerting experts when a species appears outside its known range. IRIDEON's smart-traps utilize Machine Learning to identify and count mosquitoes using wingbeat analysis from optical sensors, providing real-time data via IoT. These approaches offer efficient, cost-effective surveillance, significantly reducing time and resources compared to traditional methods.

AI transforms mosquito monitoring by automating identification and boosting data processing, key for improving surveillance methods. Mosquito Alert enables rapid, scalable, and cost-effective monitoring, enhancing traditional methods which rely on physical trapping and manual identification. IRIDEON delivers precise, real-time population data, allowing stakeholders to swiftly respond to potential mosquito-borne disease outbreaks with unprecedented time-scale resolution.

E4WARNING Eco-Epidemiological Intelligence for early Warning and response to mosquito-borne disease risk in Endemic and Emergence settings

COORDINATING ENTITY:
AGENCIA ESTATAL CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (SPAIN)

PARTICIPATING COUNTRIES:
Spain, Belgium, Greece. Germany, Switzerland, UK

PROJECT WEBSITE:
https://www.e4warning.eu/

START/END DATE:
1 January 2023 / 31 December 2026

EU contribution (€):
4 082 528,75

AI in mosquito surveillance delivers real-time data crucial for policymakers, equips pest control businesses with precise information to target interventions, and empowers citizens with proactive monitoring tools. This approach improves response times and accuracy in decision-making. Risks include dependency on user participation, which may limit data coverage, and AI's limitations in identifying mosquito species lacking unique features across diverse scenarios.

The efficacy of machine learning, particularly deep learning, is heavily reliant on extensive training datasets. However, the development of methods to detect and mitigate model biases remains an important area for future research. To create these datasets, it is crucial to have robust repositories for data collection, organization, and sharing, equipped with user-friendly annotation tools and precise version control. While cloud platforms often provide these features, they may require subscriptions for private access, or offer free but public access that might not be ideal in early research stages. Therefore, careful consideration is required when selecting a suitable platform for data management and dissemination.

4.7 EERIE

European Eddy-RIch ESMs

Answers provided by Hannah Christensen (University of Oxford, Co-Investigator), Co-lead project WP12

EERIE will develop a new generation of Earth System Models that explicitly represent a crucially important part of the Earth system, the ocean mesoscale. We are using ML tools to address the associated technological challenges including: speeding up models, developing new parametrisation schemes for this resolution, and analysing the vast quantities of data produced. Different tools are used for each problem. E.g., deep graph neural networks represent the multiscale variability of the atmosphere and are therefore able to cheaply emulate its evolution when coupled to an ocean model. In contrast causal networks are easily interpretable and are used to reveal the role of the ocean in the climate system.

Compared to traditional tools, which tend to result in incremental improvements, ML models can lead to step-change advances in each of these areas. For example, it is difficult to parametrise convection at the 10 km scale because it is a grey-scale process at this resolution: it is partially resolved and partially sub-grid. Instead of taking existing parametrisations as a starting point, which were not designed for this resolution, ML approaches can learn optimal parametrisation schemes from higher resolution simulations. This can lead to substantial improvements in the fidelity of our models.

EERIE European Eddy-RIch ESMs

COORDINATING ENTITY:
ALFRED-WEGENER-INSTITUT HELMHOLTZ-ZENTRUM FUR POLAR- UND MEERESFORSCHUNG (GERMANY)

PARTICIPATING COUNTRIES:
Germany, Spain, South Africa, UK, France, Belgium, Cameroon, Netherlands, Switzerland

PROJECT WEBSITE:
https://eerie-project.eu/

START/END DATE:
1 January 2023 / 31 December 2026

EU contribution (€):
7 789 004,48

In EERIE we are also using ML to bridge the gap between Earth System Model output and the information that is needed by end users by exploiting Integrated Assessment Models (IAMs). IAMs are essential tools for quantifying and predicting interactions and feedback between the human and Earth system in areas including energy production, agriculture, and policy decisions. However, the representation of the physical climate system in IAMs is less detailed compared to ESMs. We will use ML tools to emulate the richness of behaviour observed in EERIE simulations, and embed this into an IAM, allowing for improved feedbacks between climate variability/extremes and human behaviour.

For optimal progress in this area, we urgently need a new generation of scientists versed in both AI and environmental science. These scientists will be able to use the most sophisticated of ML approaches, yet will also have the scientific know-how to ask the right question of the algorithms in the right way. This latter requirement is key when using ML tools. Only with this "natural intelligence" can artificial intelligence give useful results.

4.8 EIFFEL

Revealing the Role of GEOSS as the default digital Portal for building Climate Change Adaptation & Mitigation Applications

Answers provided by Dr Angelos Amditis (ICCS, Research and Development Director), Project Coordinator

In EIFFEL, AI is used in cognitive search tools, Explainable AI (XAI) models, and stochastic modelling for spatiotemporal data enhancement/augmentation. AI-based cognitive tools improve data discoverability via Neural Language Processing (NLP), outperforming traditional keyword searches. XAI ensures transparency, aiding climate adaptation decisions. Stochastic modelling enhances data resolution and quality, such as improving Sentinel-2 data from 20-60m to 10m. These methods provide accurate insights, handling large datasets efficiently, aiding decision-making with user-friendly interfaces.

AI is crucial for EIFFEL, enhancing data discoverability, quality, and resolution. Cognitive search tools surpass traditional methods by interpreting natural language queries. Stochastic modelling and super resolution improve spatiotemporal data accuracy, vital for climate actions. XAI ensures transparency, building stakeholder trust. These advances enable informed decision-making, pushing beyond state-of-the-art geospatial data processing, providing detailed, reliable insights crucial for climate change mitigation.

Benefits:

- Policymakers: AI tools offer high-resolution, timely data for precise climate policy-making and effective adaptation strategies.
- Businesses: Enhanced data accuracy and predictive capabilities support better decision-making in sectors, leading to the development of new cutting-edge added value services.
- Citizens: Improved environmental monitoring systems and access to climate services to provide early warnings for natural disasters, contributing to enhanced public safety and health.

Risks:

- Data privacy and security concerns.
- Complexity for non-technical users.
- Dependence on high-quality input data.
- Resource-intensive AI model development and maintenance, especially during training phase.

Research needs include developing efficient, scalable AI algorithms, advancing XAI for better decision transparency, integrating AI with emerging technologies like IoT and edge computing, and establishing robust data privacy frameworks. Interdisciplinary collaborations can drive innovation, addressing complex environmental challenges effectively and continuing to advance AI/ML applications in Earth Observation and climate science.

EIFFEL Revealing the Role of GEOSS as the default digital Portal for building Climate Change Adaptation & Mitigation Applications

COORDINATING ENTITY:
EREVNITIKO PANEPISTIMIAKO INSTITOUTO SYSTIMATON EPIKOINONION KAI YPOLOGISTON (GREECE)

PROJECT WEBSITE:
https://www.eiffel4climate.eu/

EU contribution (€):
4 999 466,25

PARTICIPATING COUNTRIES:
Greece, Spain, Netherlands, UK, Lithuania, Switzerland, Finland, Germany

START/END DATE:
1 June 2021 / 31 May 2024

4.9 EO4EU

AI-augmented ecosystem for Earth Observation data accessibility with Extended reality User Interfaces for Service and data exploitation

Answers provided by Prof. Stathes Hadjiefthymiades (National and Kapodistrian University of Athens), Project Coordinator and Dr. Vasileios Baousis (European Centre for Medium-Range Weather Forecasts), Technical Manager

EO4EU implements AI-empowered EO data processing workflows (from the EO sources all the way to advanced user interfaces). In this complex setup, AI is adopted in almost all discrete stages. The discovery and possible linking of information is facilitated through a dynamic Knowledge Graph. A wide spectrum of ML algorithms can be applied to the already selected datasets through a distributed Kubernetes environment (with information fusion as a notable example for the "blending" of heterogeneous EO- and in-situ data). Very specific ML tasks are integrated in the platform to facilitate the compression of exchanged information (learned compression) and the self-supervised learning to significantly reduce annotation effort. Inference is a key ingredient of the platform. Deployed models include CNNs, LSTMs and random forests.

Two of the objectives of the project were the significant reduction of annotation effort when constructing task specific supervised training datasets and the development of learnt compression methods for earth observation data. Both objectives have at their core ML methods and models. To achieve the first objective, we relied on self-supervised and contrastive learning methods. These are task agnostic, unsupervised learning methods and resulted in significant reductions of the required annotation effort for downstream tasks compared to classical supervised learning approaches.

EO4EU AI-augmented ecosystem for Earth Observation data accessibility with Extended reality User Interfaces for Service and data exploitation

COORDINATING ENTITY:
ETHNIKO KAI KAPODISTRIAKO PANEPISTIMIO ATHINON (GREECE)

PROJECT WEBSITE:
https://www.eo4eu.eu/

EU contribution (€):
8 177 925,00

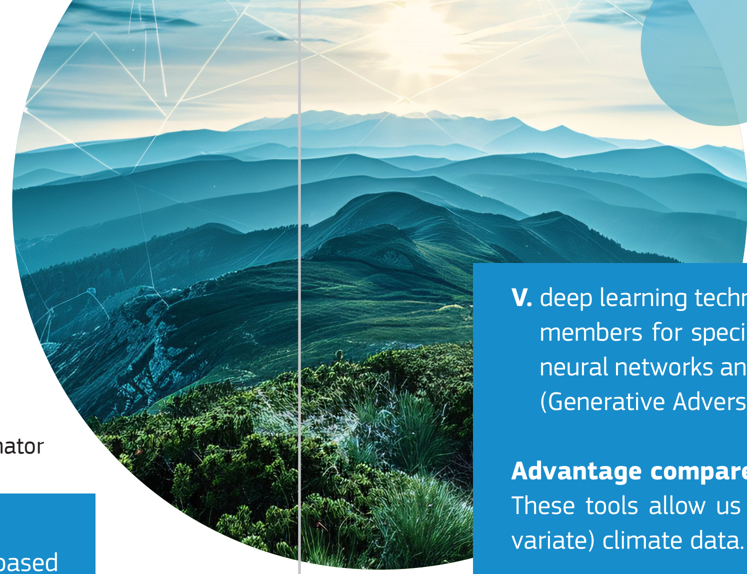
PARTICIPATING COUNTRIES:
Greece, UK, Italy, Lithuania, Latvia, Finland, Austria, Cyprus, Germany, Switzerland

START/END DATE:
1 June 2022 / 31 May 2025

With respect to the learnt compression, through the respective models (which are tuned to the distributions of processed data) we achieved better compression rates and better visual reconstruction compared to conventional compression standards (e.g., JPEG).

The platform itself simplifies the development of EO applications by abstracting all the processing steps (incl. AI ingredients) and allowing the interested user to easily specify workflows that would finally fuel existing or form entirely new applications. Therefore, the platform, by design, entails significant benefits to a very wide audience (this is also made clear through the extensive list of Use Cases that the project will implement). The abstract and generic character of the platform is also inherent in the integrated ML components. Foundational Model building is among the key tasks that the platform can achieve to reduce the annotation, along with learnt compression.

Research needs:
Leveraging of a multitude of infrastructures likes HPC for the efficient training of a wide diversity of models. Federated Learning.



4.10 EXPECT

Towards an Integrated Capability to Explain and Predict Regional Climate Changes

Answers provided by Markus Donat (Barcelona Supercomputing Center), Project Coordinator

EXPECT is using AI to:

I. improve the quality and spatiotemporal coverage of observations-based datasets of climate extremes
II. develop a ML-infused data assimilation scheme that provides more accurate and more temporally homogeneous initial conditions for hindcast simulations.
III. identify and quantify drivers of extremes
IV. process-based evaluations of models against observations
V. generate synthetic high-resolution climate model ensembles

To achieve this we use a variety of tools and methods:

I,II. For the reconstruction of incomplete observational data and initial states for climate prediction we use ML-based image pinpointing techniques, based on convolutional neural networks (CNN) and diffusion models

III,IV. We identify feature importance using different XAI methods such as permutation importance or SHAP values that provide a degree of explainability to the ML model. These will be expanded by other local model inspection techniques, e.g. layer-wise relevance propagation, which will help to reveal the most relevant regional patterns based on which the ML model grounded its prediction. To evaluate the data-driven results against physical domain knowledge, we will use causal inference methods to disentangle the mechanisms of how detected features are causally linked to the target variable.

EXPECT Towards an Integrated Capability to Explain and Predict Regional Climate Changes

COORDINATING ENTITY:
BARCELONA SUPERCOMPUTING CENTER
CENTRO NACIONAL DE SUPERCOMPUTACION
(SPAIN)

EU contribution (€):
6 648 370,00

PARTICIPATING COUNTRIES:
Spain, Germany, UK, Netherlands, Belgium, Italy, Portugal, Canada

START/END DATE:
1 April 2024 / 31 March 2028

V. deep learning techniques will be used to learn to generate synthetic ensemble members for specific dates of climate trajectories based on generative deep neural networks and long climate integrations with small, km-scale ensembles (Generative Adversarial Networks, and U-Net architectures)

Advantage compared to more traditional methods:
These tools allow us to learn and exploit empirical relationships in the (multi-variate) climate data.

AI allows us to learn and exploit empirical relationships in and across different climate variables, learnt from both observations and climate model data. This provides independent approaches to either traditional statistical methods or numerical climate model experiments (depending on specific task/application), which all have their specific shortcomings.

These applications ultimately contribute to
- Better understanding of extreme climate events based on more accurate datasets
- More accurate climate predictions based on better estimates of initial states
- Understanding the processes driving extremes, which enables more accurate predictions and projections of extremes

This more accurate knowledge of past and future climate including climate extremes is the basis for credible attribution and predictions of climate and related hazards, which is of benefit to any decision-makers dealing with climate-related risks and opportunities.

A major challenge remains the meaningful adoption of advanced AI methods to the climate science domain, as it requires deep knowledge in different disciplines (i.e. data science and earth/climate science), which is still rare expertise when depth in both is required (often AI experts still lack sufficiently deep understanding of the earth system).
Another major challenge remains the use of meaningful benchmarks and tools, documentation and protocols that ensure full reproducibility of research results.

4.11 FAIRICUBE

F.A.I.R. Information Cubes

Answers provided by Kathi Schleidt (EPSIT, Italy), Stefan Jetschny (NILU, Norway), Project coordination team

Within FAIRiCUBE, we are investigating how AI can be used to complement existing analysis pipelines. However, most FAIRiCUBE analysis tasks are more at the level of machine learning (ML), as scientific work requires dependable results.

That said, we do use AI for certain aspects of our work, e.g.:
- Gap-Filling in large datasets
- Automated clustering of cities or regions with similar characteristics
- Automatic recognition of building types from aerial or satellite images.

Generally, ML methods applied within the FAIRiCUBE use case work can identify and describe relationships between environmental data for the prediction of a desired output where there is no numerical or empirical relationship that can be formulated. In one FAIRiCUBE use case, the prediction results achieved by AI are far better than those achieved by conventional means, e.g. when performing gap filling, the data generated by AI was far more accurate than by conventional approaches. In other cases, ML/AI approaches are more efficient in providing the required outputs.

FAIRICUBE F.A.I.R. Information Cubes

COORDINATING ENTITY:
STIFTELSEN NILU (NORWAY)

PROJECT WEBSITE:
https://fairicube.eu/

EU contribution (€):
3 202 843,75

PARTICIPATING COUNTRIES:
Norway, Luxembourg, Austria, Germany, Netherlands, Italy, Spain

START/END DATE:
1 July 2022 / 30 June 2025

The more efficient and in-depth analysis of complex socio-environmental processes are leading to better information for stakeholders, as well as enabling policy makers to make well founded decisions taking many and complex aspects into account. ML aided gap filling techniques in particular can increase the data completeness which in return improves the stability, reliability, and meaningfulness of the analysis of input data. ML based clustering exercises scale well and provide multi-parameter.

There is clear focus withing the FAIRiCUBE work to balance accuracy and explainability during the process of selecting the appropriate ML method. More work is needed to enable explainable AI, i.e. providing researchers with methods and technologies to better understand the ML responses provided. That can imply to quantify explainability and to derive user-friendly metrics from the ML model to give insights in the decision processes of FAIRiCUBE ML applications.
Better metadata on available datasets, making it possible to understand potential bias ensuing from the data collection/creation process.

4.12 HARMONIA

Development of a Support System for Improved Resilience and Sustainable Urban areas to cope with Climate Change and Extreme Events based on GEOSS and Advanced Modelling Tools

Answers provided by Alexandros Psychas (ICCS Institute of Communication and Computer Systems, Greece), senior research engineer

In the HARMONIA project, AI is utilized in several key areas: analyzing satellite and in-situ data, predicting climate impacts, and assessing urban resilience. Specifically, we use machine learning techniques such as convolutional neural networks (CNNs) for image processing and deep learning models for data integration and analysis. Tools like data cubes optimize the storage and querying of large datasets. Compared to traditional methods, AI allows us to handle vast amounts of data more efficiently, providing quicker, more accurate predictions and enabling detailed, real-time monitoring of urban environments.

AI is essential for the HARMONIA project because it significantly enhances our ability to analyse and interpret large volumes of complex data from various sources. Traditional methods often fall short in managing the scale and diversity of data we deal with. AI enables real-time data processing, integrates satellite and in-situ data seamlessly, and offers high-precision models for predicting climate impacts. By doing so, AI helps us move beyond conventional techniques, allowing for a more dynamic and comprehensive understanding of urban resilience and supporting more effective climate adaptation and mitigation strategies.

HARMONIA Development of a Support System for Improved Resilience and Sustainable Urban areas to cope with Climate Change and Extreme Events based on GEOSS and Advanced Modelling Tools

COORDINATING ENTITY:
POLITECNICO DI MILANO (ITALY)

PROJECT WEBSITE:
https://harmonia-project.eu/

EU contribution (€):
4 999 794,00

PARTICIPATING COUNTRIES:
Italy, Greece, Switzerland, Poland, Romania, Sweden, Bulgaria, Finland, UK, Belgium

START/END DATE:
1 June 2021 / 31 January 2025

The HARMONIA project provides several benefits for end users. Policy makers gain access to detailed assessments of urban resilience, which aid in crafting effective climate adaptation policies. Businesses can identify risk areas, improving their strategic planning and investment decisions. Citizens benefit from enhanced urban planning and reduced risks associated with climate impacts, leading to safer and more sustainable living environments. However, potential risks include concerns about data privacy and the need for regular updates to AI models to maintain accuracy. Limitations could arise from the variability in data quality and availability.

Future research in this domain should focus on several areas: developing more sophisticated AI algorithms to improve prediction accuracy for climate impacts, creating better data integration techniques to handle diverse datasets efficiently, and enhancing real-time data processing capabilities. There is also a need for robust frameworks to ensure data privacy and security and methods to support continuous learning and adaptation of AI models as new data becomes available. Additionally, exploring ways to make AI tools more accessible and user-friendly for non-experts could significantly broaden their application and impact.

4.13 OEMC

Open-Earth-Monitor Cyberinfrastructure

Answers provided by Tom Hengl (Director OpenGeoHub foundation), project coordinator

The study by Masolele et al. demonstrated AI's potential in pinpointing cocoa farm locations across the tropics using deep learning and Sentinel-1 and Sentinel-2 satellite data. That work has also been expanded to all of Africa's many deforestation related commodity conversions. Hackländer et al. show how AI technology (primarily Machine Learning on time-series of EO images) can be used to detect significant land degradation areas using the global MODIS 250 m 10 day FAPAR time-series (covering 2000–2022). Such large volumes of data would be otherwise difficult (even impossible) to process without optimized Machine Learning frameworks running on High Performance Computing centres. Processing of the whole MODIS dataset (about 6 TB of data) took on the end only a few days, which means that we can easily update and improve these outputs. Bonnanella et al., used machine learning to map critical forest species probability shifts in key European forest species. To summarize, OEMC project uses AI technology / ML: (1) to process large volumes of EO and produce value-added predictions of key environmental variables (automated predictive mapping frameworks), (2) to extract / digitize information from high resolution satellite images (e.g. extract farm borders or location of trees, landscape patterns), (3) to simulate Earth Systems and predict states of vegetation, ecosystems under different climate scenarios, (4) to help integrate in-situ and EO data and summarize data without losing any significant information.

OEMC Open-Earth-Monitor Cyberinfrastructure

COORDINATING ENTITY:
STICHTING OPENGEOHUB (NETHERLANDS)

PROJECT WEBSITE:
https://earthmonitor.org/

EU contribution (€):
12 720 045,00

PARTICIPATING COUNTRIES:
Netherlands, Austria, Germany, Serbia, Italy, Slovenia, Croatia, Romania, Spain, Switzerland

START/END DATE:
1 June 2022 / 31 May 2026

Optimized Machine Learning enables processing of vast data volumes, unveiling complex relationships crucial for ecosystem and climate monitoring. The Open Earth Monitor Cyberinfrastructure project fosters open, transparent, and reproducible frameworks, accelerating environmental data utilization. With machine learning, large-scale tasks are completed faster, enhancing result accuracy, and reducing human errors. AI techniques integrate in-situ and satellite data for comprehensive environmental assessments, such as forest/land use and carbon stocks monitoring.

Benefits: terabytes of geospatial data for informed decision-making at reduced costs can be generated by fast, cost-efficient modelling processes with large EO data. Open access enables millions of people to monitor environmental changes, aiding national responses. Nations lacking data can identify land ownership; insurers track disasters; farmers validate deforestation-free production.
Risks: Managing terabytes of complex data from multiple mapping missions poses challenges of comprehension. Users may find it overwhelmed by 4–5 parallel global land cover products, and confused of the data source and pattern meaning. For users having multiple sources of the same type of data can become confusing (which one should I use?) or even overwhelming. In addition, with automated mapping systems (as with LLM systems) users do not always understand where the data come from? Where does some pattern come from and what does it mean? Some strategies to reduce such obscurity of complex data include reproducible research / computational notebooks published via Github/Gitlab and similar, and quality flags including prediction errors provided "per pixel", so that any user is aware at any time of data sources and risks associated with any decision.

One important AI research need is the need to invest in reference/training data and training procedures to make best use of them. Another key area of evolving research needs is for combining key datasets (satellites, in-situ, drones etc.) to underpin open-source foundation models for different key land surfaces processes (i.e. land use/management, carbon stocks and changes, biodiversity). There is today a clear need for public registers to have standard examples of EO images to represent different land cover classes, degradation states, events and similar.

4.14 OneAquaHealth

Protecting urban aquatic ecosystems to promote One Health

Answers provided by Maria João Feio (University of Coimbra), project coordinator

In OneAquaHealth AI/Machine learning supports the development of an Environmental Surveillance System, focus on guaranteeing the conservation of freshwater ecosystems (urban streams) inside cities/urbanized areas, as well as the health and wellbeing of citizens regarding potential diseases derived from the water through animals and plants (vector borne or water borne diseases). More specifically we will develop AI/models that are able to predict risks for human health from the degradation of freshwater ecosystems caused by urbanization and to determine early warning indicators. The early warning indicators will be displayed in maps in the OneAquaHealth Hub. In addition to ground data, indices derived from satellite imagery will be used as predictors of risks for human health in our AI/ML models.
ML models can extract complex features, providing valuable insights into environmental processes and trends.

Using advanced models like Artificial Neural Networks, the OneAquaHealth will can forecast the impact of urbanization on freshwater ecosystems and predict health risks associated with ecosystem changes. These novel AI methods will lead to improved problem solving and scalability as well as enhanced decision support with simulations and scenario analysis, to improved early warning systems. Furthermore, techniques like Transfer Learning can leverage pre-trained models on similar tasks or datasets to improve performance in areas where data may be sparse.

By doing so, we could enrich the areas where we currently have limited ground measurements, increasing the level of confidence in the model's predictions. Lastly, Ensemble Methods, which combine multiple models, can improve predictive performance and insights into urban freshwater ecosystems.

The use of AI in OneAquaHealth offers numerous benefits to end users. The predictive models enable early intervention, reducing the impact of environmental degradation.

- For policymakers, it provides real-time data and predictive insights about urban freshwater ecosystems, supporting the development of environmental surveillance systems.
- Businesses, especially in sectors like tourism and water management, can use this enhanced environmental data for risk mitigation and sustainable practice implementation.
- Citizens benefit directly from improved water quality and ecosystem health with AI-powered tools and applications that directly improve public engagement and awareness.

Research needs:

1. Data fusion / multimodal datasets: Research is needed to explore methods for integrating data from multiple sources, including satellite imagery, ground-based measurements (and climate models). This involves developing algorithms for data fusion and assimilation to improve the accuracy and reliability of environmental predictions and assessments.
2. Uncertainty quantification: Addressing uncertainties in Earth Observation data and model predictions is crucial for decision-making and policy formulation. Research is needed to develop methods for quantifying and propagating uncertainties through AI/ML models, as well as techniques for visualizing and communicating uncertainty
3. Interpretability and Explainability: As AI/ML models become more complex, there is a growing need for methods to interpret and explain model predictions. Research is needed to develop techniques for model explainability in the context of Earth Observation, allowing users to understand the factors driving model outputs and assess their reliability.

ONEAQUAHEALTH Protecting urban aquatic ecosystems to promote One Health

COORDINATING ENTITY:
UNIVERSIDADE DE COIMBRA (PORTUGAL)

PROJECT WEBSITE:
https://www.oneaquahealth.eu/

EU contribution (€):
4 939 558,00

PARTICIPATING COUNTRIES:
Portugal, France, Norway, Greece, Austria, Spain, Italy, Israel, Belgium, Switzerland

START/END DATE:
1 January 2023 / 31 December 2026

4.15 PREVENT

Improved Predictability of Extremes over the Mediterranean From Seasonal to Decadal Timescales

Answers provided by Christina Anagnostopoulou (Department of Meteorology and Climatology, AUTH), Project Coordinator

PREVENT uses AI to understand why extreme weather events happen in the Mediterranean. By combining different Machine Learning methods, we aim to comprehend the causes and patterns behind these extreme events. Moreover, PREVENT's goal is to create new forecast Machine Learning models for extreme weather in the Mediterranean. These models will be used to improve how we prepare for and respond to extreme weather events in the future. The use of AI in PREVENT represents a significant advancement beyond traditional methods, providing us with powerful tools to better understand, predict, and ultimately mitigate the impacts of extreme weather events in the Mediterranean.

AI is crucial for achieving our project objectives enabling us to go beyond the current state-of-the-art methods in predicting extreme weather events in the Mediterranean. Traditional approaches often fall short in capturing the causality involved in these events. Causal discovery algorithms and machine learning can effectively analyse vast amounts of data to identify the physical drivers and pathways leading to extremes. It enhances our ability to validate dynamical models but also allows us to assess the impact of climate change on these drivers. By leveraging AI, more accurate seasonal forecast models are developed, ultimately helping us better anticipate and adapt to extreme conditions.

The outcomes of PREVENT offer significant benefits to diverse end users. Policy makers gain valuable insights into Mediterranean extreme weather drivers, enabling them to develop more targeted and effective policies for climate adaptation and disaster risk reduction. Enhanced forecasting benefits businesses, improving planning and risk management in vulnerable sectors like agriculture and tourism. Citizens benefit from increased preparedness and resilience, ensuring safety and well-being.
Potential Risks and limitations: Climate system complexity & model uncertainties may affect prediction accuracy. AI needs careful interpretation & validation for reliability.

Research needs in the intersection of AI/ML and Climatology include refining algorithms for better climate prediction, addressing data quality and quantity challenges, integrating AI into climate models effectively, and exploring AI's role in understanding complex climate phenomena like extreme events and long-term trends. Additionally, ethical, societal, and policy implications of using AI/ML in climatology should be explored, including issues related to data privacy, bias in algorithms, and the equitable distribution of climate information and resources.

PREVENT Improved Predictability of Extremes over the Mediterranean From Seasonal to Decadal Timescales

COORDINATING ENTITY:
ARISTOTELIO PANEPISTIMIO THESSALONIKIS (GREECE)

PARTICIPATING COUNTRIES:
Greece, Cyprus, Germany, France, Netherlands, Israel, Morocco

PROJECT WEBSITE:
https://preventmed-climate.eu/

START/END DATE:
1 October 2023 / 30 September 2026

EU contribution (€):
2 997 875,00

4.16 SOCIO-BEE

Wearables and droneS fOr CIty Socio-Environmental Observations and BEhavioral ChangE

Answers provided by project coordinator Anastasios Drosou and deputy coordinator Evangelos Kopsacheilis (Centre for Research & Technology Hellas)

Complex Machine & Deep Learning methods are used for:

1. AI-based Data Enhancement (Data Imputation, Data Fusion, Anomaly Detection)
2. Prediction Engine (Air Pollution Forecasting, Unmonitored regions pollution estimation)
3. Exploring user behaviour (Means of Transportation Extraction, Most followed paths extraction, User Clustering)

The above AI methods lead to higher quality, accuracy and performance of the overall platform, in comparison to traditional ones.

The outcomes of the AI methods help reach the project objectives as they will provide more engaging, useful and accurate data for the end users.
The state of the art for data enhancement, prediction and user behaviour exploration -when the project was introduced- were mainly based on "conventional" methods and the introduction of AI methods which are constantly fine-tuned based on real data provides a breakthrough in the state-of-the-art.

SOCIO-BEE Wearables and droneS fOr CIty Socio-Environmental Observations and BEhavioral ChangE

COORDINATING ENTITY:
ETHNIKO KENTRO EREVNAS KAI TECHNOLOGIKIS ANAPTYXIS (GREECE)

PARTICIPATING COUNTRIES:
Greece, Belgium, Spain, Netherland, Italy, Germany, Norway

PROJECT WEBSITE:
https://socio-bee.eu/

START/END DATE:
1 October 2021 / 30 September 2024

EU contribution (€):
4 999 858,91

The end users will benefit through the provision of more accurate and constantly up-to-date Air Quality data, resulting in a more personalised and impactful experience compared to "traditional" methods.
There are no potential risks as personal data are not exposed to the AI algorithms, neither do users get drastic or important decisions through the AI outcomes.

Research needs:
- Citizen training & awareness on usage and benefitting from the AI
- Comprehensive user interactions & revealing the principles of decision support from AI
- Setting the appropriate framework for Ethical & Legal concerns

4.17 SUSTUNTECH

Sustainable tuna fisheries through advanced earth observation technologies

Answers provided by Begoña Vila Taboada (Marine Instruments) Project Coordinator

ML is used to obtain the fuel oil consumption and species distribution models developed in the project. The first step is to choose the most appropriate environmental data from a large catalogue of compatible datasets, which is done by applying the Correlation-based Feature Selection algorithm to select the most relevant features. Once this step is complete, the selected subset of inputs is then fed into a series of ML algorithms with different configurations, to identify the one which performs best by fine-tuning them and comparing their predictions with the expected outputs.

ML techniques combined with powerful computers allow to speed up the process of selecting the most appropriate algorithm from those applicable to any given problem. In our case, ML has allowed us to select the most relevant environmental data from different sources with numerous variables, reproducing multiple times the training process by modifying certain algorithm parameters, until the best results were achieved.

SUSTUNTECH Sustainable tuna fisheries through advanced earth observation technologies

COORDINATING ENTITY:
MARINE INSTRUMENTS SA (SPAIN)

PROJECT WEBSITE:
https://www.sustuntech.eu/

EU contribution (€):
2 618 968,11

PARTICIPATING COUNTRIES:
Spain, Norway, UK, Italy, Germany

START/END DATE:
1 May 2020 / 30 April 2024

In general, AI and ML have the potential to accelerate the process of finding solutions to complicated problems that affect today's societies: policy makers can take decisions with more accurate information, businesses can plan for the future with less uncertainty, and citizens can be assured that those decisions are taken in the interests of all. But there are risks: a more transparent approach to these new technologies should be pursued, explaining the need and advantages of using them, with the aim of keeping the general public onboard and to avoid a potential reluctance in their uptake.

4.18 SYLVA

A SYstem for reaL-time obserVation of Aeroallergens

Answers provided by Res. Prof. Mikhail Sofiev (Finnish Meteorological Institute), Project Coordinator

SYLVA uses Machine Learning in three major directions

- "classical AI"-type neural networks are used as a part of bioaerosol recognition algorithms for the new in-situ devices, the core of the project
- Machine Learning algorithms of various kinds are used in bioaerosol recognition algorithms of lidars and satellites, as well as in model identification procedures. They are not based on neural networks since lower complexity and amount of data allow more transparent approaches
- Bioinformatics is the third line: a bioaerosol recognition algorithms from environmental DNA sequences

Multi-dimensionality of the problems is the main reason. Wherever possible, we prefer machine learning based on principles other than neural network – primarily because of a better transparency of the obtained solutions.

SYLVA is developing background for a new-generation bioaerosol monitoring instruments, which will open a new era in aerobiology, with high-quality and high-resolution data generated by automatic instruments in real time. For various stakeholders, that would mean drastically better ground for decisions and information support. Since the data will be open, business opportunities are also huge. Citizen will obtain the accurate and up-to-date information on biological components present in the atmosphere, which is an implementation of the "right to know" with regard to biological air quality.

SYLVA A SYstem for reaL-time obserVation of Aeroallergens

COORDINATING ENTITY:
ILMATIETEEN LAITOS (FINLAND)

PROJECT WEBSITE:
https://sylva.bioaerosol.eu/

EU contribution (€):
2 998 677,25

PARTICIPATING COUNTRIES:
Finland, Belgium. Germany, Spain, Norway, Lithuania, Italy, Serbia, Switzerland

START/END DATE:
1 January 2023 / 31 December 2026

Among the risks, an over-statement of capabilities of neural networks (the actual mechanism behind the buzzy "AI" slogan) is arguably the most-significant (a self-evident risk of abuse of this technology is well known and does not need to be repeated). One should understand that neural networks are just tools that register apparent correlations between multi-dimensional processes. They cannot figure out the difference between apparent correlations and actual dependencies, their predictions are supported by data exclusively within the ranges for which the data are available, etc.

ML capabilities in predicting atmospheric composition and weather are among important directions potentially capable of significant reshaping of the whole area of meteorology and air pollution research and services. I do not expect ML to replace physical models (this is impossible in principle), but I see it as a possibility to substantially boost the efficiency and capabilities of such models, potentially replacing them in some niches, where ML instruments prove their viability.

4.19 XAIDA

Extreme Events: Artificial Intelligence for Detection And Attribution

Answers provided by Pascal Yiou (IPSL) and Melinda Galfi (VU Amsterdam), Project Co-coordinators

XAIDA focuses on the use of AI for extreme event attribution. The XAIDA consortium includes two partners who are specialized in AI research. AI is used in extreme event attribution through:

- identification of factors that cause extremes (e.g. causal inference)
- phase space reduction for data analysis and data mining (e.g. Variational Auto-Encoders)
- design of climate emulators (e.g. circulation analogues) to simulate ensembles of extreme events
- assistance to attribution reports with generative systems (e.g. ClimateQA, chatGPT)

The main advantage of AI is that it can learn complex relationships from the data, which would be very hard or even impossible to detect otherwise.

AI is at the core of the research lead in XAIDA. AI appears in all steps of extreme event attribution, from data mining, statistical inference and report writing. The goal of XAIDA research is to design and integrate objective tools of event attribution that are agile (and can be applied systematically and quickly) but scientifically rigorous. AI helps achieving those goals.
The use of generative AI was not envisaged at the beginning of XAIDA. It now strongly helps in synthetising the information and contextualising events (e.g. collaboration with ClimateQA).

XAIDA Extreme Events: Artificial Intelligence for Detection And Attribution

COORDINATING ENTITY:
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS (FRANCE)

PARTICIPATING COUNTRIES:
France, Netherlands, United Kingdom, Germany, Spain, Switzerland

PROJECT WEBSITE:
https://xaida.eu

START/END DATE:
1 September 2021 / 31 August 2025

EU contribution (€):
5 999 241,75

The main benefits are the relative ease of use of extreme event attribution tools and the extremely quick response after (and sometimes before) an event occurs. This greatly facilitates efficient communication with stakeholders, within hours to days, rather than weeks to months.
The main obvious limitation is that AI uses available knowledge based on the training data. It always provides results, but these have to be validated by human expertise.

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct centres. You can find the address of the centre nearest you online (european-union.europa.eu/contact-eu/meet-us_en).

On the phone or in writing

Europe Direct is a service that answers your questions about the European Union.
You can contact this service:

– by freephone: **00 800 6 7 8 9 10 11** (certain operators may charge for these calls),

– at the following standard number: **+32 22999696**,

– via the following form: european-union.europa.eu/contact-eu/write-us_en.

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website (european-union.europa.eu).

EU Publications

You can view or order EU publications at op.europa.eu/en/publications.
Multiple copies of free publications can be obtained by contacting Europe Direct or your local documentation centre (european-union.europa.eu/contact-eu/meet-us_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex (eur-lex.europa.eu).

EU open data

The portal data.europa.eu provides access to open datasets from the EU institutions, bodies and agencies. These can be downloaded and reused for free, for both commercial and non-commercial purposes. The portal also provides access to a wealth of datasets from European countries.

Artificial Intelligence technology plays an important role in harnessing the capacity of Earth Observations to support evidence-based decisions and improve sustainable performance in many economic sectors.
In this publication, nineteen Horizon projects that use artificial intelligence technology in various research tasks, express their views on what is offered by AI, how they use the different AI tools, what are the benefits, the risks and the limitations. Finally, they express the research needed to facilitate the uptake and use, as well as maximize the benefit, of AI in research.

Studies and reports

Publications Office
of the European Union