FIU | Applied Research Center
FLORIDA INTERNATIONAL UNIVERSITY

U.S. DEPARTMENT OF ENERGY

INFORMATION TECHNOLOGY/ARTIFICIAL INTELLIGENCE

PROJECT: Artificial Intelligence for EM Problem Set (Soil and Groundwater) – Machine Learning Modeling to Identify Temporal and Spatial Relationships between Inland and Shoreline Hexavalent Chromium [Cr(VI)] Concentrations in 100 Areas

CLIENT: U.S. Department of Energy
SITE: Pacific Northwest National Laboratory (PNNL)
PRINCIPAL INVESTIGATOR: Dr. Leonel Lagos

Description:

Hexavalent chromium (Cr(VI)) is one of the primary contaminants in the 100 Areas at the U.S. Department of Energy's (DOE's) Hanford Site. Various cleanup efforts are ongoing to remediate this waste site since the late 1990s. To estimate the effects of these cleanup efforts and plan future cleanup actions, it is necessary to analyze Cr(VI) dynamics in the groundwater and surface water. The monitoring data available for groundwater wells and aquifer tubes, as well as water table levels and river stage monitoring sampled at 100 Areas, can be used with Artificial Intelligence/Machine Learning (AI/ML) algorithms to understand complex hydrogeological processes and interactions among the aquifers and the dynamic river stage in the Columbia River. AI/ML algorithms can leverage high-performance computing to predict the spatial and temporal distribution of Cr(VI) and also help in identifying any spatiotemporal relationship in the dataset.

The overall goal is to couple long-term monitoring data of hexavalent chromium Cr(VI) with AI/ML models to identify temporal and spatial relationships of subsurface chromium transport that reduces uncertainties in the conceptual site model (CSM).

Objectives:

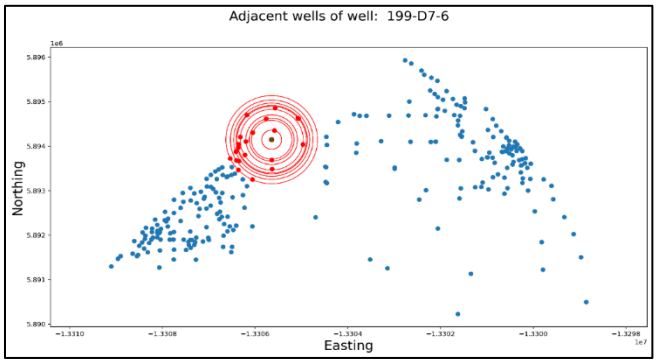
- Data source identification and pre-processing for data cleansing, discretization, transformation, and reduction to ensure data quality.
- Perform exploratory data analysis using state-of-art statistical methods and various machine learning algorithms.
- Develop AI/ML models to explore spatiotemporal relationships of subsurface hexavalent chromium transport.

Benefits:

- Identify the missing values in soil and groundwater legacy dataset.
- Extract usable data subset using data filter for AI/ML implementation.
- Explore spatiotemporal relationships between various groups for monitoring different well datasets.

Accomplishments:

- Developed an AI-based algorithm to resolve the key issue of missing values in soil and groundwater legacy datasets from 100 Areas, using Prophet Forecasting Model.



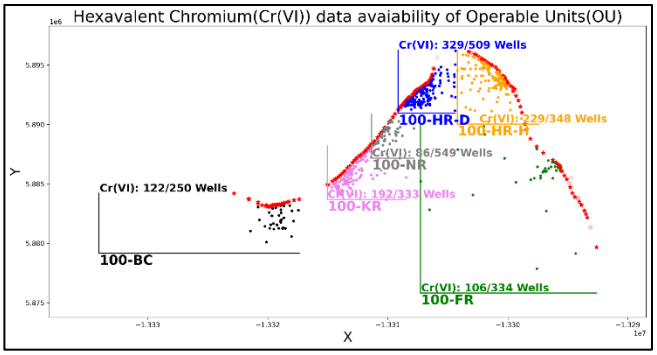
Adjacent wells identification for regressor implementation using prophet algorithm

- Deep learning-based anomaly detection system is implemented using the 100 Area Cr (IV) concentration dataset. The system utilizes the LSTM autoencoder (AE)-based anomaly detection approach.

ABOUT

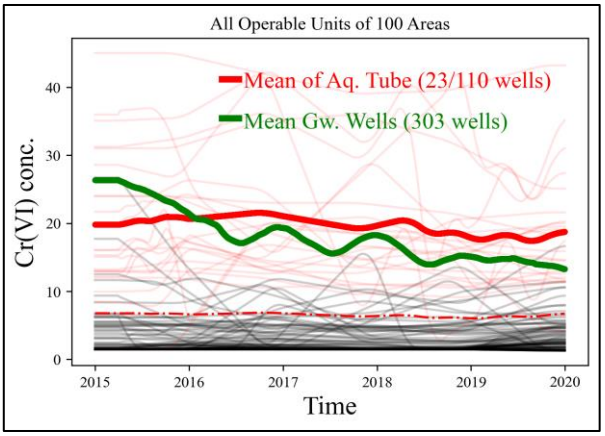
Since 1995, the Applied Research Center at Florida International University has provided critical support to the Department of Energy's Office of Environmental Management mission of accelerated risk reduction and cleanup of the environmental legacy of the nation's nuclear weapons program. ARC's research performed under the DOE-FIU Cooperative Agreement (Contract#DE-EM0005213) can be classified as fundamental/basic, proof of principle, prototyping and laboratory experimentation.

Project Contact:
Dr. Himanshu Upadhyay
Ph: (305) 348-6603
Email: upadhyay@fiu.edu
10555 W. Flagler Street, EC 2100
Miami, FL 33174
arc.fiu.edu



Hexavalent Chromium Cr (VI) data availability of wells in 100 Areas

- Based on the explored properties of the dataset, a filtering mechanism is developed to filter the usable subset of the dataset for time-series pattern recognition. Analyzed the overall statistical features of the time-series using metrics such as cosine similarity and Pearson correlation coefficient on the filtered dataset.



Ground water and surface water Cr(VI) time series similarity with data filter

- A method was developed to explore any probable spatiotemporal relationships between inland monitoring wells and shoreline Cr(VI) concentrations using data filter and machine learning models.