# Machine Learning Techniques for Determination of Groundwater Level in Mining

**Author:** Amirhossein Najafabadipour, PhD

April 2024

## Problem: Deepening Mines and Groundwater

- Increasing the mining depth causes the mine pit to be situated below the water table.
- Groundwater influx disrupts operations, raises costs, and threatens safety.
- Accurate groundwater level estimation is crucial for managing water resources.

## Traditional Methods: Limitations

- Traditional methods rely on complex physical models.
- These models require an in-depth understanding of subsurface conditions, often unavailable.
- Difficulty in handling non-linear relationships between variables.

## Machine Learning: A Promising Approach

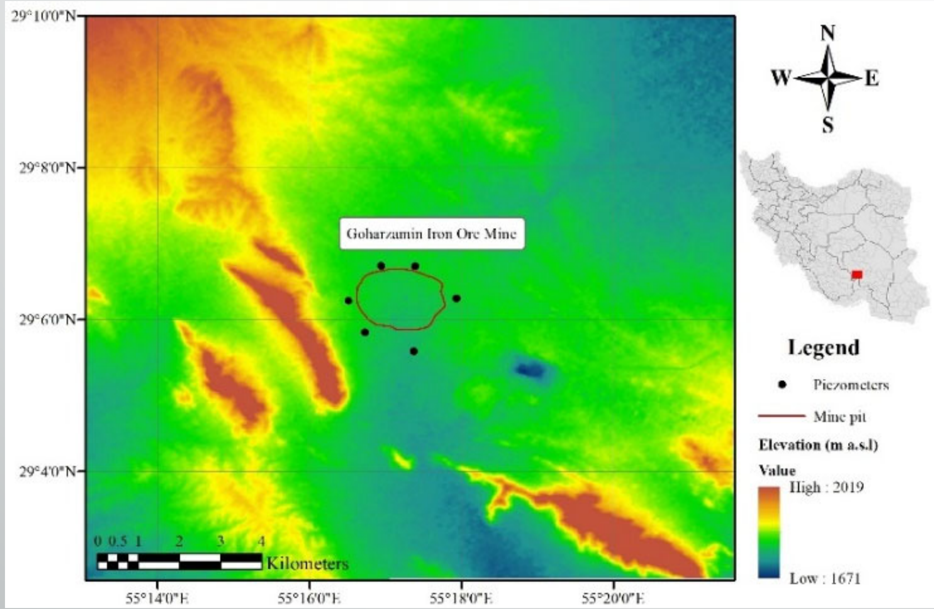
- Machine learning offers a data-driven alternative.
- It can handle complex relationships from large datasets.
- No need for an intricate understanding of physical parameters.

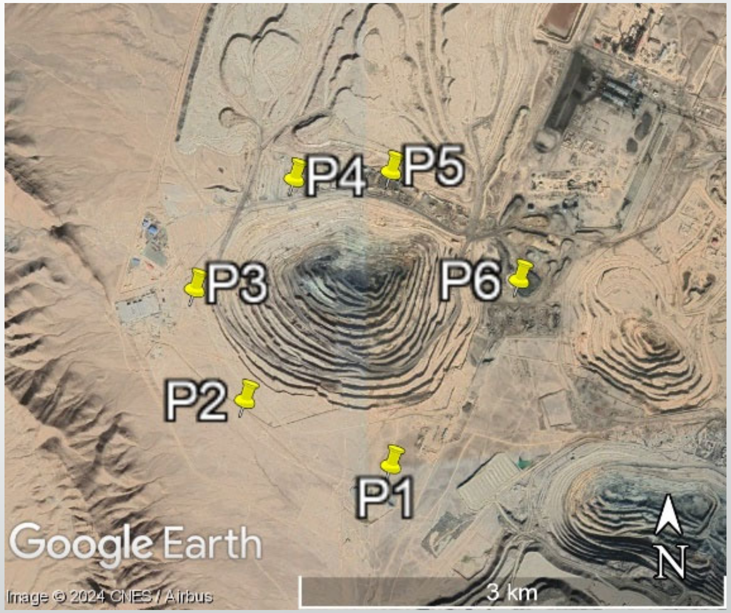
## Our Research Objectives

- Develop machine learning models to predict groundwater levels in mining.
- Leverage spatial and temporal data for prediction.
- Analyze the impact of input features on groundwater levels.
- Investigate trends in drainage within the developed models.

# Study Area: Gohar Zamin Iron Ore Mine

- Located in the Middle East, a major mining district.
- Six separate anomalies with a 1.2 billion-ton deposit.
- Area of 40 square kilometers.
- Young erosion and alluvial deposits blanket the region.




The studied area (Gohar Zamin iron ore mine) along with the location of piezometers

## Data Collection

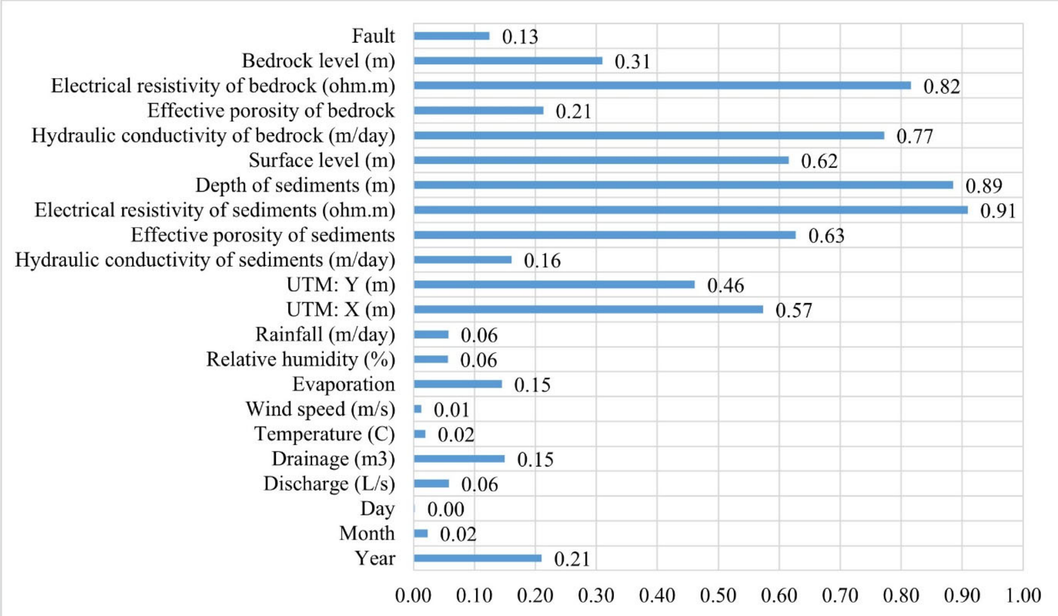
- 3534 data points from six piezometers for groundwater level determination.
- 10 temporal features (year, month, day, discharge, drainage, temperature, wind speed, humidity, evaporation, rainfall).
- 12 spatial features (latitude, longitude, effective porosity of bedrock and sediments, hydraulic conductivity of bedrock and sediments, depth of sediments, the electrical resistivity of bedrock and sediments, surface level, and fault).

## Machine Learning Models

- Multilayer Perceptron optimized with Batch Training (MLP-B) - a type of neural network.
- Cascade Forward optimized with Gradient Descent (CF-GDA) - another neural network structure.
- Multivariate Adaptive Regression Splines (MARS) - a regression technique for piecewise linear models.
- Random Subspace Ensemble (RS) - utilizes random feature subsets for training multiple models.

# Feature Selection

- Relevancy factor analysis was employed to pinpoint the features that have the most significant impact on groundwater level.
- The electrical resistivity of sediments was found to be the most influential factor affecting groundwater levels.
- After careful consideration, fifteen features were chosen for machine learning model development.
- The groundwater level in the complex mining area remains largely unaffected by rainfall due to the predominance of fossil water. Consequently, rainfall does not significantly alter the groundwater level, while variations in the electrical resistivity of sediments, influenced by water presence, exert the most significant influence on groundwater levels.




Feature selection for temporal and spatial input features that impact groundwater level

# Data Preprocessing

- Data cleaning is essential to ensure the quality and integrity of the data used for model training.
- Outliers, data points that deviate significantly from the norm, were identified and removed appropriately.
- Data normalization was performed to ensure all features were on a similar scale.

# Model Development

- Training data provides the models with the information they need to learn the patterns and relationships between the input features and the target variable (groundwater level).
- 70% of the data was designated for training the models.
- The remaining 30% of the data was reserved for testing the models' performance.

# Evaluation Metrics

- Average Absolute Relative Error (AARE): Measures the average difference between the predicted and actual groundwater levels. Lower AARE indicates higher accuracy.
- R-squared ($R^2$): Represents the proportion of variance in the target variable explained by the model. A value closer to 1 signifies a better fit.

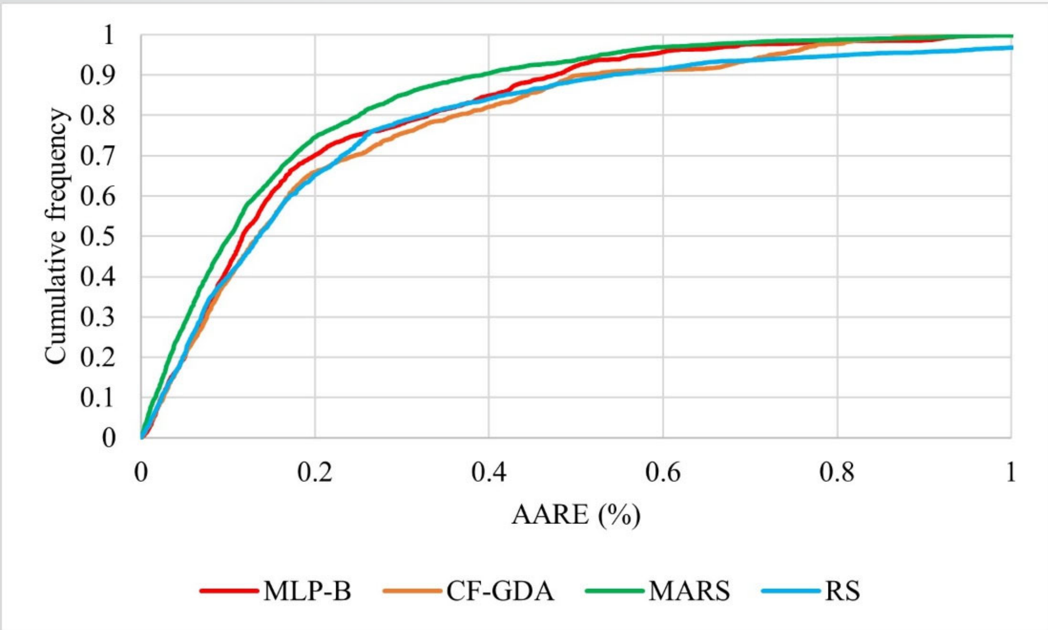
Statistical parameters of all developed models

| Models | MLP-B | Cf-GDA | MARS | RS |
|---|---|---|---|---|
| **AARE (%)** | 0.188 | 0.213 | 0.158 | 0.223 |
| **$R^2$** | 0.9917 | 0.9894 | 0.9936 | 0.9863 |

MARS > MLP-B > CF-GDA > RS

# Evaluation Metrics

For the graphical comparison of the models, the cumulative frequency of the average absolute relative error (AARE) is plotted for all the developed models.

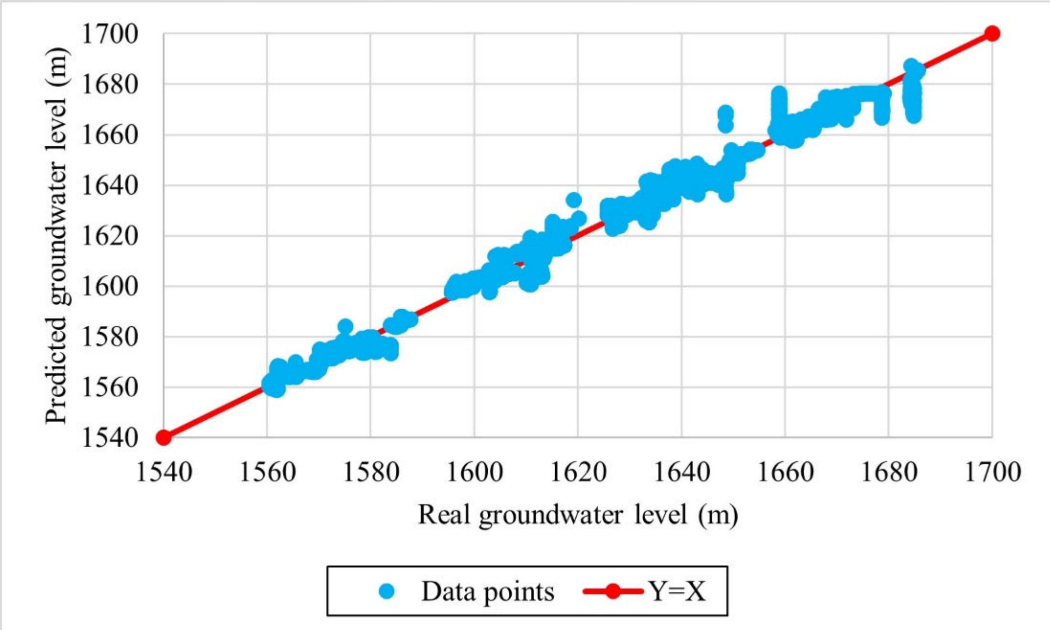



Cumulative frequency vs AARE for all developed models

Introduction → Material and Methods → Results and Discussion → Conclusion

# Evaluation Metrics

To validate the accuracy of the MARS model, the real groundwater level data are plotted against the determined groundwater level.

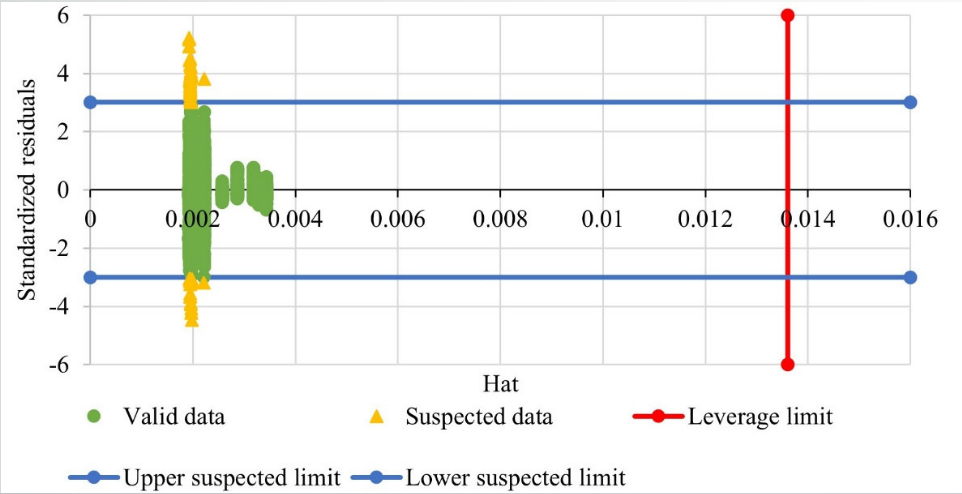



Cross plots of the MARS model for the groundwater level

Introduction → Material and Methods → Results and Discussion → Conclusion

# Applicability Domain of the Model

- The majority of determined points fall within the feasible domain of the developed model ($0 \le hat \le 0.013$ and $-3 \le R \le 3$), demonstrating the high reliability and statistical validity of the MARS model.
- The analysis revealed that approximately 98% of the data points fell within the acceptable range for the MARS model.




Applicability domain of the MARS model

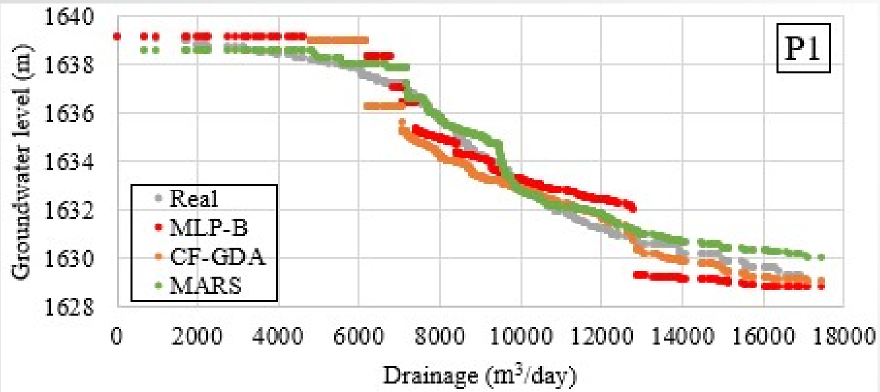


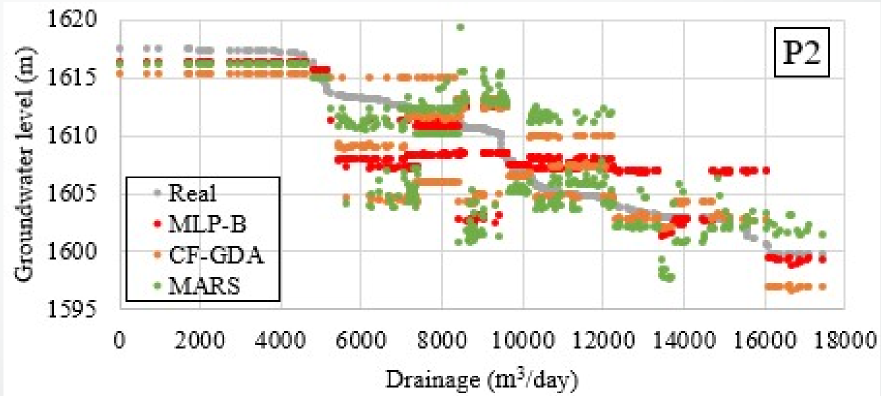
Introduction → Material and Methods → Results and Discussion → Conclusion

# Trend Analysis: Drainage and Groundwater Levels

- The models were used to investigate the relationship between drainage and groundwater levels.
- The analysis revealed a clear trend: as drainage intensity increases, groundwater levels decrease.
- This finding aligns with our understanding of how increased pumping removes water from the ground, lowering the overall water table.

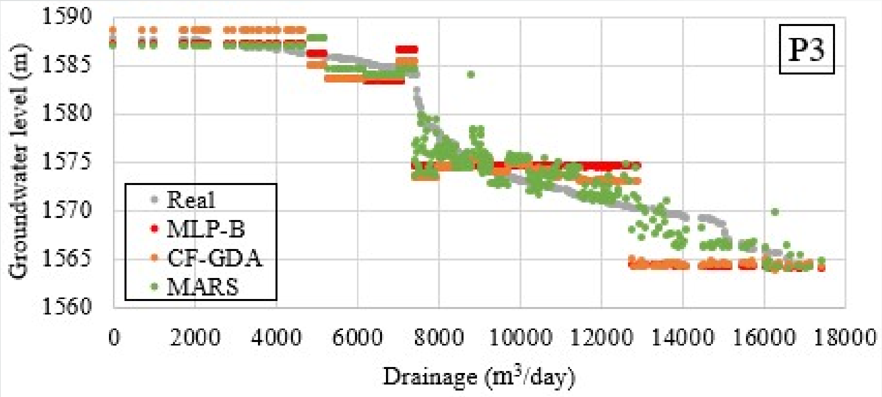
# Impact of Drainage on Groundwater Levels

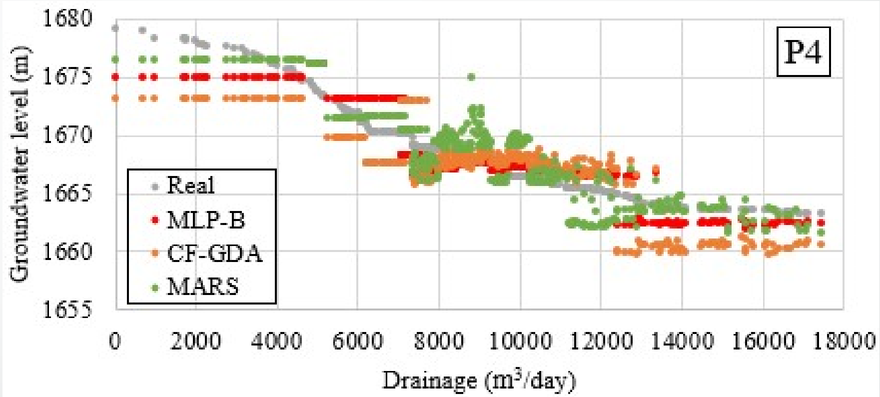



Comparing the groundwater level variation with drainage for the MLP-B, CF-GDA, and MARS models

# Impact of Drainage on Groundwater Levels

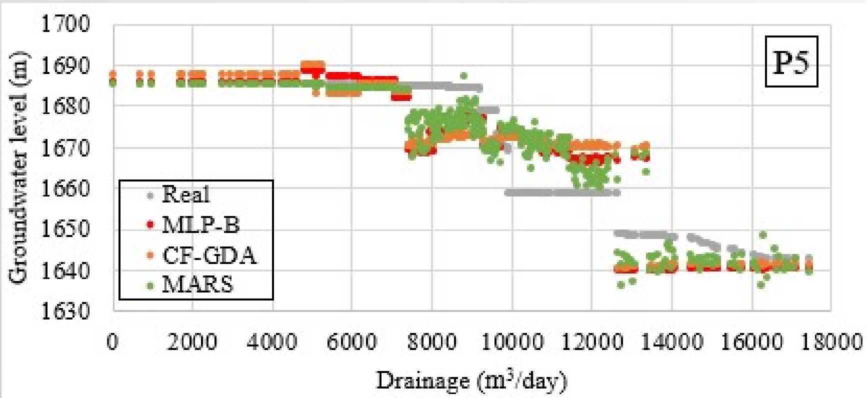


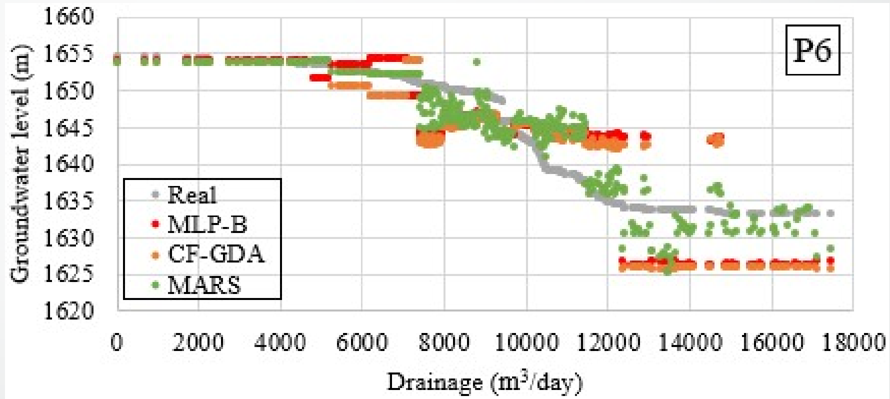




Comparing the groundwater level variation with drainage for the MLP-B, CF-GDA, and MARS models

Introduction → Material and Methods → **Results and Discussion** → Conclusion

# Impact of Drainage on Groundwater Levels




Comparing the groundwater level variation with drainage for the MLP-B, CF-GDA, and MARS models

Introduction → Material and Methods → Results and Discussion → Conclusion

# Conclusion

- Machine learning offers a powerful tool for predicting groundwater levels in mining operations.
- The MARS model achieved the most accurate results in this study, demonstrating the potential of machine learning for real-world applications.
- The capability of the MARS model to determine new data can be significant in areas where data scarcity is a common challenge.
- This study highlights the importance of drainage management for controlling groundwater levels in mines.
- The electrical resistivity emerged as a crucial factor influencing groundwater level fluctuations.
- Expanding the quantity of input data is recommended to enhance the precision of machine learning techniques.

# Thanks for your attention

najafabadipour@eng.uk.ac.ir
amirhossein.najafabadipour@gmail.com