Developing an Intelligent Advanced Warning System for Odors

at MWRDGC's Thornton Composite Reservoir

INTELLIGENT WATER SYSTEM CHALLENGE 2018

TEAM

A collaboration between the Metropolitan Water Reclamation District of Greater Chicago (MWRDGC), the University of Illinois, and Ensaras, Inc.

Dr. Kuldip Kumar, Senior Environmental Soil Scientist, MWRDGC Chicago (Team Leader and Utility Research), has more than 30 years of experience in the field, with research background in fate and transport of environmental stressors in the terrestrial and aquatic environment. He is a member of WERF Utility Innovation project and for the past few years has been involved in evaluation of new and innovative technologies to improve water and resource recovery at the MWRDGC.

Dr. Dominic Brose, Environmental Soil Scientist, MWRDGC Chicago, (Utility Research), evaluates new technologies and processes for water resource recovery and solids treatment for the capital planning group at the MWRDGC. He has an extensive background in applying statistical and risk-based methods to understanding environmental processes.

John Mulrow, PhD Candidate, University of Illinois at Chicago, (Data Management and Analytics Co-Lead), is a member of UIC's Complex and Sustainable Urban Networks Laboratory, a team of researchers applying machine learning methods to civil and environmental engineering problems. He has extensive experience working in the compost and solid waste management industries.

Nina Kshetry, PE, President, Ensaras Inc., Champaign, IL (Data Management and Analytics Co-Lead), is founder and president of Ensaras, Inc., a company providing advanced analytics and artificial intelligence solutions for water resource recovery management. She holds SB and SM degrees in Environmental Engineering from MIT and has over a decade of water resource recovery experience. She currently serves as chair of the IWEA-LIFT committee.

Dr. Lav Varshney, Chief Scientist, Ensaras, Inc., and Assistant Professor, University of Illinois at Urbana-Champaign (Data Analytics Co-Lead), is a professor of electrical and computer engineering, computer science, and neuroscience. His data analytics and artificial intelligence research—together with innovative industrial and social applications—have achieved worldwide acclaim. He currently serves on the advisory board of the IBM Watson AI XPRIZE ($5 million).

Darshan Jain, MS Candidate, University of Illinois at Chicago, (Research Affiliate), is a data analytics and programming specialist in UIC's Complex and Sustainable Urban Networks Laboratory.

PLAN

Our goal was to use 2.5 years of H_2S sensor, environmental, and operational data to train an intelligent predictor of local odor complaints, for the purpose of implementing targeted odor prevention measures.

PROBLEM STATEMENT A key component in the MWRDGC Tunnel and Reservoir Plan (TARP) is the Thornton Composite Reservoir (TCR). The TCR covers an area of 90 acres, is 300 ft. deep, and has a 7.9 billion gallon storage capacity. The reservoir was completed in 2015 and provides flood protection benefits and water quality improvements to more than 500,000 people in 14 communities in the MWRDGC's Calumet service area in the southern Chicago region. The MWRDGC operates the TCR to store excess combined sewer overflow (CSO) until wet weather events cease and the water in the TCR can be pumped via a 5-mile long 30-ft. diameter tunnel to the Calumet Water Reclamation Plant (CWRP) for treatment. The plant staff at the CWRP operate the pumps and gates between the plant, tunnel system, and TCR to control the volume of water the plant receives from the system.

Soon after the TCR was operational, the MWRDGC received odor complaints from the surrounding neighborhoods, see Fig. 1. The exact source of odor creation, whether from the solids settling in the reservoir or within the tunnel system, has not been identified. In either case, the production and dispersion of hydrogen sulfide (H_2S) is believed to be strongly associated with odor complaints. See Fig. 2, which depicts odor production, dispersion, and human perception/behavior as three key steps.

Figure 1. Odor Complaint Word Frequency

Given the complex nature of the system shown in Fig. 2, H_2S sensor data must be analyzed together with other data on weather and TCR operations—such as TCR gate opening/closing, and water levels—to better predict odors from the large-scale reservoir and develop effective odor mitigation strategies.

CHARACTERIZATION OF THE INTELLIGENT WATER SYSTEM The MWRDGC installed a network of Odalog low-range (0.01 to 2.0 ppm) H_2S loggers around the TCR and surrounding areas in early 2016. The loggers take readings at 15-min intervals and store the data until it is manually downloaded every two weeks. From 2016 to 2018, two to six loggers were operational at any given time and were moved between seven locations around the TCR. The MWRDGC also regularly monitors reservoir gate operations and the tunnel and TCR water elevations. Finally, the MWRDGC's Incident Reporting System contains records of all odor complaints on file. The following is a summary of all data sources used for this project:

- H_2S concentrations from Odalog low-range loggers
- Log of public complaints from the MWRDGC's Incident Reporting System
- Calumet TCR and tunnel system water elevation levels and gate positions
- Meteorological data from the NOAA Climate Data Online database

PLAN An algorithm has been developed to combine odor logger data with data from other physical, hydraulic, meteorological, and chemical processes to predict episodes of odor occurrence and complaint in the TCR system. Our goal in this project is to develop an advanced warning system for odors that will be used by the MWRDGC to predict odor events at the TCR. We decided to use supervised machine learning to achieve this objective, using H_2S sensor, weather, and TCR operational data as training data, and odor complaints as labels. Time series of H_2S was also considered (see Appendix C), but dispersion effects of H_2S through the reservoir and surrounding neighborhoods would be hard to model with the limited number of sensors available, as would the human perception and social factors that prompt someone to call-in a complaint. Using machine learning to predict odor events provides the MWRDGC with *actionable insight* that is currently not available. Our approach to develop an algorithm for an advanced warning system for odors consisted of: (1) Data processing and Quality

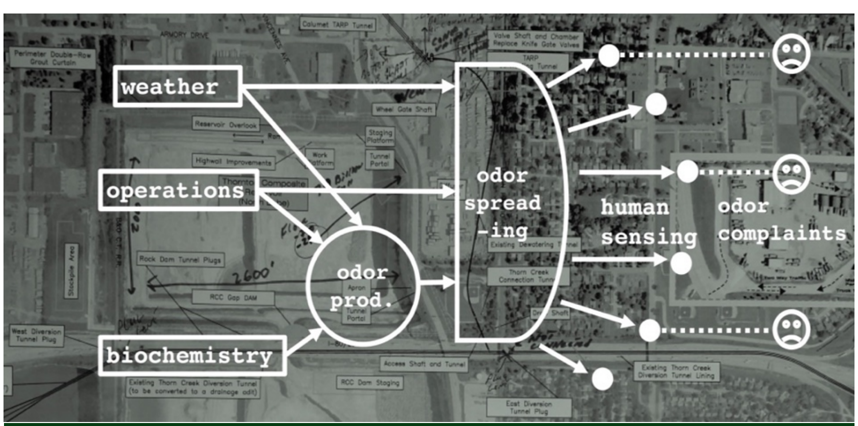

Figure 2. Odor production, spreading, and human perception.

Assurance/Quality Control (QAQC) from Mar.-May 2018; (2) Algorithm development from May-Jun. 2018; and (3) Algorithm testing and evaluation from Jul.-Aug. 2018. Further planned actions are to (4) Develop action plans and guide in technology selection to reduce odor and complaints in Sep.-Oct. 2018; and (5) Full-scale implementation and deployment from Sep. 2018 to Sep. 2019. Note that once the algorithm was tested, it was handed over to MWRDGC through an open-source and license-free software implementation using Python.

Beyond odor event prediction itself, algorithm development helped identify: i) highest priority locations for emissions of odorant gases and resource allocation for sensors, ii) correlation of H_2S concentrations with odor complaints, iii) threshold concentration of H_2S triggering complaints from the community, iv) attributes amongst all data sets most important in predicting odor complaints, and v) key insights for the development of odor mitigation strategies.

IMPLEMENTATION

Data pre-processing and exploration was followed by training and testing of a supervised classification algorithm. Results verify operational changes in H_2S sensing and odor management at TCR.

DATA With the goal of predicting an odor event with at least three days advance warning and suitable tradeoff between missed detections and false alarms, we trained several supervised learning algorithms. To do so, we compiled a data matrix with 872 observations (one observation for each data collection day between January 14, 2016 and July 2, 2018) and 312 quantitative attributes. A day level matrix was used since some data was only available at the day level, and therefore, other parameters were brought to the day level. The large number of data attributes for each day is due to the creation of several prior-day data attributes, and mathematical manipulations (e.g., minimum, mean, maximum values during a day or range of days), necessary for training a high performing advanced warning system. A key challenge faced during this project was the choice of a proper *target variable* that would adequately represent the occurrence of an *odor event*. Algorithm performance was ultimately tested on two different target variables: number of odor complaints logged through MWRDGC's odor hotline, and maximum recorded H_2S concentration. Past data compilation, target variable selection, and QAQC measures implemented are described in further detail below.

Past Data Compilation MWRDGC required a target advance warning period of at least three days, so it was decided to train the prediction algorithm on data from 3, 4, 5, 6, and 7 days prior as well as on summary data for the one and two-week periods starting three days prior.

Target Variable Selection Choosing a target variable to represent odor events was an important step in setting up a proper supervised machine learning analysis. There were three data types that could be used to represent the kind of odor event that MWRDGC desires to predict and prevent in the future: 1) Odor Patrol data, collected by MWRDGC staff from 2016-2017; 2) Odor Complaints, collected through the MWRDGC odor complaint hotline and provided by citizens reporting odor in surrounding areas; or 3) H_2S concentration levels, which analysis showed could serve as a significant proxy for sensible odor (but does not capture human perception/behavior or aspects of dispersion; see Appendix C). Each data type had pros and cons related to consistency, accuracy, and cost of future monitoring. These are described in Table A1 (Appendix A). Odor complaints were selected as the final target variable because they represent direct observations made by the general public, which the MWRDGC seeks to minimize.

QAQC Compiling disparate data types, formats, and frequencies into a single data matrix, with correct information for each day during the study period, was one of the main challenges faced during this project. Several months were spent processing and understanding the origins of the many datasets provided by the MWRDGC. QAQC steps were required to ensure the machine learning training data was unaltered in content from original data sources. Additionally, QAQC was required during algorithm training and testing. The following actions constituted our QAQC approach:

- **Spot-checks** of data files compiled from sub-spreadsheets (Odalogs and TCR elevation data) or from data in MS Word documents (Odor Patrol). Spot-checks conducted by QAQC Lead, J. Mulrow.
- **Use of shared Dropbox**, enabling version tracking and file editing history.
- **Regular email/phone check-ins** between Data Team (UIC and Ensaras) and MWRDGC staff.
- **Site visits to TCR** to observe sensor locations and TCR Operations; these visits also allowed the data team to speak with MWRDGC technicians about odor sensing and control experiences since 2016.
- **Use of Jupyter Notebooks**, a platform for sharing data analysis scripts among the team, for verification of methodology and results.

Further description of the data processing and QAQC steps taken for each data source is provided in Appendix B.

ANALYSIS AND INTERPRETATION Supervised machine learning techniques were used for understanding the mapping to odor complaints from the almost 2.5 years of data compiled from the many sources described above. Results were used to not only develop a deployable predictive algorithm, but also confirm and reveal relationships important to MWRDGC odor monitoring and reservoir management operations. Upon exploring several kinds of supervised learning classification algorithms, the random forest family of algorithms emerged as most effective. This section focuses on the results of random forest classification and regression, trained and tested on all attributes and subsets of attributes, using *Odor Complaints* as the target variable. It is important to note that the data attributes represent data measurements taken at least three days *prior* to the incidence of odor complaints.

Methods

Random Forest (RF) Classification & Regression is a specific type of supervised learning algorithm, in which an ensemble of decision trees trained on random subsets of data is used to make a final classification decision or quantitative prediction. In our case, binary classifiers were developed to predict odor and non-odor days, or regressors were developed to predict the number of odor complaints expected. This type of algorithm has several advantages beneficial in our odor prediction problem, including the ability to overcome missing data (as many of the MWRDGC datasets are incomplete over the study period), and to extract feature importance from multi-dimensional data.

Random Oversampling (ROS) is used to increase the number of observations used in creating decision trees within the random forest ensemble. This is required to deal with the *class imbalance* problem, since odor complaints were received on only 15% of the observation days included in the study period (129 of 871 days), but identifying these correctly is the basic premise of our investigation. By randomly over sampling from the training data to create an equal number of complaint versus no-complaint days in the training dataset, better supervised training results can generally be achieved. Indeed, four of the top five algorithms we present below make use of random over-sampling.

Attribute Subset Comparison was required to compare operational scenarios using various attribute subsets, so as to determine best resource allocation for implementing an advance warning system. Subset comparison provided information more applicable to operations than feature importance alone, where relevant features can become hidden if they are correlated with other important features. We use Receiver Operating Characteristic (ROC) curves to compare the performance of a variety of random forest algorithms applied to various data subsets. The area under these curves (AUC's) are used as a quantitative measure of the quality of each trained algorithm.

MWRDGC Performance Criteria that the algorithm should predict odor events with greater than 60% accuracy (true positives) and less than 25% false positives was established by MWRDGC through systems-level concerns for suitable operational impact. Algorithm performance was measured against this stated goal for an Advance Warning System for Odor.

Results

ROC Comparison Fig. 3 depicts ROC curves for all trained classifier and regressor algorithms (35 total), each based on a different subset of data attributes. Because the goal of an odor prediction algorithm is to maximize the prediction of true odor events (true positives) and to minimize the prediction of odor events on non-odor days (false positives), a desirable ROC is concave down with increasing slope and maximum area under curve (AUC). Thus, the best prediction algorithm will generally demonstrate a curve that passes closest to the top-left corner of the graph. Algorithms that meet MWRDGC's required performance criteria pass through the blue shaded area. The data subsets with best and worst-performing predictive power are clearly seen in green and red, respectively. This collection of curves demonstrates the importance of combining H_2S concentration data with weather and operational data from the MWRDGC in modeling and predicting odor events.

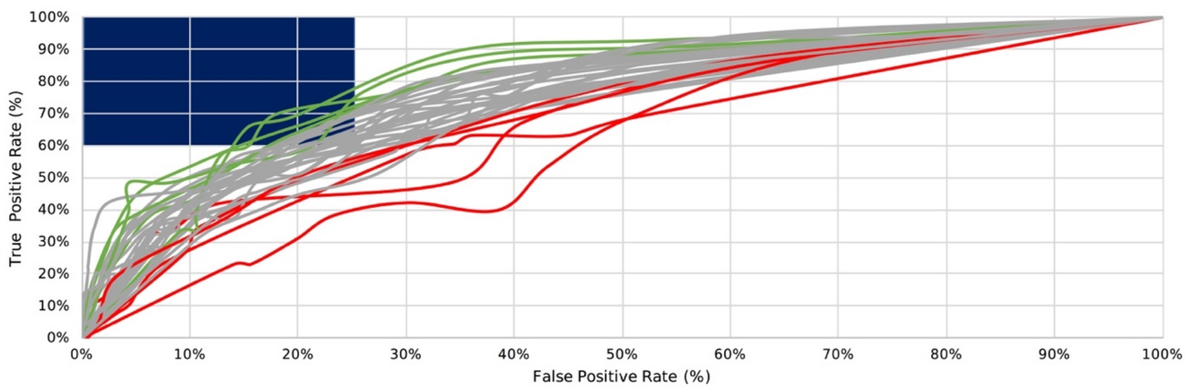

Figure 3. All data and data subset ROCs, with target accuracy shaded area

AUC Comparison A quantitative comparison of ROC curves is given by calculating AUCs, with a larger area representing better performance. When algorithms are ranked by AUC, we find the top performer is an RF Classifier with ROS, using weather and operational data together with H_2S data only from the NE, SE, and NW corners of the TCR. AUCs are compared in Table 1 It is seen that algorithms trained on a combination of at least three odalog locations, weather, and operational data result in the highest performing predictive algorithms.

The importance of H_2S sensors, weather, and TCR operational data in predicting odor events accurately, points to the complexity of this system, involving odor production, dispersion, and human behavior components. Given the MWRDGC has only three odalogs for ongoing monitoring, algorithms that use only three odalog readings were tested. In our results the three corner (NW, NE, and SE) locations slightly outperformed the drop shaft locations. The SW corner odalog was not considered as a location given the poor association with odor complaints.However, given that the drop shafts are where the highest H_2S readings are observed, it was decided that going forward odalogs will be kept at the dropshaft locations, therefore, the RFClassifier_ROS_DropShaft_All was the algorithm finally selected for implementation. This was the fourth best performing algorithm of all 35 tested. **The most important attributes in this algorithm are: Wet Well Drop Shaft H_2S, relative humidity, wind speed, wind direction, dry bulb temperature, and TCR elevation.**

Table 1. Five Best and Worst Performing Algorithms, by AUC

Algorithm	Data Subset Description	AUC
Highest AUC's		
RFClassifier_ROS_ThreeCorners_All	- Odalog data from NW, NE, and SE corners - All weather and operational data	0.826
RFRegressor_ROS_ThreeCorners_All	- Odalog data from NW, NE, and SE corners - All weather and operational data	0.812
RFClassifier_All	- All available Odalog, weather and operational data	0.801
RFClassifier_ROS_DropShaft_All	- Odalog data from each of the 3 drop shafts - All weather and operational data	0.796
RFRegressor_ROS_DropShaft_All	- Odalog data from each of the 3 dropshafts - All weather and operational data	0.792
Lowest AUC's		
RFRegressor_NoOdalogs	- All weather and operational data	0.705
RFClassifier_ROS_WeatherOnly	- All weather data	0.698
RFClassifier_ROS_DropShaftOnly	- Odalog data from each of the three dropshafts	0.682
RFClassifier_ROS_ThreeCornersOnly	- Odalog data from NW, NE, and SE corners	0.641
RFRegressor_ROS_NoOdalongs	- All weather and operational data	0.609

KEY INSIGHTS FROM DATA ANALYSIS AND ALGORITHM TRAINING AND EVALUATION

- Three-day prior Odolog, weather, and operational data are able to predict odor events with true and false positive rates that exceed MWRD's minimum standard.
- Using all categories of data resulted in high-performance odor complaint prediction. H_2S readings alone, or weather/operational data alone predicted odor complaints with lower accuracy.
- Data from a three-sensor system will be able to exceed accuracy criteria.
- Important features include wet well drop shaft reading, relative humidity, wind speed, wind direction, dry bulb temperature, and TCR elevation level.
- The high performance of the drop shaft-based algorithm, along with operational considerations, supports the decision to implement ongoing monitoring at the three drop shafts.
- Trained algorithms exceed the MWRDGC accuracy requirement, achieving over 75% true positive and less than 25% false positive identification rates

COMMUNICATION AND USE The algorithm's original purpose was to provide a three-day advance notice of possible odor events to the MWRDGC, triggering the dosing of an odor-mitigation treatment at key sites throughout the TCR and Drop Shaft tunnel systems. The algorithm was meant to ensure efficient deployment of an expensive treatment system. Mid-way through the project, MWRDGC ruled that even given high-efficiency dosing schedules, a chemical treatment would be too costly at the TCR. While this changed our implementation scheme, the predictive algorithm still provided value and improvement to the odor control system at the reservoir. The predictive algorithm will now be used to evaluate odor-mitigation strategies that are implemented at the TCR, as well as deployed at other reservoirs connected to the tunnel system that are believed to be causing odors.

Odor Control Implementation Plan The MWRDGC is now planning the installation of an activated carbon filter in the wet well dropshaft on the northeast side of the TCR to absorb H_2S before it can be dispersed to neighboring communities. The algorithm and analysis results strongly suggest the drop shafts as the source of H_2S, and the wet well drop shaft specifically as an important feature. The sensors at the corners of the TCR are detecting elevated concentrations on days when dispersion of the fugitive odors is poor; this would explain why drop shaft and corner locations are both good predictors of odor complaints. The algorithm will be used to compare the prediction of odor complaints with actual odor complaints received following the installation of this odor mitigation strategy. For effective odor mitigation, predicted odor complaints should be higher than actual complaints. This implementation scheme is described in Fig. 4.

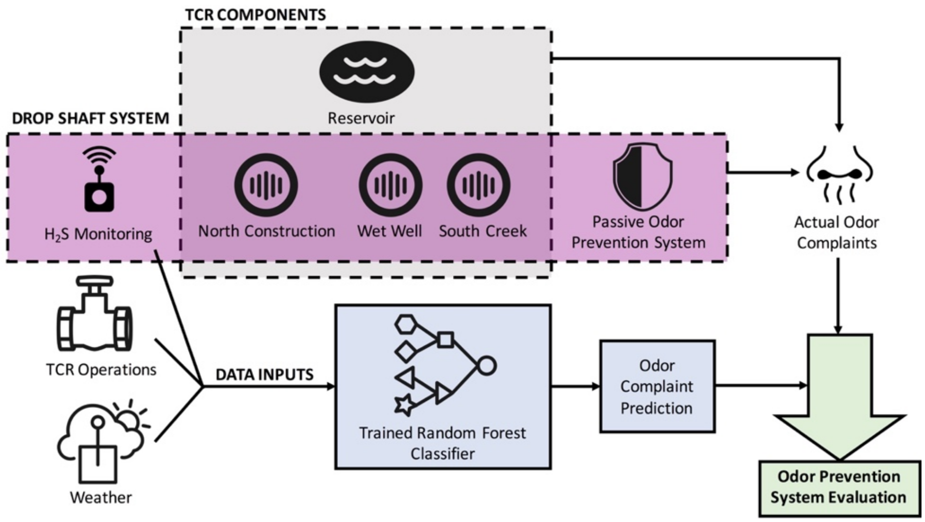

Figure 4. Algorithm deployment within the updated odor control system at TCR drop shafts

To address dispersion of future H_2S concentrations emitted from the reservoir, the MWRDGC will conduct Large Eddy Simulations at the corners of the TCR to study how trees (number and orientation) may impact dispersion of H_2S. Trees mitigate odor both by taking up sulfur compounds and hosting microbial communities that attenuate odor.

Data Collection Improvement Plan The MWRDGC is exploring other sensor technology that allows for wireless telemetry of the data to a remote location. The manual downloading of the Odalog sensors every two weeks was recognized as a cumbersome data management practice. Starting in August 2018, two Altech sensors, one with a detection limit of 10 ppb H_2S and the other at 30 ppb H_2S, are being piloted and compared to Odalog sensor data to determine the feasibility of discontinuing the Odalog sensor technology for prediction of odor complaints by the algorithm.

Deployment Outside of the TCR The odor prediction approach and data management improvement plans presented here are not specific to the TCR; this supervised machine learning approach can be implemented at other utilities that have similar reservoirs and tunnel systems. The MWRDGC itself has plans to implement this odor prediction and data management approach at other reservoirs that are connected with the tunnel system.

CONCLUSION

This project confronted many of the problems that modern water resource recovery treatment facilities must tackle in mitigating odor emissions to their surrounding communities. The team was challenged to handle a variety of data types and qualities, as well as combine data-driven machine learning approaches with the MWRDGC's expert knowledge of odor production and dispersion. Finally, the team developed the next stage of odor control measures at TCR, resulting in improved data sensor technology, better data handling, and deployment of a trained odor prediction algorithm for verifying the effectiveness of the new odor prevention scheme. We believe these efforts are relevant and informative for every water resource recovery utility that manages tunnel and reservoir systems. More broadly, our work provides a concrete example of how artificial intelligence and machine learning can be used along with sensor and operational data to solve important water resource recovery industry problems.

APPENDIX A

Data Compilation & Target Variable Selection Figures

Table A1. Target Variable Pros and Cons

Data Type	Pros	Cons	Used in algorithm training?
 **Odor Patrol**	• Daily **in-person** observations • Made at **consistent locations** around TCR • Human sensor able to identify **odor characteristics** (rotten eggs, sewage, oil, etc.)	• **High cost** of sending employees out on a daily basis • **Inconsistency** of odor patrol personnel • MWRDGC employees potentially **less sensitive to odors** (ie. biased observations)	✗
 **Odor Complaints**	• Registered odor complaints are **direct observations** of public disturbance • Location of complaint contributes **geospatial information**	• **Subjectivity** of data. Some residents are more likely to take time to call and complain • **Inconsistency**. Not all odor events get reported and non-odor days are not registered at all.	✓
 **H_2S Concentrations (Odalog sensors)**	• **Consistent** and known sensor locations • **Quantitative** measure • **Reliably logged** in usable spreadsheet format	• Assumed as a proxy of odor; **No human observation** • **Stationary** observations • **Unable to capture wide area** surrounding TCR	✓

APPENDIX B

QAQC Descriptions

Table B1. Data Attributes, Processing and QAQC Summary

Data Attribute	Source	Description	Processing & QAQC
H_2S Concentration	MWRDGC provided odalog summary spreadsheets compiled at 2-week intervals	***OdaLog® Low Range H_2S Loggers***: 2-6 sensors were positioned around the TCR reservoir from Jan 2016 to June 2018. Each sensor takes an aerial H_2S reading every 15 minutes. **Primary sensor locations:** • TCR corners: NE, SE, NW, SW • Tunnel drop shafts: N. Wet Well, N. Construction, S. Creek	• 15-minute interval data was compiled for every day and every location into a single large array. • The compiled file was checked for duplicate entries, out-of-range H_2S readings, and correct location. • Day-level summary of H_2S readings was compiled using Max and Mean readings at each location.
Precipitation	NOAA weather station	Daily precipitation records	• Data units confirmed as American units (inches, Fahrenheit, etc.)
Temperature, Humidity and Wind Speed/Direction		Hourly temperature, humidity and wind speed/direction readings	• Day-level summary of weather readings was compiled using Min, Max, and Mean readings.
TCR Elevation	MWRDGC provided spreadsheet with data beginning Oct 2016.	Daily reading of reservoir water elevation	• Confirmed no readings showing lower than -306 CCD (Chicago City Datum)
TCR Tunnel Gate Position		Position of the two gates leading from TCR to the tunnel leading to Calumet WRP. Daily summary of the % of time gate was open. (e.g. 0.5 = gate open for half of the day)	• Translated text records (e.g. OPEN, CLOSED, X% OPEN, etc.) into numerical values, 0.0 – 0.1. • Created summary variable: Average of Gate 1 and Gate 2 positions.
TARP Tunnel Elevation		Elevation of water in TARP tunnel, behind the gate leading to TCR.	• Confirmed no readings outside the expected range of TARP depths.
Odor Patrol Observations	MWRDGC provided Microsoft Word document records	Daily odor observations, recorded by MWRDGC Technicians on-site. Observations made at each corner of the TCR and at tunnel drop shafts.	• Data held in separate Word documents was transferred to spreadsheet format • Spreadsheet spot-checked for alignment with original documents
Odor Complaints	MWRDGC provided spreadsheet log	Record of all citizen-reported odor issues occurring between January 2016 and June 2018. Reports are made through citizen hotline and online incident reporting system.	• Complaint location checked to confirm proximity to TCR. • Data summarized at the day-level, showing number of total complaints each day during study period.

APPENDIX C

Additional Analyses

Figure C1. Percent of days with H_2S readings above and below 0.1ppm threshold
a) Among days with 1 or fewer complaints, b) Among days with more than 1 complaint

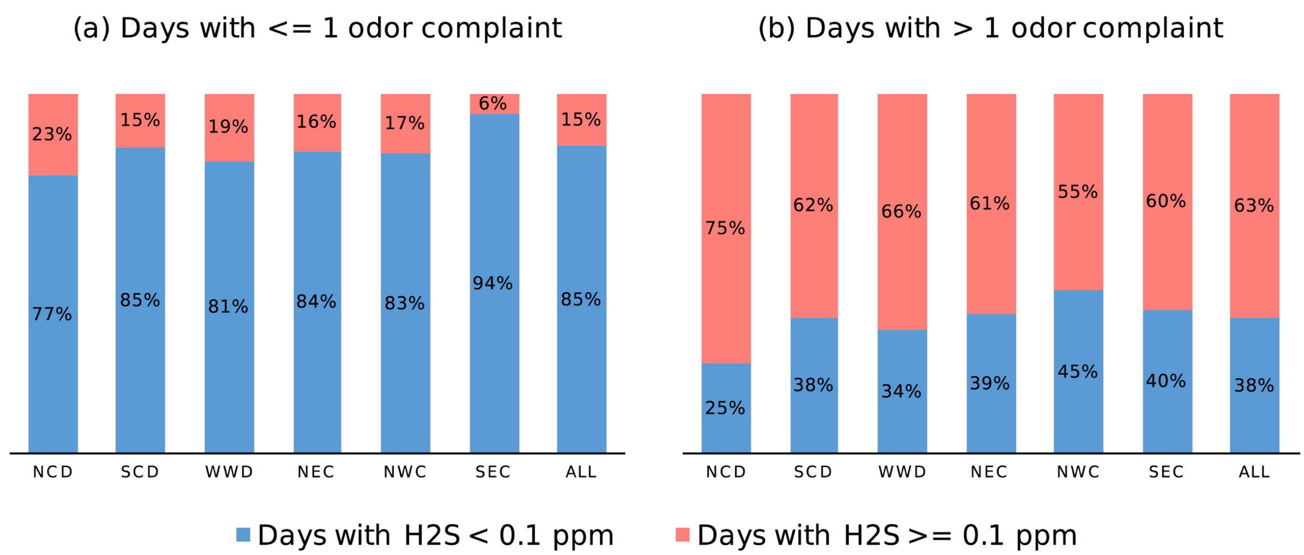

Figure C2. H_2S Readings At Each Location – Time Series