EasyChair Preprint

№ 14815

GPU-Accelerated Machine Learning Models for Drug Discovery in Computational Biology

Abilly Elly

September 12, 2024

GPU-Accelerated Machine Learning Models for Drug Discovery in Computational Biology

Author

Abilly Elly

Date: September 11, 2024

Abstract

Drug discovery is a complex and time-consuming process that requires the identification of potential therapeutic compounds from vast chemical libraries. In recent years, the integration of machine learning (ML) with computational biology has revolutionized this field by enabling faster and more accurate predictions of drug-target interactions, molecular properties, and toxicity profiles. This paper explores the use of GPU-accelerated machine learning models in drug discovery, focusing on how the parallel processing capabilities of GPUs significantly reduce the computational burden associated with large-scale simulations and predictive modeling. Key advancements in deep learning architectures and generative models, such as convolutional neural networks (CNNs) and graph neural networks (GNNs), are discussed in the context of their application to molecular dynamics, protein-ligand binding, and virtual screening. Additionally, the study highlights the role of transfer learning and active learning strategies in enhancing model accuracy and adaptability in drug discovery workflows. Through case studies, this research demonstrates the potential of GPU-powered ML models to accelerate the identification of novel drug candidates, improve lead optimization, and ultimately shorten the drug development timeline.

Keywords: GPU acceleration, Drug discovery, Machine learning, Computational biology

Introduction

Drug discovery is the process of identifying new candidate medications based on the biological understanding of diseases and their molecular targets. This multifaceted process involves several stages, including target identification, lead compound discovery, preclinical testing, and clinical trials. The ultimate goal of drug discovery is to develop safe and effective therapies that can alleviate disease symptoms, treat underlying causes, or provide a cure. Given its crucial role in healthcare innovation, drug discovery has a direct impact on public health outcomes and the pharmaceutical industry, where reducing the time and cost of development is a priority.

The Role of Computational Biology in Drug Discovery

Computational biology has emerged as a key player in modern drug discovery by enabling the simulation and analysis of biological systems at the molecular level. It combines biological data, algorithms, and computational techniques to model interactions between drug candidates and

their molecular targets, such as proteins, enzymes, and receptors. By leveraging in silico methods such as molecular docking, molecular dynamics simulations, and bioinformatics tools, researchers can predict the behavior of drug candidates in a virtual environment before laboratory testing. This greatly accelerates the process of identifying promising compounds and reduces the need for costly and time-consuming experimental procedures.

Limitations of Traditional Drug Discovery Methods

Traditional drug discovery methods rely heavily on experimental approaches, such as high-throughput screening (HTS) and empirical testing of chemical libraries, which are resource-intensive and often slow. These methods require significant investments in laboratory infrastructure, skilled personnel, and time, with many potential drug candidates failing at later stages of development. Furthermore, traditional techniques are limited in their ability to handle the complexity of biological systems and the vast chemical space that needs to be explored, leading to low hit rates and high rates of failure in clinical trials. These inefficiencies highlight the need for more powerful computational tools that can streamline the discovery process.

GPU Acceleration and Its Potential Benefits

Graphics Processing Units (GPUs) have traditionally been used for rendering graphics in video games and visual applications. However, their parallel processing capabilities have made them highly suitable for handling large-scale computations in scientific research, particularly in machine learning (ML) and computational biology. GPU acceleration allows for the execution of thousands of simultaneous computations, significantly speeding up complex tasks such as molecular simulations, deep learning model training, and data analysis. By utilizing GPUs, machine learning models for drug discovery can process vast datasets more efficiently, enhancing the ability to predict drug-target interactions, optimize lead compounds, and conduct virtual screenings at a fraction of the time required by traditional methods. The introduction of GPU-accelerated machine learning thus promises to revolutionize drug discovery by enabling faster, more scalable, and more accurate predictions, ultimately leading to a reduction in drug development timelines and costs.

Machine Learning Techniques for Drug Discovery

In drug discovery, machine learning (ML) techniques provide powerful tools to analyze complex datasets, predict molecular behaviors, and design optimal drug candidates. These techniques are applied at various stages, from predicting chemical properties to optimizing drug design. Below are the key ML approaches used in the drug discovery process:

Supervised Learning

Supervised learning algorithms are highly effective in drug discovery when labeled data, such as experimental outcomes or known molecular properties, is available. These techniques are employed to predict specific molecular features or interactions:

- **Regression Models**
  Regression models, including linear regression and support vector regression (SVR), are

used to predict continuous properties of molecules such as solubility, toxicity, and pharmacokinetic behavior. These models allow researchers to quantify the relationship between molecular descriptors and desired drug attributes, facilitating early decision-making in the drug discovery pipeline.

- **Classification Models**
  Classification algorithms, such as random forests (RF) and gradient boosting machines (GBMs), are used to predict categorical outcomes, such as whether a molecule will exhibit biological activity or bind effectively to a target protein. These models help filter out inactive compounds early, focusing efforts on those more likely to succeed in preclinical testing.

Unsupervised Learning
Unsupervised learning techniques are invaluable for discovering hidden patterns in data, particularly when there are no predefined labels. These methods are often used to group molecules or identify relationships between biological targets:

- **Clustering Algorithms**
  Clustering algorithms, such as k-means and hierarchical clustering, are used to group similar chemical compounds or biological targets. This can help identify chemical families with similar biological activities or classify molecular structures based on their predicted behavior in a biological system.

- **Dimensionality Reduction Techniques**
  Techniques like Principal Component Analysis (PCA) and t-distributed Stochastic Neighbor Embedding (t-SNE) reduce the complexity of high-dimensional drug-related data, enabling better visualization and understanding of molecular relationships. These techniques are particularly useful for feature selection and exploratory data analysis in drug discovery.

Reinforcement Learning
Reinforcement learning (RL) is gaining traction in drug design as it allows the optimization of sequential decision-making processes. Unlike supervised learning, which relies on labeled data, RL algorithms learn by interacting with an environment and receiving feedback (rewards or penalties).

- **Deep Q-Networks (DQN)**
  Deep Q-networks combine deep learning and Q-learning to optimize drug design by predicting the best sequence of modifications to improve molecular properties. DQN can help optimize compounds for specific goals, such as improving binding affinity or reducing toxicity.

- **Policy Gradient Methods**
  Policy gradient methods are used to directly optimize the policy that controls how decisions are made in the drug design process. This approach is particularly useful in

optimizing the structure of drug candidates, as it can iteratively improve a molecule's properties based on feedback from virtual screening or molecular simulations.

GPU Acceleration in Machine Learning Models

In recent years, GPU (Graphics Processing Unit) acceleration has become a cornerstone of machine learning (ML) advancements, particularly in areas requiring large-scale computation such as drug discovery. GPUs offer substantial performance improvements over traditional CPU-based systems, thanks to their architecture and optimization for parallel processing tasks.

GPU Architecture and Parallel Processing

Differences Between CPUs and GPUs

The central difference between CPUs (Central Processing Units) and GPUs lies in their design. CPUs are designed for general-purpose computing and are optimized for single-threaded performance. They typically have fewer cores, each capable of handling a wide variety of tasks, which makes them suitable for sequential, complex operations. On the other hand, GPUs have thousands of smaller, simpler cores designed to perform parallel tasks efficiently. While a CPU might have a few dozen cores optimized for high-performance sequential execution, a GPU could have thousands of cores specialized in handling multiple tasks concurrently.

GPU Parallel Processing Optimization

GPUs are specifically optimized for parallel computations. In machine learning, particularly in tasks like matrix multiplication, image processing, or simulations, there is a high degree of parallelism, which makes GPUs an ideal choice. Each GPU core executes the same operation on different pieces of data simultaneously, following a SIMD (Single Instruction, Multiple Data) architecture. This architecture enables GPUs to process vast datasets quickly, dramatically speeding up the training and inference phases of machine learning models, such as those used in drug discovery.

GPU Libraries and Frameworks

To leverage the power of GPU acceleration, several specialized libraries and frameworks have been developed, enabling efficient integration with machine learning models:

- **CUDA (Compute Unified Device Architecture)**
  CUDA, developed by NVIDIA, is a parallel computing platform and application programming interface (API) that allows developers to use NVIDIA GPUs for general-purpose computing. CUDA provides direct access to the GPU's virtual instruction set and parallel computing elements, making it the backbone of many GPU-accelerated ML models.

- **OpenCL (Open Computing Language)**
  OpenCL is an open standard for cross-platform parallel programming of diverse processors, including GPUs. While it supports a broader range of hardware (from various

vendors), it is less specialized than CUDA, which is designed specifically for NVIDIA GPUs.

- **TensorFlow and PyTorch**
  TensorFlow (by Google) and PyTorch (by Facebook) are two of the most popular deep learning frameworks. Both libraries provide native support for GPU acceleration, allowing seamless execution of neural networks on GPUs. TensorFlow uses CUDA and cuDNN for GPU optimization, while PyTorch has similar integrations, offering flexibility and ease of use for both research and production environments.

Features and Benefits for Machine Learning
The primary benefit of these libraries is their ability to offload computationally intensive tasks to GPUs, thus speeding up training times for large models. CUDA and cuDNN (CUDA Deep Neural Network library) allow TensorFlow and PyTorch to run optimized neural network computations, such as matrix multiplications, convolutions, and other operations fundamental to deep learning. These libraries enable researchers to scale up machine learning experiments, facilitating more complex models and deeper networks in drug discovery.

Performance Optimization Techniques

To maximize the performance of GPU-accelerated machine learning models, several optimization techniques are employed:

- **Memory Management Strategies**
  Efficient memory management is crucial to prevent bottlenecks in GPU computing. Two key techniques include:
  - **Pinned Memory**: Pinned (or page-locked) memory allows faster data transfers between the CPU and GPU, as it prevents the operating system from swapping memory to disk. This leads to reduced latency in data movement.
  - **Unified Memory**: Unified memory provides a single memory address space accessible by both the CPU and GPU, simplifying memory management for developers. It reduces the complexity of manually copying data between the host and device, though it may be less efficient in terms of raw performance compared to explicitly managing memory transfers.
- **Kernel Optimization**
  GPU kernels, which are the functions executed on GPU cores, can be further optimized to improve performance:
  - **Loop Unrolling**: This technique involves unrolling loops within the kernel to reduce the overhead of loop control, thus enhancing the throughput of each thread.
  - **Thread Block Synchronization**: Ensuring that threads within a block are properly synchronized minimizes race conditions and ensures correct results

without sacrificing performance. Careful management of thread block size and organization improves resource utilization and execution speed.

- **Hardware Acceleration with cuDNN and cuBLAS**
  GPU-accelerated libraries such as cuDNN (CUDA Deep Neural Network library) and cuBLAS (CUDA Basic Linear Algebra Subroutines) provide highly optimized implementations of common deep learning and linear algebra operations. cuDNN accelerates the computation of deep neural networks, including convolutional layers and backpropagation, while cuBLAS accelerates matrix operations, crucial for machine learning models. These libraries are highly tuned for NVIDIA hardware, providing significant speedups for applications like drug discovery, where such operations are computationally intensive.

Applications of GPU-Accelerated Machine Learning in Drug Discovery

GPU-accelerated machine learning (ML) has revolutionized drug discovery by significantly reducing computation times, enabling the analysis of massive datasets, and enhancing prediction accuracy. These advancements have transformed several key areas of the drug development pipeline, from virtual screening to drug repurposing. Below are some critical applications of GPU-accelerated ML in drug discovery:

Virtual Screening

Large-Scale Screening of Compound Libraries

Virtual screening involves computationally evaluating large libraries of chemical compounds to identify those most likely to interact with a specific biological target. Traditional methods for virtual screening can be slow and resource-intensive. GPU acceleration enables the parallel processing of thousands of compounds simultaneously, significantly speeding up the process. Machine learning models trained on experimental data can quickly rank compounds based on their predicted biological activity, reducing the number of compounds that require in-lab testing.

Accelerating Docking Simulations and Scoring Functions

Molecular docking simulations, which predict how small molecules (ligands) bind to target proteins, are central to virtual screening. Scoring functions are used to estimate the binding affinity of ligands. GPU-accelerated ML models, such as deep learning-based scoring functions, can improve the accuracy and speed of docking simulations by quickly identifying compounds with high binding affinity. This acceleration allows researchers to screen larger chemical libraries in a fraction of the time.

De Novo Drug Design

Generating Novel Drug-Like Molecules

De novo drug design involves the generation of new molecular structures from scratch that have desirable therapeutic properties. Generative models, such as variational autoencoders (VAEs) and generative adversarial networks (GANs), are used to create novel drug-like molecules. GPU

acceleration enables these models to explore vast chemical spaces efficiently, generating diverse molecular candidates with optimized properties. These models can rapidly design molecules that meet specific criteria, such as binding affinity, solubility, or toxicity.

Optimizing Molecular Properties Using Generative Models

Once novel molecules are generated, their properties must be optimized to enhance their therapeutic potential. Reinforcement learning techniques, combined with generative models, allow for the iterative optimization of molecular structures. GPU-accelerated reinforcement learning can quickly explore possible molecular modifications, leading to compounds with improved pharmacological properties, such as higher potency or lower toxicity.

Protein-Ligand Interaction Prediction

Predicting Binding Affinity and Selectivity

Accurately predicting how well a drug molecule binds to a target protein is crucial in drug discovery. GPU-accelerated deep learning models, such as convolutional neural networks (CNNs) and graph neural networks (GNNs), are capable of predicting binding affinity and selectivity for potential drug candidates. These models can analyze the 3D structures of proteins and ligands and predict their interactions, helping researchers identify high-affinity drug candidates faster than traditional computational methods.

Identifying Potential Drug Targets

Beyond predicting interactions with known targets, ML models can help identify new, previously unknown drug targets. By analyzing large biological datasets, such as genomics and proteomics data, GPU-accelerated ML models can uncover patterns and relationships between proteins and diseases. This can lead to the discovery of novel therapeutic targets, opening up new avenues for drug development.

ADMET Prediction

Predicting Absorption, Distribution, Metabolism, Excretion, and Toxicity

ADMET properties (Absorption, Distribution, Metabolism, Excretion, and Toxicity) are critical factors in determining a drug's success in clinical trials. Predicting these properties early in the drug discovery process can prevent costly failures during later stages of development. GPU-accelerated ML models can analyze large datasets of chemical compounds and their associated ADMET profiles to predict how new drug candidates will behave in biological systems. These models can quickly assess whether a compound is likely to have favorable pharmacokinetic and toxicity properties.

Drug Repurposing

Identifying New Uses for Existing Drugs

Drug repurposing involves finding new therapeutic applications for already approved drugs, often by identifying new biological targets for these compounds. GPU-accelerated ML models can analyze vast amounts of biomedical data, such as gene expression profiles, chemical structures, and clinical trial outcomes, to identify potential new uses for existing drugs. This approach can significantly shorten the drug development timeline and reduce costs since

repurposed drugs have already been tested for safety in humans. For example, ML models have been used to identify drugs with potential efficacy against emerging diseases, such as COVID-19, by predicting their interactions with viral proteins.

Challenges and Future Directions in GPU-Accelerated Machine Learning for Drug Discovery

While GPU-accelerated machine learning (ML) has significantly advanced drug discovery, several challenges still hinder its full potential. These challenges range from data-related issues to technical constraints and the integration of computational methods with experimental validation. Below are key challenges and the potential future directions in overcoming them:

Data Quality and Quantity

Data Scarcity, Bias, and Noise

High-quality, diverse, and sufficiently large datasets are crucial for training robust ML models. However, in drug discovery, obtaining vast amounts of high-quality labeled data is a challenge. Many experimental datasets are either scarce or incomplete, leading to a high reliance on synthetic data or small datasets that may not generalize well to real-world scenarios. Moreover, datasets are often biased toward well-studied compounds or targets, leading to a lack of diversity in molecular structures or biological activities represented in the training data.

- **Future Direction**: To address data scarcity, new initiatives for sharing preclinical and clinical data in open-access repositories could help expand available datasets. Active learning techniques, transfer learning, and generative modeling could also mitigate data limitations by leveraging knowledge from related domains or generating synthetic data that enriches existing datasets.

Model Interpretability

Understanding the Decision-Making Process of Black-Box Models

Many of the most powerful ML models used in drug discovery, such as deep neural networks, are often regarded as "black boxes" because their decision-making processes are difficult to interpret. This lack of interpretability poses challenges in drug discovery, where researchers need to understand why a model predicted a particular compound as promising, especially when it comes to safety and efficacy predictions.

- **Future Direction**: The development of explainable AI (XAI) techniques is critical for improving model interpretability. Techniques such as attention mechanisms, feature importance scoring, and model-agnostic interpretability tools like SHAP (Shapley Additive Explanations) can help clarify how models reach decisions, allowing researchers to better trust and refine predictions.

Hardware and Software Limitations

Memory Constraints and Computational Bottlenecks

Although GPUs provide substantial computational power, they are not without limitations. Large-scale drug discovery tasks, such as simulating protein-ligand interactions or processing high-dimensional biological data, can run into memory constraints or require excessive computational resources. Additionally, parallelizing certain ML tasks effectively can be challenging, especially when working with very large datasets or complex models that require constant communication between the CPU and GPU.

- **Future Direction**: Advances in hardware, such as the development of more powerful GPUs with larger memory capacities and the introduction of specialized hardware like Tensor Processing Units (TPUs), can alleviate these bottlenecks. On the software side, optimizing ML frameworks and libraries for better memory management, as well as developing algorithms that distribute computations more efficiently, can further reduce hardware constraints.

Integration with Experimental Validation

Bridging the Gap Between In Silico and In Vitro Studies

While computational methods have become indispensable in early drug discovery, experimental validation remains critical. A major challenge is ensuring that predictions from in silico models align with real-world biological data from in vitro or in vivo experiments. Models can often be overfitted to computational datasets, leading to results that do not translate effectively into experimental settings.

- **Future Direction**: To bridge the gap between computational predictions and experimental validation, tighter integration of ML with high-throughput experimental techniques, such as automated drug screening and lab robotics, is needed. Hybrid approaches, where ML models guide experiments and experiments refine ML predictions in a feedback loop, can provide more reliable results. Additionally, efforts to improve the biological realism of computational models through better simulations of molecular interactions will enhance their predictive power.

Conclusion

GPU-accelerated machine learning (ML) has made transformative contributions to drug discovery, offering unprecedented speed and efficiency in tasks such as virtual screening, de novo drug design, protein-ligand interaction prediction, and ADMET profiling. By leveraging the parallel processing power of GPUs, ML models can handle large datasets, perform complex simulations, and predict molecular properties faster than traditional methods. These capabilities enable researchers to explore vast chemical spaces, design novel compounds, and optimize molecular properties with greater precision, drastically reducing the time and cost involved in drug development.

The potential impact on the pharmaceutical industry is immense. By accelerating early-stage drug discovery, GPU-powered ML can shorten the time from target identification to clinical trials, facilitating the development of novel therapies for a wide range of diseases. Furthermore, GPU-accelerated techniques in drug repurposing can bring new applications for existing drugs, helping address emerging health threats more efficiently. This shift could also lower research and development costs, making drug discovery more accessible to smaller biotech firms and research institutions, thus fostering innovation across the sector.

Looking ahead, future research must focus on overcoming key challenges, including improving data quality, addressing the interpretability of complex ML models, and optimizing hardware and software for large-scale drug discovery tasks. Additionally, integrating computational predictions with experimental validation will be crucial for ensuring that in silico findings translate into real-world therapeutic outcomes. As these challenges are addressed, GPU-accelerated machine learning is poised to play an even more pivotal role in revolutionizing the pharmaceutical industry.

References

1. Chowdhury, R. H. (2024). Advancing fraud detection through deep learning: A comprehensive review. *World Journal of Advanced Engineering Technology and Sciences*, *12*(2), 606-613.

2. Akash, T. R., Reza, J., & Alam, M. A. (2024). Evaluating financial risk management in corporation financial security systems. *World Journal of Advanced Research and Reviews*, *23*(1), 2203-2213.

3. Abdullayeva, S., & Maxmudova, Z. I. (2024). Application of Digital Technologies in Education. *American Journal of Language, Literacy and Learning in STEM Education* , *2* (4), 16-20.

4. Katheria, S., Darko, D. A., Kadhem, A. A., Nimje, P. P., Jain, B., & Rawat, R. (2022). Environmental Impact of Quantum Dots and Their Polymer Composites. In *Quantum Dots and Polymer Nanocomposites* (pp. 377-393). CRC Press

5. 209th ACS National Meeting. (1995). *Chemical & Engineering News, 73*(5), 41–73.

   https://doi.org/10.1021/cen-v073n005.p041

6. Chowdhury, R. H. (2024). Intelligent systems for healthcare diagnostics and treatment. *World Journal of Advanced Research and Reviews*, *23*(1), 007-015.

7. Zhubanova, S., Beissenov, R., & Goktas, Y. (2024). Learning Professional Terminology With AI-Based Tutors at Technical University.

8. Gumasta, P., Deshmukh, N. C., Kadhem, A. A., Katheria, S., Rawat, R., & Jain, B. (2023). Computational Approaches in Some Important Organometallic Catalysis Reaction. *Organometallic Compounds: Synthesis, Reactions, and Applications*, 375-407.

9. Bahnemann, D. W., & Robertson, P. K. (2015). Environmental Photochemistry Part III. In ˜ *The œhandbook of environmental chemistry*. https://doi.org/10.1007/978-3-662-46795-4

10. Chowdhury, R. H. (2024). The evolution of business operations: unleashing the potential of Artificial Intelligence, Machine Learning, and Blockchain. *World Journal of Advanced Research and Reviews*, *22*(3), 2135-2147.

11. Zhubanova, S., Agnur, K., & Dalelkhankyzy, D. G. (2020). Digital educational content in foreign language education. *Opción: Revista de Ciencias Humanas y Sociales* , (27), 17.

12. Oroumi, G., Kadhem, A. A., Salem, K. H., Dawi, E. A., Wais, A. M. H., & Salavati-Niasari, M. (2024). Auto-combustion synthesis and characterization of La2CrMnO6/g-C3N4 nanocomposites in the presence trimesic acid as organic fuel with enhanced photocatalytic activity towards removal of toxic contaminates. *Materials Science and Engineering: B*, *307*, 117532.

13. Baxendale, I. R., Braatz, R. D., Hodnett, B. K., Jensen, K. F., Johnson, M. D., Sharratt, P., Sherlock, J. P., & Florence, A. J. (2015). Achieving Continuous Manufacturing: Technologies and Approaches for Synthesis, Workup, and Isolation of Drug Substance May 20–21, 2014 Continuous Manufacturing Symposium. *Journal of Pharmaceutical Sciences*, *104*(3), 781–791. https://doi.org/10.1002/jps.24252

14. Chowdhury, R. H. (2024). AI-driven business analytics for operational efficiency. *World Journal of Advanced Engineering Technology and Sciences*, *12*(2), 535-543

15. Bakirova, G. P., Sultanova, M. S., & Zhubanova, Sh. A. (2023). AGYLSHYN TILIN YYRENUSHILERDIY YNTASY MEN YNTYMAKTASTYYN DIGITAL TECHNOLOGYALAR ARGYLY ARTTYRU. *News. Series: Educational Sciences* , *69* (2).

16. Parameswaranpillai, J., Das, P., & Ganguly, S. (Eds.). (2022). *Quantum Dots and Polymer Nanocomposites: Synthesis, Chemistry, and Applications*. CRC Press.

17. Brasseur, G., Cox, R., Hauglustaine, D., Isaksen, I., Lelieveld, J., Lister, D., Sausen, R., Schumann, U., Wahner, A., & Wiesen, P. (1998). European scientific assessment of the atmospheric effects of aircraft emissions. *Atmospheric Environment*, *32*(13), 2329–2418. https://doi.org/10.1016/s1352-2310(97)00486-x

18. Chowdhury, R. H. (2024). Blockchain and AI: Driving the future of data security and business intelligence. *World Journal of Advanced Research and Reviews*, *23*(1), 2559-2570.

19. Babaeva, I. A. (2023). FORMATION OF FOREIGN LANGUAGE RESEARCH COMPETENCE BY MEANS OF INTELLECTUAL MAP. *Composition of the editorial board and organizing committee* .

20. Ahirwar, R. C., Mehra, S., Reddy, S. M., Alshamsi, H. A., Kadhem, A. A., Karmankar, S. B., & Sharma, A. (2023). Progression of quantum dots confined polymeric systems for sensorics. *Polymers*, *15*(2), 405.

21. Chrysoulakis, N., Lopes, M., José, R. S., Grimmond, C. S. B., Jones, M. B., Magliulo, V., Klostermann, J. E., Synnefa, A., Mitraka, Z., Castro, E. A., González, A., Vogt, R., Vesala, T., Spano, D., Pigeon, G., Freer-Smith, P., Staszewski, T., Hodges, N., Mills, G., & Cartalis, C. (2013). Sustainable urban metabolism as a link between bio-physical sciences and urban planning: The BRIDGE project. *Landscape and Urban Planning*, *112*, 100–117. https://doi.org/10.1016/j.landurbplan.2012.12.005

22. Chowdhury, R. H., Prince, N. U., Abdullah, S. M., & Mim, L. A. (2024). The role of predictive analytics in cybersecurity: Detecting and preventing threats. *World Journal of Advanced Research and Reviews*, *23*(2), 1615-1623.

23. Du, H., Li, N., Brown, M. A., Peng, Y., & Shuai, Y. (2014). A bibliographic analysis of recent solar energy literatures: The expansion and evolution of a research field. *Renewable Energy*, *66*, 696–706. https://doi.org/10.1016/j.renene.2014.01.018

24. Marion, P., Bernela, B., Piccirilli, A., Estrine, B., Patouillard, N., Guilbot, J., & Jérôme, F. (2017). Sustainable chemistry: how to produce better and more from less? *Green Chemistry*, *19*(21), 4973–4989. https://doi.org/10.1039/c7gc02006f

25. McWilliams, J. C., Allian, A. D., Opalka, S. M., May, S. A., Journet, M., & Braden, T. M. (2018). The Evolving State of Continuous Processing in Pharmaceutical API Manufacturing: A Survey of Pharmaceutical Companies and Contract Manufacturing Organizations. *Organic Process Research & Development*, *22*(9), 1143–1166. https://doi.org/10.1021/acs.oprd.8b00160

26. Scognamiglio, V., Pezzotti, G., Pezzotti, I., Cano, J., Buonasera, K., Giannini, D., & Giardi, M. T. (2010). Biosensors for effective environmental and agrifood protection and commercialization: from research to market. *Microchimica Acta*, *170*(3–4), 215–225. https://doi.org/10.1007/s00604-010-0313-5

27. Singh, S., Jain, S., Ps, V., Tiwari, A. K., Nouni, M. R., Pandey, J. K., & Goel, S. (2015). Hydrogen: A sustainable fuel for future of the transport sector. *Renewable and Sustainable Energy Reviews*, *51*, 623–633. https://doi.org/10.1016/j.rser.2015.06.040

28. Springer Handbook of Inorganic Photochemistry. (2022). In *Springer handbooks*. https://doi.org/10.1007/978-3-030-63713-2

29. Su, Z., Zeng, Y., Romano, N., Manfreda, S., Francés, F., Dor, E. B., Szabó, B., Vico, G., Nasta, P., Zhuang, R., Francos, N., Mészáros, J., Sasso, S. F. D., Bassiouni, M., Zhang, L., Rwasoka, D. T., Retsios, B., Yu, L., Blatchford, M. L., & Mannaerts, C. (2020). An Integrative Information Aqueduct to Close the Gaps between Satellite Observation of Water Cycle and Local Sustainable Management of Water Resources. *Water*, *12*(5), 1495. https://doi.org/10.3390/w12051495

30. Carlson, D. A., Haurie, A., Vial, J. P., & Zachary, D. S. (2004). Large-scale convex optimization methods for air quality policy assessment. *Automatica*, *40*(3), 385–395. https://doi.org/10.1016/j.automatica.2003.09.019