ILLINOIS Water Environment Association

Enhancing the Illinois Water Environment through Education and Collaboration.

2025 IWS Challenge Sample Topics

You are encouraged to develop your own topic to participate in this challenge. This could be a topic from your current course project, graduate project, or a question or problem you observed and are interested in solving using data analytics. However, you may instead select a challenge topic from below. The topic should focus on resolving problems in environmental field.

Challenge Plan Topic 1 – Estimate Plant Influent Carbon Fractionation

Leverage existing data from one/several Wastewater Reclamation Plants to create a model that can estimate carbon fractions (e.g., total carbonaceous oxygen demand (COD), soluble and colloidal COD (scCOD, a.k.a. filtered COD (fCOD)), soluble COD (sCOD, a.k.a., flocculated and filtered COD (ffCOD)) as well as other fractions that are difficult or impossible to measure directly (e.g., fraction of total COD that is comprised of particulate unbiodegradable organics (Xu), fraction of colloidal unbiodegradable organics in colloidal COD (Cu), and fraction of total COD that is biomass). These fractions are a critical component of all process modeling software packages (e.g., SUMO, GPS-X, BioWin). The fraction names vary across software packages, but typically represent the same component. These fractions can fluctuate greatly depending on the time scale being considered, so time it a critical component to consider.

Data

- Data must be collected by Team

Problems

- Build a ML model that can estimate COD fractionation based on influent data
- User should be able to input a set of values (e.g., flow rate, COD and/or BOD, TKN and/or ammonia, etc.) and receive an estimate of appropriate fractions for use in a plant modeling package (e.g., SUMO, GPS-X, BioWin, etc.)

Things to Consider

- Make necessary assumption(s) if needed, but clearly state your assumption(s) in your plan.

ILLINOIS Water Environment Association

Enhancing the Illinois Water Environment through Education and Collaboration.

Challenge Plan Topic 2 – Predict influent flow from historic flow and weather forecasts

Leverage local weather information as well as historic influent flow data to estimate near-term (e.g., in the next few hours or days) influent flows to a wastewater facility.

Data

- 3 years historic weather data
- 3 years historic influent flow data

Problems

- Make predictions of flow

Things to Consider

- Make necessary assumption(s) if needed, but clearly state your assumption(s) in your plan.
- Model framework should be flexible enough to be extended to other facilities at other locations.

ILLINOIS Water Environment Association

Enhancing the Illinois Water Environment through Education and Collaboration.

Challenge Plan Topic 3 – Dewatering process optimization for a large Water Resource Recovery Facility (WRRF)

Develop a data driven tool to optimize post digestion centrifuge operation at a WRRF to minimize the use of polymer while increasing process throughput and ensuring a minimum of 95% of sludge total solids captured by the process to avoid impacting downstream processes. The post digestion centrifuge machines serve to dewater – or further concentrate – human waste sludge to minimize disposal costs associated with the wastewater treatment process. Post digestion centrifuge processes are developed based on empirical data and no theoretical formula is known to predict the process output.

Data

- Six months of operation data for 21 post digestion centrifuges in operation.
- Additional data is available but requires a significant effort to transcribe by hand.

Available parameters

machine	An integer identifying the machine (1-21)
HrsOps	Hours of operations on a given date for a given machine (0-24hrs)
Month	Month of the year
Day	Day of the month
Tsfeed	% Total solids of the sludge feeding the centrifuge machine
TVSfeed	% Total volatile solids of the total solids in the sludge feeding the centrifuge machine
CKTS	Cake, or outflow solids discharge containing the captured solids, % total solids
CNTS	Centrate, or outflow slurry water cleared of captured solids, % total solids
RawPolyTS	% Total solids of the coagulating polymer in its raw form, prior to diluting it with water
DilPolyTS	% Total solids of the coagulating polymer after it has been diluted for use
SludgeFlow	The flow of sludge, in gpm (gallons per minute) being fed into the centrifuge machine
PolyFlow	The flow of coagulating polymer, in gpm (gallons per minute) being fed into the centrifuge machine
PinionSpeed	The rotation speed of the centrifuge conveyer, in rpm (revolutions per minute)
Torque	The torque produced by the centrifuge motor, in in-lbs (inch pounds)
BowlSpeed	The rotation speed of the centrifuge bowl, in rpm (revolutions per minute)
CaptureP	A calculated value for the % of the sludge total solids captured by the process, excluding the total solids introduced by the polymer flow

Mentor

- Eddie Paulino (paulinoe@mwrd.org)

ILLINOIS Water Environment Association

Enhancing the Illinois Water Environment through Education and Collaboration.

Challenge Plan Topic 4 – Predict flow and influent wastewater characteristics

Leverage existing data from one/several Wastewater Reclamation Plants to estimate influent flow and/or characteristic predictors (e.g., 5-day biological oxygen demand (BOD5), total phosphorus (P-TOT), soluble phosphorus (P-SOL), suspended solids (SS), total Kjeldahl nitrogen (TKN), and/or ammonia-N (NH3-N)) for a Wastewater Reclamation Plant.

Data

- 20 years of complete influent data for Plant A and B.
- 20 years of complete influent data for plant C, except that only 2 years of influent P-SOL data are available.
- Only 2 years complete data is available for plant D.
- Weather data available at https://www.ncdc.noaa.gov/cdo-web/search for the CHICAGO O HARE (ORD) station and/or https://mesonet.agron.iastate.edu/request/download.phtml?network=IL_ASOS

Conditions

- Due to the lab test holding time, all influent characteristic data for the four plants are lagged by 5 days. This lag does not apply to flow data. All four plants expect a new P-TOT limit in their National Pollutant Discharge Elimination System (NPDES) permit and are planning to implement enhanced biological phosphorus removal (EBPR) with chemical phosphorus removal polishing. They would like to predict plant flow, influent BOD5, influent P-TOT, and influent P-SOL concentration to help optimize EBPR operation and determine chemical needed for P polishing.

Problems

- Make predictions of flow and influent wastewater characteristics (particularly BOD5, P-TOT, and P-SOL) for one or more plant

Things to Consider

- A complete list of influent characteristics will be provided. Please select the proper parameters to build model.
- Make necessary assumption(s) if needed, but clearly state your assumption(s) in your plan.

Mentor

- Fenghua Yang (YangF@mwrd.org)