Deep Neural Networks Application in Environmental and Water Resources Simulations

Mohammad Mahdi Rajabi

Researcher at the University of Luxembourg

mahdi.rajabi@uni.lu

Overview

- Background
- Data-Driven Machine Learning Models
- Deep Neural Networks (DNNs)
- Why are DNNs Particularly Important in EWRS?
- Diverse DNN Architectures
- Physics-informed Neural Networks
- Prospects for Future Research

Target Audience:

- This presentation is mostly designed for environmental and water resources engineers and scientists who are new to using Deep Neural Networks (DNNs) and Physics-Informed Neural Networks (PINNs).
- It is also valuable for experienced researchers in the field seeking insights into future research opportunities.

Background

Environmental and Water Resources Simulation: (EWRS) A computational process that models the behavior and interactions of environmental systems and water resources such as rivers, lakes, groundwater, and rainfall patterns to predict changes and plan management strategies.

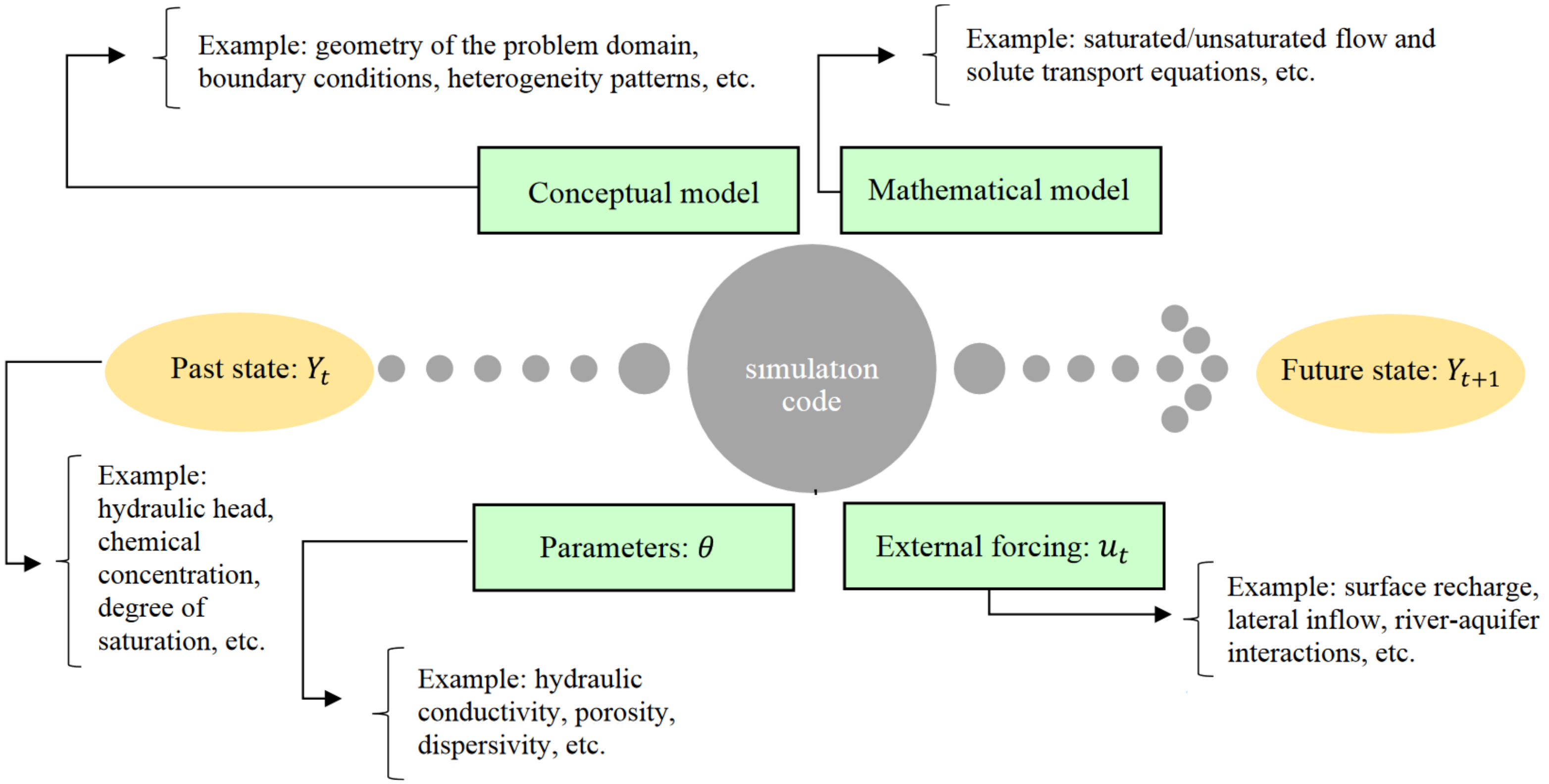
Scope: Encompasses a wide range of topics including, but not limited to, modeling hydraulic processes in surface waters, groundwater flow dynamics, the spread of pollution, sediment transport, water consumption patterns, and air pollution modeling.

Traditional Approach: EWRS has been traditionally based on **physics-based analytical and numerical models**.

Physics-based Analytical and Numerical Models

Simulation code examples:

- Groundwater models: MODFLOW, SEAWAT, etc.
- Rainfall run off models: SWMM, etc.
- Lake models: CE-QUAL-W2
- Water distribution network models: EPANET
- …



Source: Rajabi, M. M., Ataie-Ashtiani, B., & Simmons, C. T. (2018). Model-data interaction in groundwater studies: Review of methods, applications and future directions. *Journal of hydrology*, *567*, 457-477.

Challenges of Physics-based Models

Challenges with Effectiveness:

- ***High Computational Demand:*** Traditional models often require extensive computational resources and lengthy simulation times, especially for large-scale or repeated simulations.

  → These computational demands make tasks such as uncertainty analysis, simulation-optimization, and real-time predictions challenging and often unfeasible.

- ***Incomplete Understanding:*** Often, there is an incomplete understanding or imperfect mathematical representation of physical processes in traditional models.

  → This leads to frequent discrepancies between model predictions and actual field observations, affecting reliability.

Shift to Data-Driven Models: Increasing adoption of machine learning (ML) techniques in EWRS to overcome traditional model limitations.

Data-Driven ML Models employ algorithms to learn from data patterns and automatically map inputs to outputs, without relying on physical laws or explicit programming.

Data-Driven ML Models

Advantages of ML Models: Often faster compared to physics-based models, requiring less detailed information, can learn from data alone.

Traditional ML Models: Include techniques like random forests, support vector machines, polynomial chaos expansion, and tree-based regression models.

Common Feature:

- Use statistical methods to interpret and predict data patterns.
- All rely on structured datasets for training, learning from historical data to generalize insights to new scenarios without requiring knowledge of underlying physical processes.

History:

- These model began gaining popularity in EWRS in the late 1990s and early 2000s as computational resources became more accessible and datasets larger.
- They are still widely used today due to their effectiveness in many practical applications, robustness, and ease of understanding and implementation.

Use Cases of ML Models in EWRS

Surrogate or Meta-Modeling:

- ***Purpose:*** Replace original, computationally intensive physics-based models with faster ML models.
- ***Method:*** Use physics-based numerical models to generate input-output pairs, then train ML models on this data to create a surrogate model.
- ***Applications:*** Widely used for tasks requiring numerous simulations such as uncertainty propagation analysis, focusing on enhancing computational efficiency.

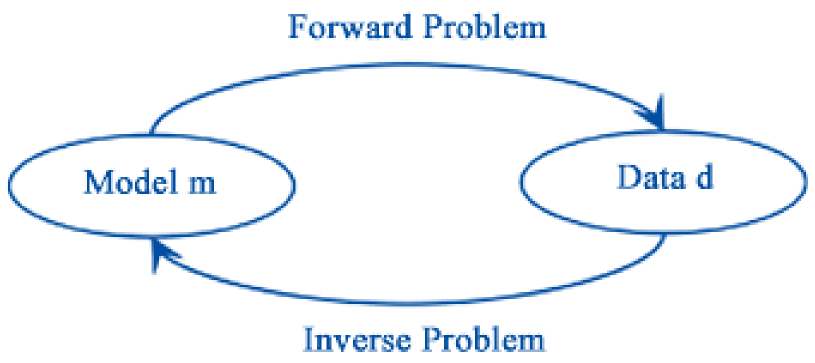
Developing Statistical Mappings:

- ***Challenge:*** Limited or non-existent mathematical formulations that incorporate physical laws.
- ***Solution:*** Use ML models to create statistical mappings between observable parameters and various states.
- ***Benefit:*** Enables predictions beyond the training data's scope, particularly useful in scenarios where traditional physical formulations are inadequate.

Use Cases of ML Models in EWRS

Inverse Modeling:

- ***Issue***: Many EWRS scenarios lack direct analytical or physical equations linking model inputs (like parameters and external forces) to predicted states, leading to complex, often ill-posed optimization problems.
- ***ML Approach***: Replace traditional optimization processes with ML models that establish direct mappings between states and inputs.
- ***Training***: Utilize forward simulations of the model or actual observations as training data, simplifying the inverse modeling process and improving problem tractability.



Use Cases of ML Models in EWRS

Feature Extraction:

- ***Objective***: Identify and isolate significant features from large datasets that impact system behavior and outcomes.
- ***Method***: Employ ML models to automatically detect and prioritize critical variables and interactions within environmental data.
- ***Applications***: Useful in complex environmental systems where traditional analytical methods may overlook subtle yet influential patterns. This capability enhances model accuracy and informs more effective resource management strategies.

Examples:

Independent Task Example: In water quality monitoring, feature extraction can independently analyze sensor data to identify key indicators of contamination events, aiding in immediate response and remediation efforts without further simulation.

Precursor to Simulation Example: For flood risk assessment, feature extraction might first identify critical geographical and climatic factors from historical data; these features are then used to simulate potential future flooding scenarios under various climate change models.

Challenges with Traditional ML in EWRS

i. Difficulty handling rare, unpredictable black swan events.

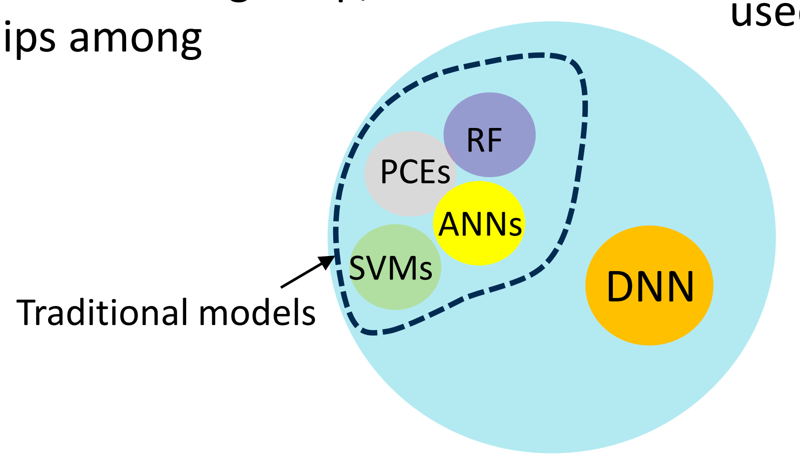
ii. Struggles to adapt to scenarios not present in training data.

iii. Inefficiencies in managing large data volumes.

iv. Limited capability in uncovering deep, complex relationships among variables.

The Alternative: Deep Neural Networks (DNNs)

- DNNs offer greater flexibility and higher prediction accuracy, especially with large datasets.
- Their advanced learning capabilities position DNNs as a leading tool in EWRS research.
- DNNs gain popularity around 2012 and became widely used in EWRS by the end of 2010s.



Machine learning

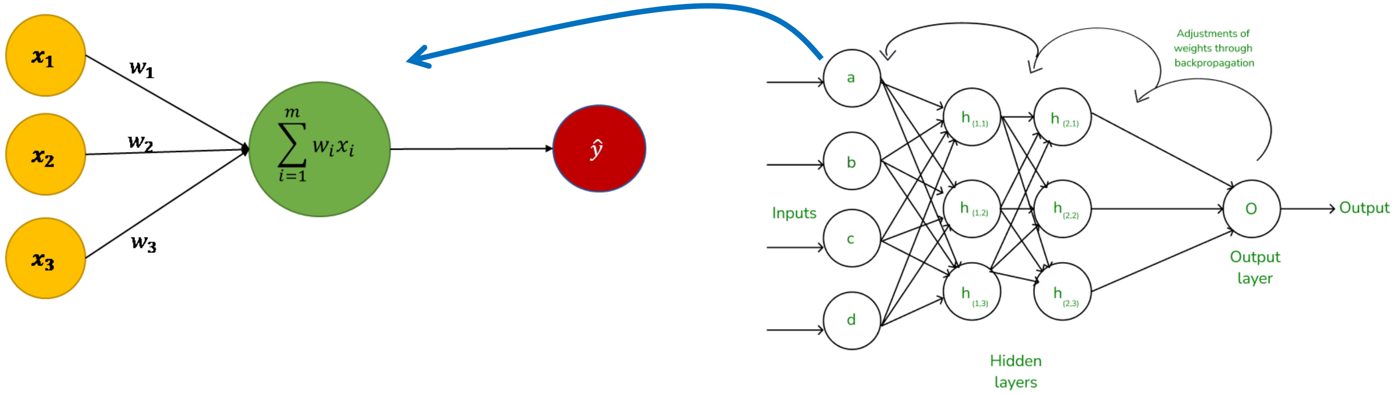
Deep Neural Networks

Neural Networks: Computational models inspired by the human brain, composed of interconnected layers of nodes (neurons). Each layer processes inputs from the previous layer using a structured network of nodes.

Processing Mechanism: Neurons compute a weighted sum of inputs and pass this through an activation function such as ReLU or Sigmoid. This step introduces non-linearity, enabling the network to perform complex pattern recognition.

Learning and Adaptation: During training, the network adjusts weights and activation function parameters to minimize output prediction errors. This fine-tuning helps improve accuracy over time.

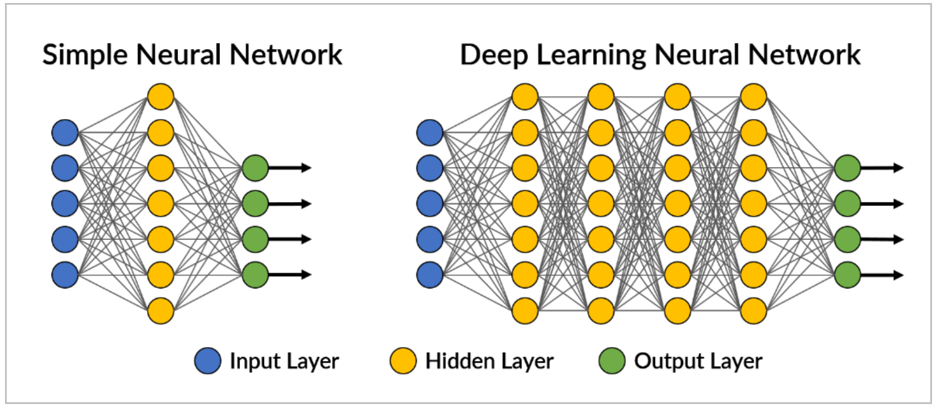
Backpropagation: A critical learning technique where the network updates its weights and biases based on the error gradient of the predictions. This method helps the network learn efficiently from mistakes by minimizing the discrepancy between predicted and actual outcomes.



Deep Neural Networks

- An extension of basic neural networks that include multiple hidden layers. These additional layers allow the extraction of progressively more abstract features from the data, enhancing the network's learning and predictive capabilities.

- **NNs vs. DNNs:** The distinction between neural networks and deep neural networks often blurs; generally, 'deep' refers to networks with multiple hidden layers that can capture deeper data insights, though there's no strict boundary defining when an NN becomes a DNN.



- DNNs are famously known for being **global approximators**, for their ability to approximate any continuous function to an arbitrary degree of accuracy, regardless of its complexity or non-linearity, over a certain domain.

- In other words, given a sufficiently large and well-designed Neural Network architecture, it can approximate any function that maps inputs to outputs with a high degree of accuracy.

Why are DNNs Particularly important in EWRS?

Handling Complexity: DNNs can manage the high complexity of environmental systems, capturing non-linear relationships and interactions that are often missed by traditional ML models due to simpler predictive algorithms, and physics-based models, which often require explicit mathematical formulations not always available for all processes.

Example: In hydrological modeling, DNNs can integrate various data types (e.g., rainfall, soil moisture, land use etc.) to better predict river flow rates.

Unlike physics-based models (e.g., SWAT) that rely on imperfect equations needing calibration for each scenario and often overlooking spatial and temporal changes, DNNs can analyze these inputs to enhance the prediction accuracy of river flow rates, capturing the complex and dynamic relationships that traditional models frequently miss.

Why are DNNs Particularly important in EWRS?

Big Data Capabilities: With the ability to process and learn from vast amounts of data, DNNs are adept at utilizing large datasets commonly found in environmental studies, improving the accuracy and reliability of predictions.

Feature Extraction: DNNs automatically identify and prioritize the most relevant features from raw data, which is crucial for accurately modeling dynamic and multi-dimensional environmental processes.

Enhanced Predictive Performance: By learning deep representations, DNNs often achieve superior predictive performance in tasks such as weather prediction, flood forecasting, and water quality management.

Adaptability: DNNs can adapt to new data without being explicitly reprogrammed, making them suitable for real-time environmental monitoring and responsive to changing conditions.

Diverse DNN Architectures:

Various DNN architectures are optimized for different data types and applications in EWRS, each offering specific advantages.

- **Transition in Usage:** The adoption in EWRS has shifted from primarily using deep feedforward neural networks to favoring convolutional neural networks (CNNs) and recurrent neural networks (RNNs), particularly long short-term memory (LSTM) networks, which are better suited for modeling spatial and temporal patterns.

- **Emerging Techniques:** Although less common, graph neural networks (GNNs) and generative adversarial networks (GANs) are gaining traction. GNNs analyze unstructured data effectively, while GANs are noted for their ability to generate synthetic data, enhancing modeling capabilities.

DNN architecture	Key features	Application examples
Feedforward Neural Networks (FFNNs)	Employs a straightforward architecture with layers connected in a feedforward manner. Commonly used for vector data.	Water demand forecasting (Gil-Gamboa et al., 2024).
Convolutional Neural Networks (CNNs)	Leverages convolutional layers for spatial hierarchy feature extraction. Go to option for image data and related tasks, such as image-to-image regression, image segmentation and image classification.	Forecasting contaminant and temperature distribution maps based on maps of material properties (Rajabi et al., 2022), Land cover prediction (Stoian et al., 2019).
Recurrent Neural Networks (RNNs)	Utilizes memory elements to process sequences and temporal dependencies.	Flood forecasting (Kao et al., 2020).
Hybrid CNN- Long short-term memory (LSTM)	Developed for spatiotemporal data analysis.	Flowrate prediction in a watershed (Li et al., 2023).
Graph Neural Networks (GNNs)	Uses graph-based data representations for node and edge analysis. Compatible with point cloud and mesh-based data.	Groundwater level forecasting (Bai and Tahmasebi, 2023).
Generative Adversarial Networks (GANs)	Involves a dueling setup where one network generates data, and another evaluates it.	Generating synthetic weather data (Ji et al., 2024), Leak detection in water distribution networks (Rajabi et al., 2023).

Emerging Trends

Integration of Advanced Architectures: Recent advancements include the use of CNNs combined with RNNs, like CNN-LSTMs, for efficient analysis of spatial and temporal data. Additionally, technologies like PointNet and various GNNs are being explored for processing point clouds from LiDAR and similar technologies for detailed environmental analysis.

Sophistication in Application: The use of DNNs in EWRS has progressed from basic regression and time series prediction tasks to more complex predictive modeling, such as image-based regression tasks (both image-to-value and image-to-image).

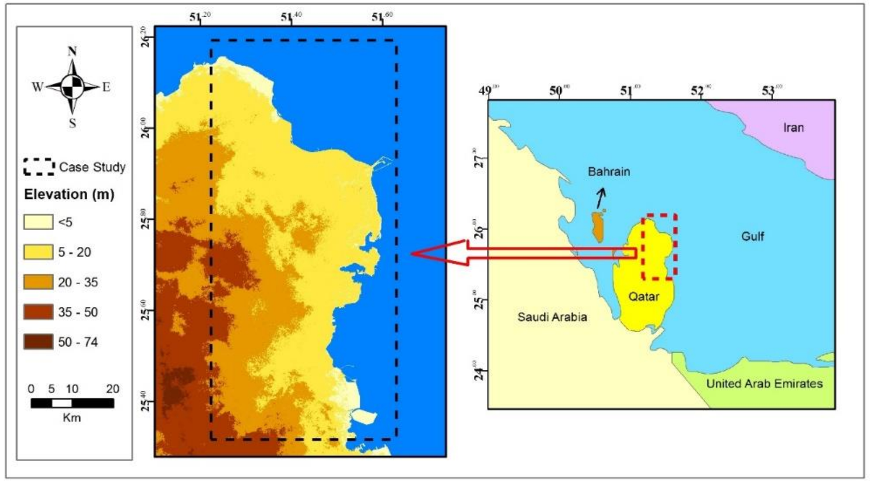
Example: Land Use Classification

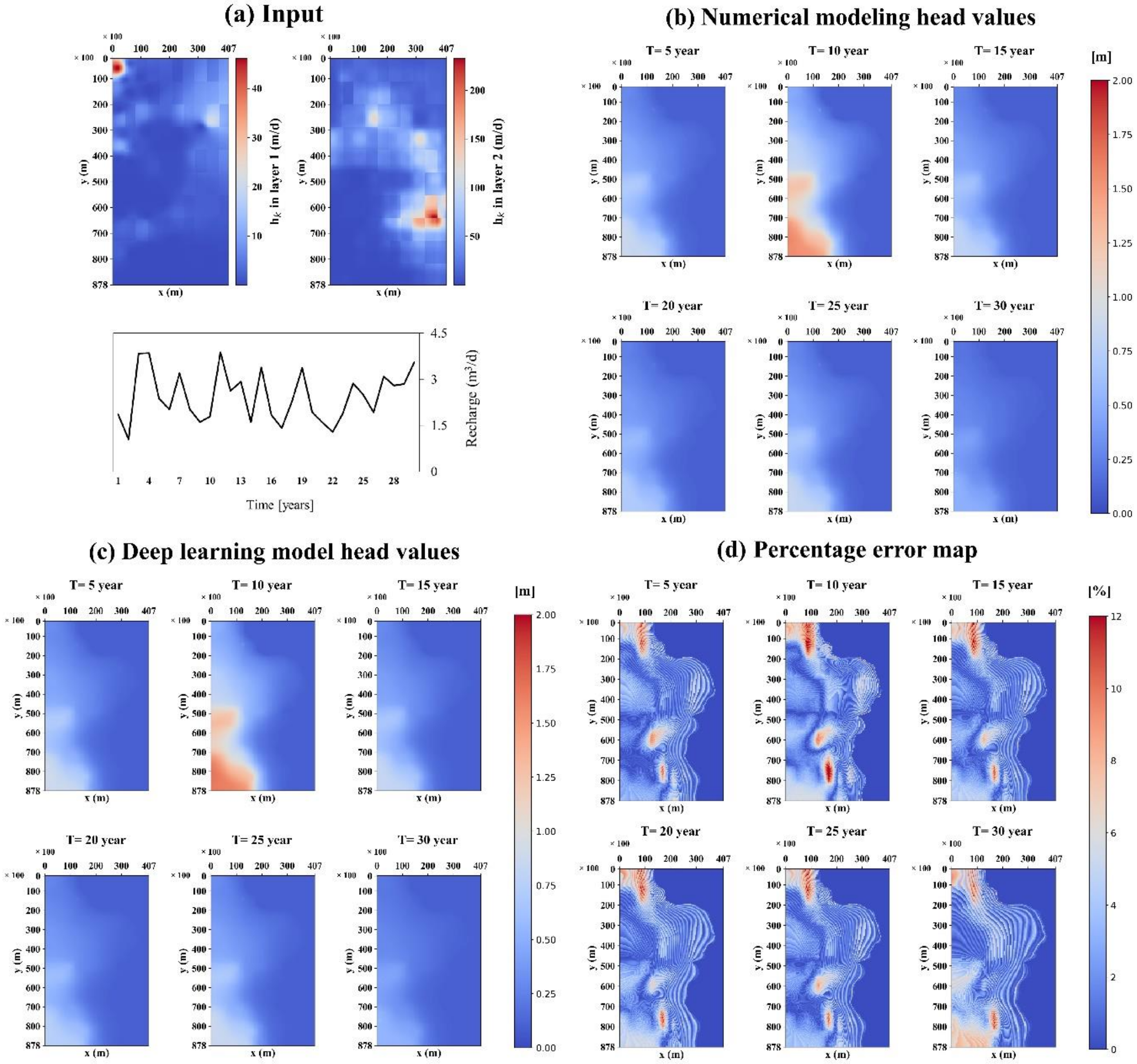
Example: Predicting flooding maps from rainfall maps

Example of Image-to-Image Regression using DNNs in EWRS

Source: Accelerating Regional-scale Groundwater Flow Simulations with a Hybrid Deep Neural Network Model Incorporating Mixed Input Types: A Case Study of the Northeast Qatar Aquifer (Under Review)

Ali Al-Maktoumi, Mohammad Mahdi Rajabi, Slim Zekri, Rajesh Govindan, Aref Panjehfouladgaran, Zahra Hajibagheri

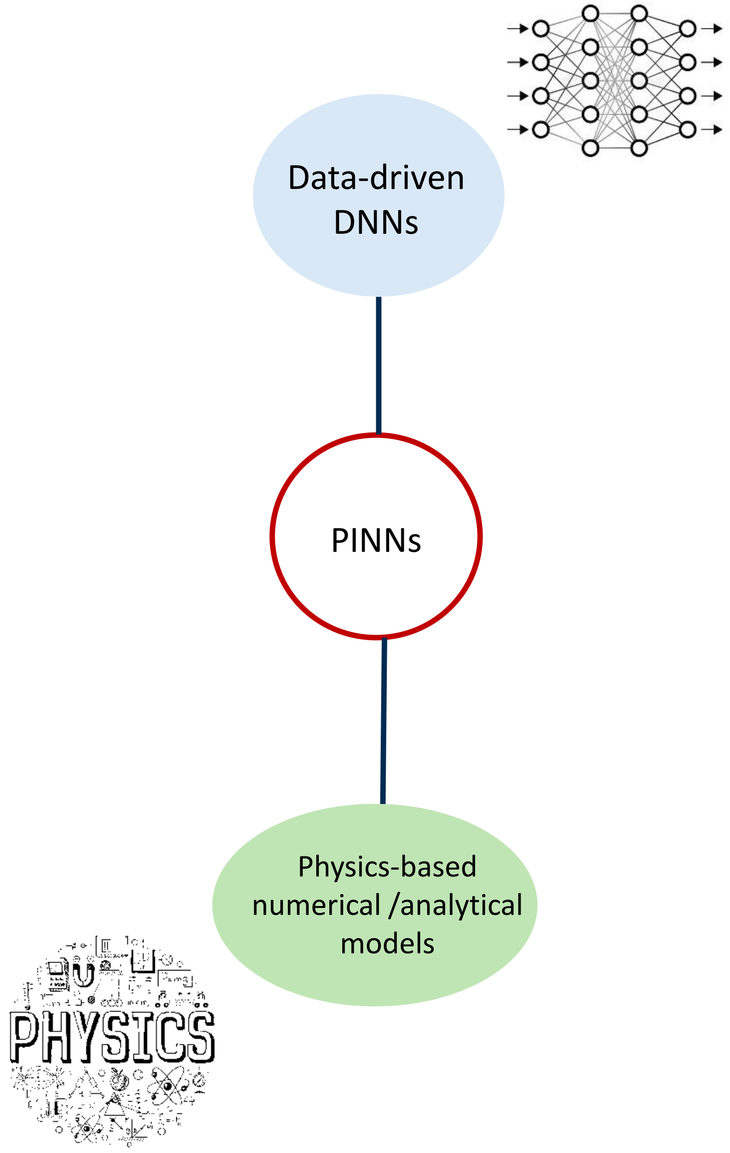




Emerging Trends

Shift to Field Data: Increasingly, DNNs are being trained using field data from sensors and aerial imagery, expanding their use to phenomena that are difficult to model with traditional physics-based approaches.

Physics-Informed DNNs: A novel approach involves embedding physical laws directly into the loss function of DNNs, creating physics-informed neural networks (PINNs). These networks utilize physics-based losses as part of the objective function in backpropagation or as constraints.



Physics-informed Neural Networks

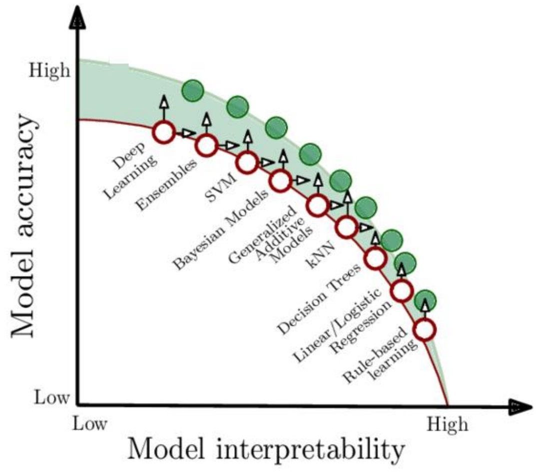
Efficient Data Utilization: PINNs efficiently train on limited data due to their integration of physical knowledge, making them ideal for scenarios where extensive data collection is challenging or costly.

Enhanced Interpretability: While data-driven neural networks are often seen as "black boxes," PINNs offer clearer insights by incorporating physically meaningful components (although they don't fully resolve all interpretability concerns).

Improved generalization: PINNs incorporate physical information, which allows them to better generalize to regions of the training space that lack training data.

Mitigated Overfitting: PINNs, by integrating physical constraints, capture the core physics of the modeled problem, thereby minimizing overfitting and enhancing performance on noisy data.

Improved accuracy: By incorporating physical information, PINNs are better suited for making accurate predictions on unseen data and are less prone to proving physically unrealistic solutions.



Computational Burden of PINNs vs. DNNs

- In general, training a PINN can be more computationally intensive than training a data-driven neural network with the same architecture -> PINNs typically require solving PDEs or other physics-based constraints as part of the optimization process
- PINNs can reduce the amount of labeled data needed for training, which can be especially valuable in cases where obtaining labeled data is computationally expensive.
- PINNs can also converge faster during training due to the additional information provided by the physics

-> So, the computational burden of PINNs vs. data-driven NNs (with the same architecture) depend on the specific problem at hand.

Why Not Use PINNs?

There are a variety of scenarios in which PINNs may not be the best choice of modeling approach:

- If our understanding of the system is limited, unreliable or unquantifiable,
- If the problem is low-dimensional and has a relatively simple alternative model-> then PINNs are too heavy of a tool to warrant their use.
- If data is too limited -> While PINNs can handle data-scarce problems, they require some amount of data for training. If data is extremely limited, other machine learning approaches may be more suitable.
- Limited computational resources -> PINNs can be computationally expensive to train and may require high-performance computing resources. If computational resources are limited, a simpler machine learning approach may be more appropriate.

Different Methods of Incorporating Physics in PINNs

Optimization-Based:

- Joint Loss Optimization: This is the most common method in PINNs. The total loss is a combination of the data loss and the physics loss. The model is trained to minimize this combined loss.
- Sequential Optimization: Initially, train the network to fit the data. Once this is achieved, train the network to minimize the physics loss. This two-step process can sometimes help in stabilizing training.

Regularization-Based:

- Physical Regularization: The physical loss acts as a regularizer to the primary data loss. While the main objective remains data fitting, the physics loss ensures the learned function doesn't deviate significantly from physical laws.

Regularization-Based: Soft vs. Hard Constraints

- In the soft regularization method, the physics loss is added as a penalty term to the primary data loss:

$$\mathcal{L}_{total} = \mathcal{L}_{data\_driven} + \underbrace{\boxed{\lambda\, \mathcal{L}_{physics-informed}}}_{} \rightarrow \text{Penalty term}$$

Regularization parameter

A larger λ will enforce the physical laws more strictly.

- In the hard regularization method, the optimization problem can be stated as:

$$\min \mathcal{L}_{data-driven}$$

$$Subject\ to\ \mathcal{L}_{physics-informed} = 0$$

Minimized the data-driven loss while ensuring that the predicted function strictly adheres to the physical laws.

Prospects for Future Research

Capturing Multi-Scale Interactions:
Creating DNNs/PINNs that can effectively handle multi-scale interactions within environmental systems to provide more accurate and comprehensive simulations of complex phenomena.

Leveraging Transfer Learning:
Investigating transfer learning techniques to use knowledge from related fields or datasets, enhancing the performance of DNNs in environmental and water resource applications.

Enhancing Scalability: Tackling scalability challenges by exploring scalable architectures and algorithms capable of efficiently processing and analyzing large-scale environmental datasets. Notably important in using RS & GIS data in EWRSs.

Integration into Decision Support Systems: Exploring how DNN-based predictive models can be integrated into decision support systems for environmental management and planning, enabling better-informed decision-making based on the insights provided by these advanced models.

Prospects for Future Research

Multi-fidelity modeling is a technique that combines models of varying levels of complexity to create a more efficient and accurate model.

- Example: Utilizes a low-fidelity model to generate training data for a high-fidelity model, effectively reducing the computational cost of training.
- *Hierarchical Modeling Approach:* Employs a series of models of increasing complexity, where a low-fidelity model provides initial conditions for a higher-fidelity model, iteratively enhancing accuracy until the desired level is achieved.

Another area of active research is the **incorporation of uncertainty into DNN & PINNS.** This involves developing new methods for estimating and propagating uncertainty through the model, which can help to improve the reliability of predictions. Bayesian PINNs are an important option in this context.

Thank You for Your Attention!

Let's Connect

- *Mohammad Mahdi Rajabi: mahdi.rajabi@uni.lu*

Looking forward to our discussion!