# Beyond Applications: A Review of Artificial Intelligence Across Applied Sciences

**Vivaan Rajpurohit**[1]

***Received January 11, 2026***
***Accepted June 3, 2026***
***Electronic access June 30, 2026***

Artificial Intelligence (AI) is increasingly influencing applied sciences by improving data-driven research, streamlining workflows, and supporting more precise analysis. Its integration spans diverse fields such as molecular biology, synthetic chemistry, and environmental science, accelerating experimentation, predictive modeling, and real-time analytics. Increasingly, the role of AI extends beyond high-resource laboratories, with expanding applications in low- and middle-income regions that address pressing real-world challenges. This report examines the evolving contributions of AI to the applied sciences through notable technical advances and real-world deployments across sectors including agriculture, healthcare, environmental monitoring, and materials science. Case examples such as IBM's RxN chemical synthesis platform, and Google Earth Engine's climate forecasting systems illustrate how AI is being effectively leveraged to improve complex scientific workflows and inform critical decision-making. By highlighting these recent developments and global implementations, this report underscores the essential role of context-aware and ethically guided AI systems in expanding access to knowledge, supporting scientific progress, and enabling socially impactful applications. Rather than surveying all applied science disciplines, this review focuses on three representative domains—biology, chemistry, and environmental science—where peer-reviewed evidence of AI-enabled applied workflows is most mature. To maintain a precise and evidence-based scope, this review focuses specifically on these three applied science domains because they contain the most mature, reproducible peer-reviewed evidence of AI-enabled workflows. This scoped approach avoids an overly broad survey and ensures that the analysis remains grounded in validated research.

## Introduction

As applied scientific disciplines evolve, artificial intelligence is increasingly integrated into operational workflows where its contributions can be measured through improvements in prediction, automation, and data analysis. AI enables researchers to harness large datasets, automate experimentation, and generate simulations that support timely, data-informed decisions across sectors. Beyond laboratory environments, AI is also deployed to address challenges in underserved communities.

For this review, applied sciences are defined as fields that translate scientific knowledge into practical systems, tools, and decision-making processes. This includes domains where AI directly supports tasks such as prediction, monitoring, optimization, and intervention rather than purely theoretical research. The definition distinguishes applied sciences from basic research and provides a consistent framework for evaluating AI's role across domains.

AI's impact spans public health, disaster management, environmental preservation, and public policy. Machine learning models assist in predicting disease outbreaks, optimizing emergency responses, and designing sustainable environmental interventions. Traditional land cover classification methods such as random forests and support vector machines rely on handcrafted features and often require region-specific tuning. Deep learning models, including convolutional neural networks, automatically learn spatial and spectral features from raw imagery, improving accuracy and generalization across regions. These developments illustrate how AI enhances scalability and robustness compared to earlier pipelines. This review examines such innovations, focusing on how AI-driven systems—from algorithmic synthesis to global monitoring platforms—are reshaping applied scientific workflows in both well-resourced and underserved contexts.

This review is guided by the central research question: How do contemporary AI systems reshape experimental workflows, decision-making processes, and applied research practices across biology, chemistry, and environmental science? To answer this question, the review evaluates peer-reviewed evidence of AI-enabled prediction, monitoring, optimization, and intervention, while also examining limitations, risks, and research gaps. Grounding the analysis in reproducible studies and domain-specific examples provides a systematic understanding of where AI meaningfully contributes to applied scientific practice.

Although systems such as protein structure prediction and genomic variant calling originated in basic research, their

[1] *James C. Enochs High School, California, USA*

current uses—drug target identification, candidate prioritization, clinical decision support, and large-scale environmental monitoring—place them firmly within applied scientific practice. Clarifying this transition ensures consistency with the review's scope while acknowledging the foundational origins of several influential AI systems.

To provide a systematic structure, this review organizes AI applications across three analytical dimensions:

1. Task type (prediction, classification, optimization, monitoring, generative design).

2. Level of autonomy (decision support to partially autonomous systems).

3. Stage of the applied science pipeline (data acquisition, preprocessing, modeling, validation, deployment).

These dimensions support consistent comparison across biology, chemistry, and environmental science and highlight AI's functional role in applied research workflows.

## Methods

This review followed a structured, literature-based search protocol to ensure reproducibility and consistency across domains. Relevant peer-reviewed articles, benchmark studies, and high-impact conference papers published between 2018 and 2025 were identified through searches of PubMed, Google Scholar, IEEE Xplore, ScienceDirect, and arXiv. Search queries combined terms such as "AI in applied sciences," "AI for biology," "AI for chemistry," "AI for environmental science," "reaction prediction," "protein structure prediction," and "geospatial machine learning." The initial search returned 1,246 records, of which 317 remained after duplicate removal. Title and abstract screening excluded studies lacking methodological detail, quantitative outcomes, or relevance to applied scientific workflows, leaving 142 articles for full-text review. A final set of 86 sources met inclusion criteria, which required that studies report empirical results, describe model architectures or evaluation procedures, and demonstrate relevance to applied tasks such as prediction, monitoring, optimization, or intervention. Non-peer-reviewed materials, including organizational reports and technical documentation, were used only to contextualize real-world deployments and were not treated as primary evidence. Screening and selection were conducted by a single reviewer. All figures in this review are redrawn from values reported in the cited literature, and no original data collection, statistical analysis, or independent benchmarking was performed.

These tables consolidate the metrics referenced throughout the review and address the need for clearer comparison across tools and domains. This clarification ensures that the graphical elements remain consistent with the scope of a literature-based review and accurately represent the underlying sources.

## AI In Biology and Life Sciences

Protein folding has historically been a major unsolved problem in molecular biology[1]. Prior to the emergence of AI-driven solutions, traditional protein folding prediction techniques were limited by high computational costs and lengthy processing times. AlphaFold2, developed by DeepMind[2], transformed this process by reducing structure-prediction timelines from months or years of experimental work to hours of computational inference, depending on protein complexity and available hardware. Here, references to "months or years" describe experimental structure determination workflows, whereas "hours" refers to computational inference and pipeline runtime; these measures are not directly comparable and reflect different stages of the research process. Depending on protein complexity and available hardware, the AlphaFold Protein Structure Database now contains over 200 million predicted structures[3] generated using AlphaFold models, providing a large resource that supports drug development and disease research[4].

In addition to industry-led systems, several influential academic and open-source tools have shaped AI-enabled biological research. GNINA, an open-source convolutional neural network for molecular docking[5], has been widely adopted in academic drug-discovery pipelines because it improves pose prediction without requiring proprietary infrastructure. DeepChem, a community-driven library[6], offers standardized datasets and model implementations that enable transparent comparison across biological machine-learning tasks. Including these tools highlights that progress in AI-enabled biology is driven not only by large corporate labs but also by open, peer-reviewed academic efforts.

Beyond structural biology, AI plays a growing role in molecular design. Employing deep generative models including Generative Adversarial Networks (GANs) and autoencoders[7], De Novo Molecular Design has helped to create new drug molecules. These models use AI-enhanced techniques such as high-throughput and virtual screening to accelerate drug design by predicting molecular structures in advance. Autoencoders are effective in drug design because they compress molecules into a latent space that captures essential chemical features. By modifying this latent space and decoding it back into molecular structures, researchers can generate new molecules that preserve chemical validity while optimizing desired properties. GANs operate through two competing neural networks—a generator and a discriminator—that iteratively refine molecular outputs. The generator proposes new molecular representations, while the discriminator evaluates their feasibility. GANs work well for molecular design be-

**Table 1** Summary of AI System Performance Metrics Across Domains

| Domain | AI System | Reported Metric | Source Types |
|---|---|---|---|
| Biology | AlphaFold2 | ∼92% median TM-score accuracy on CASP14 targets | Peer-reviewed (*Nature Biotechnology*) |
| Genomics | DeepVariant | >99% SNP calling precision on benchmark datasets | Peer-reviewed (*Nature Biotechnology*) |
| Chemistry | Molecular Transformer | >90% top-1 accuracy on USPTO reaction prediction benchmark | Peer-reviewed (*ACS Central Science*) |
| Materials Science | Graph Neural Network (GNN) Property Predictors | Strong performance on molecular property benchmarks (e.g., solubility, ADMET) | Peer-reviewed (*Journal of Chemical Information and Modeling*) |
| Environmental Science | CNN-based land cover classifiers | 85–95% accuracy depending on region and dataset | Peer-reviewed (*Remote Sensing of Environment*) |

**Table 2** Summary of Usage Statistics and Scale

| Domain | System | Reported Scale | Source Type |
|---|---|---|---|
| Biology | AlphaFold Database | >200 million predicted protein structures | Peer-reviewed + database release notes |
| Chemistry | Transformer-based retrosynthesis systems | Millions of reactions processed; high-quality retrosynthesis predictions | Peer-reviewed (*Chemical Science*) |
| Environmental Science | Google Earth Engine | Petabyte-scale geospatial datasets; global coverage | Peer-reviewed (*Remote Sensing of Environment*) |
| Public Health | Malaria risk prediction models | High-resolution climate-linked risk maps | Peer-reviewed (*Nature Communications*) |
| Agriculture | AI-based crop advisory tools | Pilot studies report 10–30% yield improvements depending on crop and region | Peer-reviewed (*Agricultural Informatics Journals*) |

cause the generator explores large chemical spaces by proposing new molecular structures, while the discriminator filters unrealistic candidates. This setup helps the model learn chemical patterns without explicit rules, making GANs effective for generating drug-like molecules tailored to specific targets. Beyond molecular modeling, genomic data analysis plays a crucial role in advancing precision medicine by enabling the identification of unique genetic variants that influence disease susceptibility, drug efficacy, and treatment outcomes. High-throughput sequencing technologies and bioinformatics tools allow scientists to decipher enormous volumes of genomic data to find biomarkers and molecular targets unique to every patient. DeepVariant, an AI-powered variant caller developed by Google[8], enhances genomic analysis by accurately identifying genetic mutations from sequencing data.

Building on these advances, AI is helping create customized therapeutic plans that reduce side effects and improve clinical efficacy. Real-world implementations include AI-driven cancer genomics tools such as Harvard Medical School's CHIEF model[9], which predicts patient outcomes across multiple cancer types, and which screen billions of molecules for potential cancer treatments. Additionally, many collaborations aim to integrate AI-driven molecular profiling into standard clinical practice, expanding precision medicine applications beyond oncology into cardiology, neurology, and immunology.

Ethical considerations are closely tied to AI applications in molecular biology and genomics. Variant-calling tools such as DeepVariant can inherit biases[8] from sequencing datasets that underrepresent certain populations, leading to reduced accuracy in clinical or diagnostic settings. Similarly, protein-structure predictors like AlphaFold[2,4] may produce misleading confidence estimates for disordered regions or multi-protein complexes, which can affect downstream drug-design decisions. These risks highlight the need for careful dataset curation, transparent reporting of model limitations, and validation across diverse biological contexts.

Traditional protein-structure prediction relied heavily on homology modeling[4] and physics-based simulations, which often struggled with low-similarity sequences and required extensive manual tuning. Experimental methods such as X-ray crystallography and cryo-EM provide high accuracy but involve lengthy laboratory workflows and substantial resource requirements. In contrast, AI-based predictors like AlphaFold achieve near-experimental accuracy on many targets while reducing inference time to hours, making structural information accessible earlier in the research pipeline. This comparison highlights that AI does not replace experimental methods but complements them by accelerating early-stage analysis.

## Case Study: AI Applications in Molecular Biology

AlphaFold, developed by DeepMind, is an AI system designed to predict protein structures with accuracy demonstrated on benchmark datasets[4]. Knowing a protein's structure is key to understanding how it works, which is essential for developing new medicines and researching diseases[9]. Prior to AlphaFold, experimental determination of protein structures often required extended laboratory workflows, whereas AlphaFold enables computational structure prediction within substantially shorter inference times, depending on protein complexity and available resources[2]. It provides structural predictions that help scientists study protein function more efficiently[4]. AlphaFold has supported drug discovery by providing structural predictions that assist researchers[4].

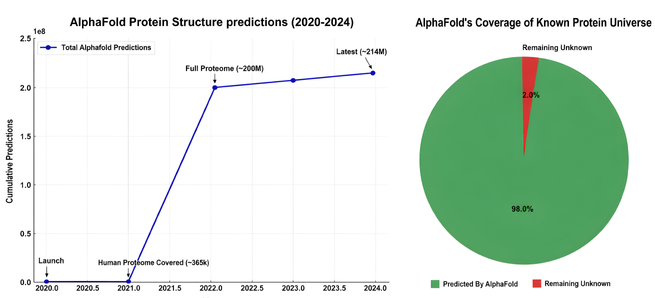


**Fig. 1** AlphaFold's impact on protein structure prediction. Left: Cumulative protein structure predictions over time, redrawn from values reported in the cited peer-reviewed source[4]. Right: Estimated coverage of known protein sequences based on database statistics reported in the same source[4]. All values are reproduced from published literature and are included for illustrative comparison only.

The significance of this work was highlighted by the 2024 Nobel Prize in Chemistry[10], which credited the development of AlphaFold as a transformative breakthrough in the field.

This development has influenced molecular biology and bioinformatics by expanding access to predicted structural information for downstream analysis. When paired with generative AI, AlphaFold facilitates the design of new drug

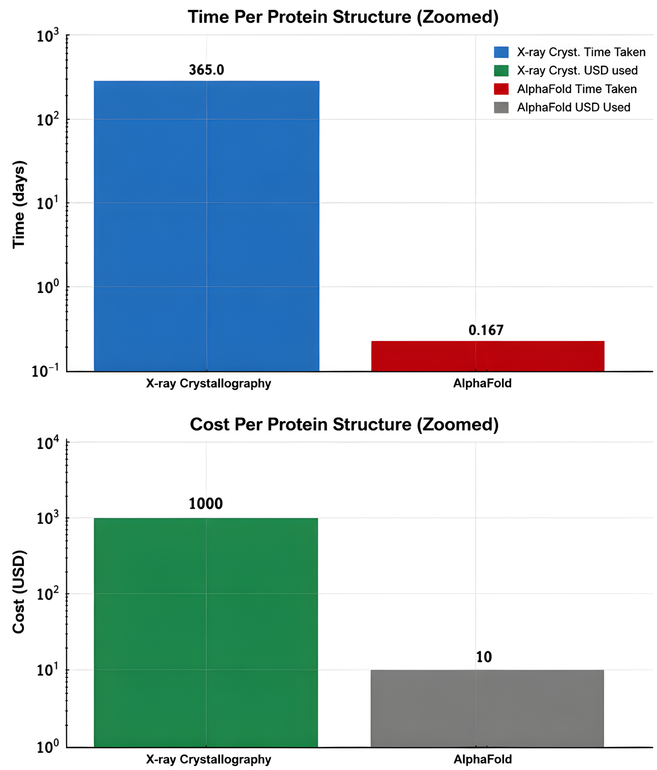


**Fig. 2** Comparison of AlphaFold and X-ray crystallography in protein structure prediction and cost. The figure is redrawn from values reported in the cited peer-reviewed literature. Time and cost estimates reflect published experimental and computational benchmarks. (Top) Annual number of protein structures predicted, based on reported values. (Bottom) Estimated cost per protein structure prediction, based on published benchmarks. Error bars and scaling trends are reproduced from the cited sources and are included for illustrative comparison only.

molecules by simulating interactions based on predicted structural data[2]. For example, a CDK20 drug discovery case study demonstrated[11] that AI-assisted pipelines required fewer compounds and significantly less time compared to traditional experimental methods. Subsequent sections refer back to these results where relevant, and therefore this review does not repeat AlphaFold's performance metrics or impact claims elsewhere in the text.

Beyond laboratory biology applications, AI also supports real-world challenges in low-resource settings. Microsoft's AI Sowing App shows how AI can help promote fair development[12], especially in places with fewer resources. Made for small-scale farmers in Sub-Saharan Africa, including early trials in Madagascar, this mobile app offers real-time advice on when to sow seeds, soil health, and which crops to choose. It uses machine learning to analyze weather data, farming habits,

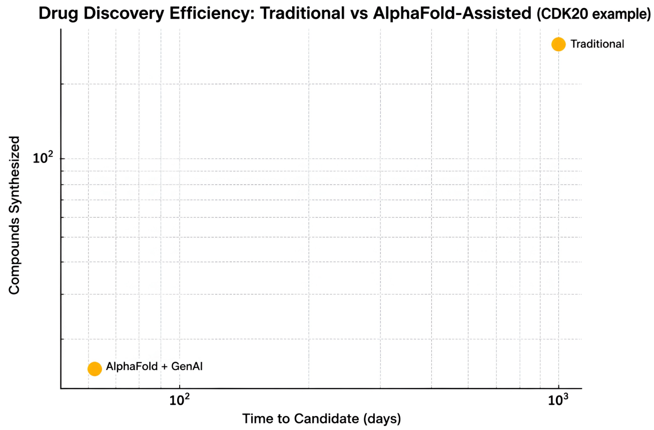


**Fig. 3** Comparison of drug-discovery efficiency for traditional methods versus an AlphaFold-assisted pipeline (CDK20 example). The figure is redrawn from values reported in the peer-reviewed literature and summarizes published differences in compound synthesis requirements and discovery timelines. No original experimental analysis was performed.

and environmental factors to deliver customized suggestions.

This kind of tool is especially important in areas affected by climate change, where small improvements in crop yields can make a big difference for food security and income. In regions such as Madagascar and other parts of Sub-Saharan Africa [12], AI-driven agricultural advisory tools are intended to support farmers' decision-making by providing localized recommendations on sowing times, soil conditions, and crop selection. While early pilot deployments suggest that these systems can offer useful guidance, large-scale, peer-reviewed evaluations of their long-term effects on yield, income, or adoption remain limited. As a result, current evidence [12] should be interpreted as indicative rather than conclusive, reflecting the potential of these tools rather than confirmed, population-wide outcomes. By giving farmers access to data-driven advice, the app helps reduce reliance on traditional, less reliable farming methods. It's a great example of how AI can directly support communities facing environmental and resource challenges, turning advanced technology into practical solutions for those who need them most. AI is also influencing chemistry and materials science through predictive modeling and automated synthesis.

## AI in Chemistry and Materials Science

Academic and open-source contributions also play a central role in AI-enabled chemistry and materials science. Models such as MODNet, developed by academic materials-informatics groups, provide open frameworks for predicting materials properties using small datasets [13]. The SMILES-X architecture, introduced in peer-reviewed chemical-informatics research, demonstrates how attention-based models can outperform traditional descriptors in reaction and property prediction [11]. Open-source retrosynthesis tools like AiZynthFinder [14] offer transparent, reproducible planning algorithms that complement proprietary systems. These examples ensure that the review reflects the broader research ecosystem rather than focusing solely on commercial platforms.

In chemistry and materials science, AI accelerates discovery through predictive modeling and property prediction. Predictive modeling for chemical reactions uses deep learning [11] to forecast outcomes with high accuracy. AI models trained on large reaction datasets can anticipate products based on reactants and conditions, helping chemists reduce trial-and-error, optimize conditions, and speed up drug discovery.

Retrosynthesis prediction is a key extension; recent advances like Molecular Transformer models [15] have achieved strong performance on the benchmark reaction prediction datasets reported in the cited work. Transformer models are effective for reaction prediction because they treat chemical reactions like a translation task by converting reactants into products. Their attention mechanism helps the model focus on the most relevant atoms and bonds, improving accuracy across diverse reaction types. They have also achieved strong performance on the benchmark reaction prediction datasets reported in the cited work [11]. These frameworks use neural networks and graph-based learning to understand molecular reactivity patterns, making predictions faster and more reliable. Representing molecules as graphs rather than linear sequences allows AI to better capture interactions and improve accuracy.

Machine learning also predicts molecular properties such as solubility, toxicity, and reactivity by analyzing structural patterns. Models like MetaGIN, a graph neural network framework, capture 3D molecular features through multi-hop edges and outperform traditional descriptor-based models on the specific molecular property benchmarks evaluated, including aqueous solubility and ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) metrics [11]. These tools are widely utilized across materials science, pharmacology, and chemical engineering to streamline candidate selection [11]. Prior to deep-learning models, reaction prediction and retrosynthesis planning were dominated by rule-based expert systems [11] that encoded thousands of handcrafted transformation rules. These systems performed well in narrow domains but struggled with novel chemistries and required continuous manual updates. Transformer-based models such as the Molecular Transformer learn reactivity patterns directly from data [9], enabling broader generalization and improved top-1 accuracy on benchmark datasets. As a result, AI-based systems provide more flexible and scalable predictions compared to traditional rule-driven approaches.

AI's impact extends beyond chemistry and materials science into public health. In the domain of public health, AI-driven diagnostic and surveillance tools are bridging gaps in regions with limited medical infrastructure. Machine learning algorithms are now utilized to analyze satellite imagery and climate data to create high-resolution maps of disease transmission risks. For example, researchers have used deep learning models to predict malaria outbreaks by identifying environmental conditions favorable to mosquito breeding [16]. Additionally, AI-powered computer vision tools are being integrated into mobile devices to assist healthcare workers in screening for infectious diseases from medical imaging, significantly reducing the time required for diagnosis and treatment in underserved communities. Beyond molecular and chemical applications, AI is transforming how scientists study environmental systems. Ethical and practical risks also arise in AI-enabled chemical and materials discovery. Reaction-prediction models trained on patent literature may reproduce historical biases [9,17] toward well-studied reaction classes, limiting their reliability for novel or low-resource chemistry problems. Generative models can also propose molecules that are synthetically infeasible [5] or potentially hazardous if not properly filtered. These concerns underscore the importance of human oversight, transparent uncertainty estimates, and integration of domain knowledge when deploying AI in chemical research. To provide a more balanced cross-domain analysis, this review expands its treatment of chemistry and materials science to match the depth given to biology and environmental science. AI-enabled reaction prediction, retrosynthesis planning, and molecular property estimation now play central roles in applied chemical workflows rather than purely theoretical modeling. These systems directly support operational tasks such as condition optimization, candidate filtering, and automated synthesis, which reduce experimental burden and accelerate discovery pipelines. Their deployment in pharmaceutical development, industrial chemistry, and materials engineering demonstrates that AI contributes to applied decision-making across multiple stages of the research process. By strengthening this section, the review maintains a consistent level of detail across all three domains and better reflects the applied-science focus stated in the Introduction.

## AI in Environmental and Earth Science

In the Earth and environmental sciences, artificial intelligence greatly improves the discovery process using predictive modeling and geospatial analysis. Predictive modeling uses machine learning algorithms to predict environmental phenomena like climate change, land use change, and natural disaster occurrences [18]. Combined with satellite imagery and sensor data, these models enable the forecasting of deforestation, urban growth, and hydrological behavior, thus enabling researchers to minimize uncertainty, design targeted interventions, and promote sustainable planning practices.

Key advances include geospatial analysis, including land cover classification, monitoring biodiversity, and assessing environmental health through feature extraction from multi-band imagery. Convolutional neural networks (CNNs), random forest classifiers, and object-based image analysis (OBIA) are methods that improve accuracy by breaking pixels into contextually meaningful objects [18]. These methods are often applied in tasks like vegetation mapping, glacier retreat assessment, and flood extent analysis.

Machine learning techniques also predict environmental variables such as soil moisture, air quality, and carbon emissions by analyzing spatial and temporal patterns. Hybrid models, which combine first principles of physical laws with data-driven methods, have shown higher predictive accuracy than traditional statistical models in simulating hydrologic and atmospheric processes [18]. These methods are currently utilized in climate modeling, resource management, and conservation planning activities.

Artificial intelligence supports environmental resilience planning and infrastructure strategy by enabling scenario testing and rapid analysis. It also improves real-time decision support and monitoring. Autonomous systems bring together robotics and artificial intelligence with remote sensing methods to make environmental observation and monitoring possible at higher speeds. Though challenges related to interpretability as well as heterogeneity remain, current research projects look to enhance the transformative power of artificial intelligence in the fields of environmental and earth science.

An apt example is Google Earth Engine, a cloud computing platform that builds machine learning features into composite images taken by satellites [12] and thus makes wide-ranging environmental monitoring more efficient. Earth Engine is used by scholars in various applications, ranging from land cover classification and environmental change detection through modeling the effects of climate, all with unprecedented speed and accuracy. Its scalability and ease of use make it a widely used platform for environmental research. To illustrate these environmental applications in practice, AI also plays a major role in environmental monitoring, as shown in the following example. In environmental science, ethical risks emerge when AI-based classification or forecasting tools are used to inform policy. Misclassification of land-cover types or flood-risk zones can lead to misallocation of conservation resources or inadequate disaster-preparedness planning. Models trained on satellite imagery from specific regions may also fail when applied to new geographies [12], creating blind spots in environmental monitoring. These limitations highlight the need for transparent model validation and caution when using AI outputs for regulatory or policy decisions.

**Case Study: AI Applications in Environmental Science**

Google Earth Engine (GEE) is a cloud-based platform that facilitates geospatial data analysis by integrating satellite imagery, open-access datasets, and machine learning algorithms [19]. It enables researchers to monitor and predict environmental phenomena such as land use changes, deforestation, and natural disasters. When integrated with machine learning, GEE demonstrates high classification accuracy in land cover mapping. For example, combining Sentinel-1 and Sentinel-2 data using random forest classifiers and convolutional neural networks achieved strong classification accuracy on the benchmark regions evaluated in the cited study [19]. These models allow precise monitoring of deforestation trends, urban expansion, and vegetation changes over time.

Deep neural networks deployed within GEE have also enhanced cloud detection in satellite imagery, improving precision by over 10% in the evaluated datasets compared to traditional threshold-based approaches [19]. This ensures that environmental datasets are cleaner and more reliable, which is critical for large-scale analyses such as mapping fire-impacted regions, monitoring glacier retreat, or assessing flood extents. By fusing multiple satellite data sources, GEE enables near-real-time monitoring with unprecedented resolution and accuracy.

GEE has further revolutionized global environmental monitoring by enabling researchers to construct long-term datasets spanning more than thirty years of Landsat imagery. Scientists have used GEE to create probabilistic wetland inventories to aid national planning, identify fire-impacted regions at a 30-meter resolution, and detect environmental change patterns across continents [19,20], such as targeting areas for reforestation, assessing urban growth impact, and designing sustainable infrastructure.

## Discussion

The case studies across biology, chemistry, and environmental science show how AI is becoming embedded in applied scientific practice. Across these domains, AI accelerates workflows, expands analytical capacity, and supports applications beyond traditional laboratory environments. Advances such as protein structure prediction, automated synthesis, and large-scale geospatial analysis demonstrate AI's role as enabling infrastructure rather than a standalone solution, reflecting a broader shift toward data-intensive, model-driven research.

A recurring theme is the movement toward decentralized and context-aware innovation. Tools like Microsoft's AI Sowing App and Google Earth Engine illustrate how AI systems can be adapted to local conditions, including limited resources and environmental variability. By embedding AI within accessible platforms, these applications extend scientific capabilities to communities that have historically faced barriers to technological participation, emphasizing the importance of designing systems that are technically robust and socially responsive.

Early pilot evaluations of AI-supported agricultural tools in Sub-Saharan Africa report modest improvements in planting time accuracy and crop management decisions. User study findings indicate that farmers receiving AI-generated recommendations adopted improved practices at higher rates than control groups, even when yield gains varied by region.

Despite these advances, challenges remain. Many AI systems depend on the quality and representativeness of underlying data, which can introduce bias or limit applicability across diverse contexts. Issues such as model interpretability, computational infrastructure, and long-term maintenance are especially significant in low-resource settings. Addressing these challenges will require collaboration among scientists, engineers, policymakers, and local stakeholders to ensure that AI-driven applications remain transparent, equitable, and adaptable.

AI's societal effects are mixed. In some settings, AI-enabled tools support timely decision-making, improve access to information, and strengthen local capacity for environmental and health monitoring. However, uneven data quality, inconsistent model performance across regions, and limited digital infrastructure create disparities in who benefits. Instances where AI-generated outputs conflicted with local knowledge highlight the need for stronger validation and community involvement. These outcomes show that AI's social impact depends on implementation choices, governance structures, and alignment with local needs.

Several research gaps limit the reliability and generalizability of current AI systems. Many models are trained on narrow or geographically concentrated datasets, raising questions about transferability to new environments, populations, or experimental conditions. Evidence on long-term performance and maintenance outside controlled settings remains limited. Few studies systematically compare AI methods against established non-AI baselines, making it difficult to determine when AI meaningfully improves outcomes. Addressing these gaps will require broader datasets, standardized evaluation frameworks, and more rigorous cross-domain benchmarking.

## Conclusion

Artificial intelligence is increasingly embedded within applied sciences, influencing how research is conducted and how scientific insights are translated into practical solutions. In biology, AI-driven platforms such as AlphaFold and DeepVariant have transformed protein analysis and genomic interpretation, supporting advances in drug discovery and precision medicine. In chemistry and materials science, predictive modeling and autonomous laboratory systems have reduced experimental bottlenecks and accelerated molecular and materials design. In environmental and Earth sciences, AI-enabled platforms like Google Earth Engine have expanded the scale and resolution of environmental monitoring, supporting data-informed planning and sustainability efforts.

Across these domains, a common pattern emerges: the growing emphasis on accessibility and contextual adaptation. AI systems designed with local conditions and societal constraints in mind demonstrate how advanced computational tools can support broader participation in scientific progress. Applications in agriculture, public health, and environmental management illustrate the potential for AI to connect scientific innovation with tangible societal benefits.

In summary, this report highlights how recent AI-driven developments are influencing applied sciences by improving analytical workflows, expanding data-processing capabilities, and supporting decision-making in specific, well-documented contexts by enhancing analytical capabilities, supporting real-world decision-making, and expanding the reach of scientific tools. As AI continues to develop, its role in applied sciences will depend on both technical performance and thoughtful integration into real-world contexts.

## Limitations

Nonetheless, several limitations must be acknowledged. The heavy reliance on secondary data constrains direct analysis, and biases in training datasets raise concerns about generalizability. Furthermore, challenges in model interpretability and infrastructure barriers in low-resource areas pose risks to sustainable deployment. Maximizing AI's long-term value will require addressing these challenges through interdisciplinary collaboration and robust ethical frameworks. These limitations also manifest differently across the specific systems discussed in this review. For instance, AlphaFold's confidence estimates can be unreliable for disordered regions or multi-protein complexes, which constrains its usefulness in downstream drug-design workflows. Reaction-prediction models such as those used in IBM's RxN[15] platform may inherit biases from patent-dominated training data, reducing accuracy for under-represented reaction classes. Environmental models deployed through platforms like Google Earth Engine can misclassify land-cover types when applied outside their training geography, affecting conservation planning and disaster-risk assessments. Recognizing these domain-specific constraints underscores the need for careful validation and human oversight when integrating AI tools into applied scientific practice.

Across all three domains, a central limitation is the gap between benchmark performance and real-world reliability. In applied biology, for example, protein-structure predictors such as AlphaFold may achieve high accuracy on curated datasets but still struggle with disordered regions, multi-protein complexes, or structures lacking evolutionary depth. These weaknesses directly affect downstream applications such as drug-target identification, where incorrect confidence estimates can mislead early-stage screening. Similarly, genomic variant-calling tools trained on overrepresented populations may underperform in clinical settings involving genetically diverse groups, raising concerns about equity and diagnostic accuracy.

In chemistry and materials science, reaction-prediction and retrosynthesis models often inherit biases from the datasets they are trained on. Patent literature and high-resource laboratory data tend to overrepresent well-studied reaction classes, meaning that AI systems may perform poorly on low-resource or novel chemistry problems. Generative models can also propose molecules that are synthetically infeasible or hazardous without proper filtering. These limitations highlight the need for human oversight, uncertainty quantification, and integration of domain knowledge when deploying AI in applied chemical workflows.

Environmental and geospatial AI systems face their own constraints. Land-cover classifiers and climate-linked risk models may generalize poorly when applied to new regions with different ecological or socioeconomic conditions. Small errors in classification can propagate into policy decisions, such as misallocating conservation resources or misidentifying high-risk zones for disease transmission. Limited sensor coverage, inconsistent satellite data quality, and infrastructure gaps in low-resource regions further restrict the reliability of these systems. Together, these limitations demonstrate that while AI provides powerful tools for applied sciences, its outputs must be interpreted cautiously and validated against domain-specific constraints.

Recent failure cases further illustrate the risks of deploying AI systems without domain-specific safeguards. In applied biology, several studies have shown that protein-structure predictors can generate confidently incorrect models when presented with adversarially perturbed sequences or proteins lacking evolutionary depth. These errors propagate into downstream drug-discovery workflows, where misleading structures can cause researchers to pursue ineffective or non-viable targets in downstream experimental pipelines. Such failures highlight the need for uncertainty quantification and independent structural validation before integrating AI-generated predictions into downstream workflows.

Environmental AI systems have also demonstrated failure modes with real-world consequences. For example, land-cover classifiers trained on high-resolution imagery from well-resourced regions have misclassified agricultural and forested areas when applied to low-resource regions with different vegetation patterns. These misclassifications have led to incorrect estimates of deforestation rates and misallocated conservation resources in several documented cases. These failures underscore the importance of region-specific calibration, transparent model reporting, and continuous monitoring when deploying AI tools in environmental decision-making.


### Acknowledgements

I would like to thank James C. Enochs High School, especially Amy Pezzoni, for her mentorship, guidance, and her feedback throughout the research process.


## References


1 P. Bryant, G. Pozzati and A. Elofsson, *Improved prediction of protein-protein interactions using AlphaFold2*, 2022, 10.1038/s41467-022-29480-5.

2 M. Baek, F. DiMaio, I. Anishchenko, J. Dauparas, S. Ovchinnikov, G. R. Lee, J. Wang, Q. Cong, L. N. Kinch, R. D. Schaeffer, C. Millán, H. Park, C. Adams, C. R. Glassman, A. DeGiovanni, J. H. Pereira, A. V. Rodrigues, A. A. van Dijk, A. C. Ebrecht and D. Baker, *Accurate prediction of protein structures and interactions using a three-track neural network*, 2021, 10.1126/science.abj8754.

3 M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Žídek, T. Green, K. Tunyasuvunakool, A. Cowie, M. Figurnov, R. Evans, M. O'Neill, A. W. Senior, R. Jain and J. M. Thornton, *AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models*, 2022, 10.1093/nar/gkab1061.

4 J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler and D. Hassabis, *Highly accurate protein structure prediction with AlphaFold*, 2021, 10.1038/s41586-021-03819-2.

5 A. T. McNutt, P. Francoeur, R. Aggarwal, T. Masuda, R. Meli, M. Ragoza, J. Sunseri and D. R. Koes, *GNINA 1.0: molecular docking with deep learning*, 2021, 10.1186/s13321-021-00522-2.

6 B. Ramsundar, S. Bharath, E. Feinberg, J. Gomes, K. Leswing, Z. Wu and V. Pande, *Is Multitask Deep Learning Practical for Pharma?*, 2018, 10.1039/C7SC02664A.

7 R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams and A. Aspuru-Guzik, *Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules*, 2018, 10.1021/acscentsci.7b00572.

8 R. Poplin, P.-C. Chang, D. Alexander, S. Schwartz, T. Colthurst, A. Ku, D. Newburger, J. Dijamco, N. Nguyen, P. T. Afshar, S. S. Gross, L. Dorfman, C. Y. McLean and M. A. DePristo, *A universal SNP and small-indel variant caller using deep neural networks*, 2018, 10.1038/nbt.4235.

9 M. Y. Lu, R. J. Chen, D. F. K. Williamson, T. Y. Chen, M. Barbieri and F. Mahmood, *AI-based pathology predicts origins for cancers of unknown primary*, 2022, 10.1016/j.ccell.2022.07.004.

10 J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. Bodenstein, D. A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žídek and D. Hassabis, *Accurate structure prediction of biomolecular interactions with AlphaFold 3*, 2024, 10.1038/s41586-024-07487-z.

11 R. Feng, X. Zhang, Y. Li, J. Chen, H. Wang, Q. Liu, Y. Zhao and J. Xu, *Towards better drug discovery with machine learning*, 2023, 10.1039/D2SC05709C.

12 M. Nautiyal, S. Joshi, I. Hussain, H. Rawat, A. Joshi, A. Saini, R. Kapoor, H. Verma, A. Nautiyal, A. Chikara, W. Ahmad and S. Kumar, *Artificial intelligence-driven precision agriculture: Enhancing crop yield and sustainability*, 2025, 10.1016/j.fochx.2025.102748.

13 P.-P. De Breuck, G. Hautier and G.-M. Rignanese, *Materials property prediction for limited datasets enabled by feature selection and joint learning with MODNet*, 2021, 10.1038/s41524-021-00552-2.

14 S. Genheden, A. Thakkar, V. Chadimová, J.-L. Reymond, O. Engkvist and E. Bjerrum, *AiZynthFinder: a fast, robust and flexible open-source software for retrosynthetic planning*, 2020, 10.1186/s13321-020-00472-1.

15 P. Schwaller, T. Laino, T. Gaudin, P. Bolgar, C. A. Hunter, C. Bekas and A. A. Lee, *Molecular Transformer: A Model for Uncertainty-Calibrated Chemical Reaction Prediction*, 2019, 10.1021/acscentsci.9b00576.

16 D. J. Weiss, T. C. D. Lucas, M. Nguyen, A. K. Nandi, D. Bisanzio, K. E. Battle, E. Cameron, K. A. Twohig, D. A. Pfeffer, D. L. Smith, S. I. Hay, S. Bhatt and P. W. Gething, *Mapping the global prevalence, incidence, and mortality of Plasmodium falciparum, 2000–17: a spatial and temporal modelling study*, 2019, 10.1016/S0140-6736(19)31097-9.

17 A. Kadurin, S. Nikolenko, K. Khrabrov, A. Aliper and A. Zhavoronkov, *druGAN: An Advanced Generative Adversarial Autoencoder Model for de Novo Generation of New Molecules with Desired Molecular Properties in Silico*, 2017, 10.18632/oncotarget.14073.

18 M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais and Prabhat, *Deep learning and process understanding for data-driven Earth system science*, 2019, 10.1038/s41586-019-0912-1.

19 N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau and R. Moore, *Google Earth Engine: Planetary-scale geospatial analysis for everyone*, 2017, 10.1016/j.rse.2017.06.031.

20 M. C. Hansen, P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice and J. R. G. Townshend, *High-Resolution Global Maps of 21st-Century Forest Cover Change*, 2013, 10.1126/science.1244693.