# Concentration of Toxic Chemical in Local Water System after Fire Seasons

Wanwei Huang

## 1 Introduction

Extended from the analysis of wildfires in the city Alamogordo, NM, we will primarily examine how the city's water system, serving more than 35 thousand people, is affected by wildfires in this analysis. These wildfires have the potential to produce toxic substances, including particulate matter, trace minerals, carbon oxides, organic chemicals, and nitrogen oxides. Since smoke from fires can be carried through the air, the city is vulnerable not only to wildfires within its vicinity but also to those occurring at a distance outside the city. Even though the poor air quality may only persist during the fire season, and the number of fires in nearby areas may not be sufficient to raise immediate concerns, we must acknowledge the possibility that byproducts could linger in various ways in our daily lives, particularly in our water supply. Therefore, it is crucial to address this issue seriously and take precautions as early as possible.

## 2 Background & Related Work

### 2.1 Background

Through water, the direct byproducts of wildfires persist as small-sized particles enter the water undetected, and toxic gasses dissolve into it, altering the concentrations of various chemical compounds. Simultaneously, less direct byproducts, such as Volatile Organic Compounds (VOC), emanate from materials susceptible to fire or heat, including burned plastics, melted cables powering water systems, and overheated surfaces of water tanks. Numerous studies highlight that these chemical compounds could elevate the risk of cancer through prolonged exposure. Consequently, we will investigate how the concentration of certain chemicals changes after fire seasons. Narrowing down the scope of chemicals to the ones most related to wildfires (Pennino et al.), we recognize a few that are available from the source of data. To begin our analysis, we examined the concentrations of different compounds that enter the water through various pathways. These include Haloacetic Acids (HAA5), which enter as VOC, and Nitrate (Ni), which predominantly enters through the air.

### 2.2 Related Works

#### 2.2.1 HAA5

To comprehensively grasp the potential harm HAA5 could inflict on the human body, it is essential to understand that Volatile Organic Compounds (VOCs) constitute a group of substances that not only readily vaporize and dissolve in water but also adhere to the soil, thereby contaminating groundwater—the primary source of drinking water in the U.S. (Peeples, 2020). Concurrently, statistics indicate a close association between VOCs and an increased risk of kidney cancer, leukemia, and other health issues. Notably, the link between wildfires and VOCs was only recently established (Peeples, 2020).

As a subset of VOCs, HAA5 comprises disinfection byproducts containing five types of acids. It has the potential to cause skin irritation and corrosiveness in the short-term if highly concentrated, as well as metabolic changes, contributing to an elevated risk of liver and bladder cancers (NHDES, 2018) in the long-term .

Furthermore, studies indicate that the concentration of HAA5 rises after wildfires, though the duration of surface- and in-water concentrations varies (Pennino et al.). Therefore, it is imperative to investigate the correlation between wildfires and HAA5 concentration, and in the future, exploring data spanning more than one year could yield valuable insights.

### 2.2.2 Nitrate

While Nitrate has been one of the most impactful contaminants in the U.S., consistently posing higher risks of methemoglobinemia (Pennino et al.) and various cancers, including colon, kidney, stomach, and specifically thyroid and ovarian cancers among women (National Cancer Institute), the post-fire concentration of Nitrate warrants more careful investigation due to its potential health implications. However, it has been noted that, despite an increase in the concentration of Nitrate in water systems after wildfires in specific sites, there is also a rise in Nitrate concentration in control sites. Consequently, fully distinguishing the reasons behind the elevated Nitrate concentrations may prove challenging (Pennino et al.).

Nevertheless, considering the expansive area under scrutiny, spanning 1250 miles in radius, it is worth considering how Nitrate might travel in the air and impact control sites—areas devoid of any recorded wildfire events for more than three decades (Pennino et al.). Therefore, we will continue to incorporate the concentration of Nitrate in our ongoing analysis.

## 2.3 Datasets

As an extension plan of the wildfire investigation to assess water quality, we accessed the New Mexico Environmental Public Health Tracking (NM-EPHT) website and identified various indices available and also reused the wildfire dataset and the AQI datasets. Specifically, the Community Water Systems Water Quality index provides data on the population served by community water systems, water contaminants by systems, and concentrations of chemical compounds in the water systems. To narrow our focus to data within fire seasons, we opted for the dataset with the smallest granularity, which is the quarterly data. However, there are limited datasets with quarterly information.

Looking into current research and trying to decide which datasets to use in the analysis, the conclusions drawn about the upwards trends of the concentrations of certain chemical compounds in the water after fire seasons (Pennino et al.) led to the choice of both Nitrate and HAA5. This decision also aligns with our earlier sections where we aimed to investigate compounds with different entry methods into the water systems.

# 3 Methodology

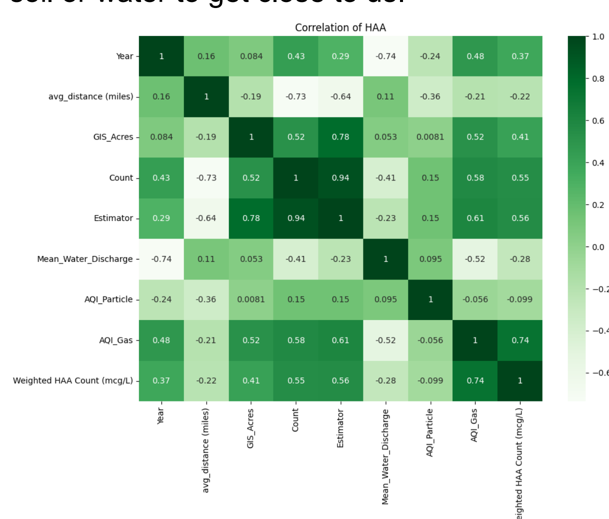
## 3. 1 Weighing Quarterly Concentrations

As mentioned earlier, we opted for the quarterly dataset due to its finer granularity. However, to specifically focus on fire seasons (May 1st to Oct 31st), we applied a weighting system to extract relevant data. Assigning weights to each quarter, we designated the first quarter as 0 out of 6, the second quarter as 2 out of 6 (including May and June, which are in the fire season), the third quarter as 3 out of 6 (entirely within the fire season), and the last quarter as 1 out of 6 (since only October is in the fire season). By summing up the weighted concentrations from different quarters, we obtained an aggregated annual data set, aligning with the annual data of fires and Air Quality Index (AQI).

## 3.2 Correlation Analysis

Looking at the correlations between the concentrations and the features from the wildfires, it was found that annual counts of fires, avg_distance (miles), GIS_Acres and the gaseous AQI have significant numbers there.

While Nitrate has positive correlation with count and gaseous AQI,and significantly negative correlation with avg_distance (miles), it might be due to that Nitrate could not be carried too far and the concentration will enhance when more fires happen close to the city. As for the concentration of HAA5, there are quite positive correlations with gaseous AQI, GIS_Acres and count of fires while less significant in the negative correlation with avg_distance (miles). This might indicate that, disregarding the distance, the larger the area it burns, the more toxic compounds go into the soil and travel with soil or water to get close to us.

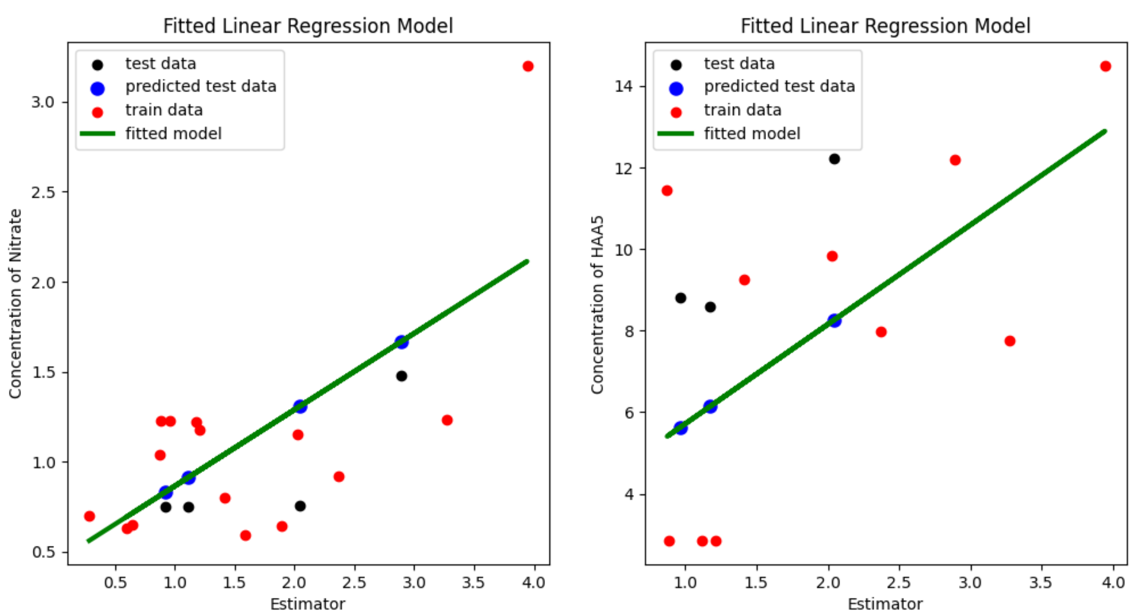
## 3.3 Estimator Design

In light of the insights gained from the correlation analysis, we opted to incorporate a select few features exhibiting high correlations with the concentrations of HAA5 and Nitrate when designing the estimator. It's important to note that, given the current lack of comprehensive information about other potentially influential features and to prevent possible overfitting in model construction, we have chosen to employ identical estimators for both chemical compounds. Furthermore, while acknowledging that AQI data updates later than the occurrence of fires, we decided not to include it in the estimator.

To underscore the significance of the impacts stemming from wildfires, we introduced a squared count of fires multiplied by the square root of GIS_Acres. This formulation anticipates that the effects of the burned area will diminish as it extends further. Subsequently, we divided this product by the average distance (in miles) to account for the square root of GIS_Acre and further normalized the result by dividing it by 10 to the power of 7, ensuring the values remain within a manageable scale.

## 3.4 Simple Linear Regression

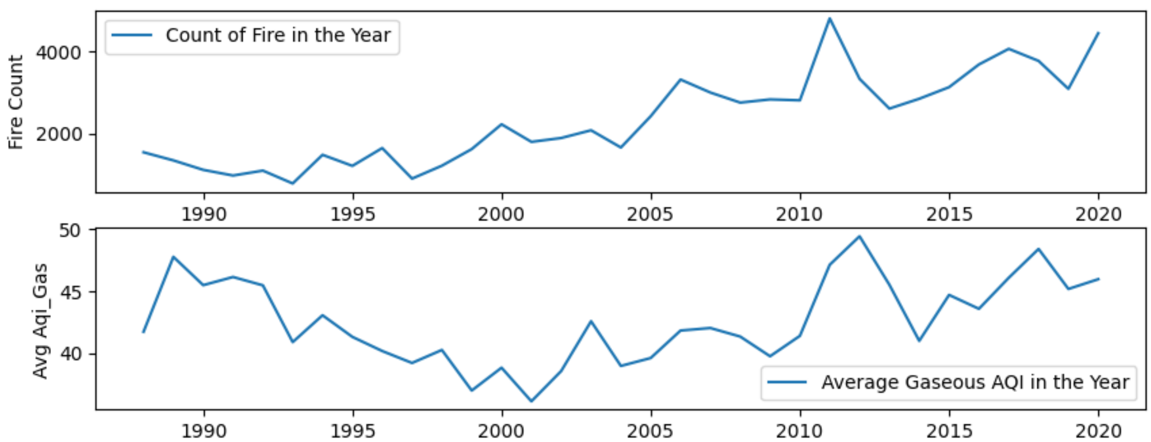
To facilitate a clearer understanding of the relationship between the selected features and the concentration of compounds in the water system, we amalgamated these features into a single estimator, as outlined earlier. Next, we employed a simple linear regression as the baseline and plotted the fitted models. However, the plots reveal that the fitted results are far from satisfactory, indicating the need to explore more sophisticated models or estimators. It is imperative to address the limitations outlined in the subsequent section for future analyses.



# 4 Findings

From the analysis, we could see the correlation between concentrations of the two compounds and features from wildfire, and a few most significant features are Acres, Average distance, counts of fires and the Gaseous AQI. So a few hypotheses from the correlation analysis are that the counts of fire impacts the concentration of both highly, area plays a greater role in bringing more VOC into the water system and the shorter distance allows toxic gasses to be carried to the city in greater quantities.

Then, the two simple linear regression models that applied the estimator and the hypotheses to predict the concentration of compounds revealed that …. Even though the R2 scores are far from ideal, we could still see a rough trend there. In the future, different estimators might be necessary for different methods how the compound might get into the water




# 5 Discussion/Implications

Considering the strong correlation identified through correlation analysis, linking the annual count of fires to the concentration of toxic chemical compounds in the water system, coupled with the observed increasing trend in the count of fires near the city, it is imperative to implement measures aimed at reducing the incidence of fires. Additionally, enhanced water purification strategies should be employed to effectively address the escalating situation. Imposing a fire ban and implementing robust water purification processes post-fire seasons are crucial steps in managing and mitigating the potential impact on water quality. However, while the concentrations of both Nitrate and HAA5 are far from the Maximum Contaminant Level (MCL) (US EPA), the city could take its time to implement what's mentioned here. In the meantime, the analysis should be carried forward and further look into the trends of concentrations of different contaminants and evaluate if there is imminent need to take action

Last but not least, as discussed later in Section 6 (Limitations), we strongly recommend the establishment of an Air Quality Index (AQI) capture site within the city. This initiative is essential for obtaining more accurate and localized data, thus improving the precision of our analysis and enhancing the city's preparedness for potential environmental challenges.

# 6 Limitations

## 6.1 Data cleaning in the Fire Dataset

Considering the various categories of fires (prescribed, likely prescribed, wildfire, etc.), determining the specific impact of fires from each category poses a challenge. Hence, incorporating indicators of fire categories may be essential when constructing new models in the future. Simultaneously, while the current grouping of fires is presented as annual data, exploring finer granularity, such as examining specific months within the entire fire season, could provide valuable insights for more nuanced analyses.

## 6.2 the site of AQI

As Alamogordo, NM lacks its own site for capturing Air Quality Index (AQI) data, the data used in this analysis is sourced from nearby cities. Despite our expansive scope with a radius of 1250 miles, relying on data from neighboring cities may introduce potential inaccuracies in the model predictions.

## 6.3 the granularity of concentration data

In the initial phase of the analysis, an attempt to extract the start date of each fire using the "Listed_Fire_Startdate" revealed inaccuracies in some dates, and certain fires did not align with fire seasons. Consequently, rather than pursuing finer granularity, the fire data was later grouped by year in the subsequent analysis. Simultaneously, the Air Quality Index (AQI) datasets were also grouped by year, with only data during fire seasons being extracted.

In anticipation of observing the impact of fire seasons on the water system and the concentrations of chemical compounds in the water, the plan was to exclusively utilize Water Quality Index (WQI) data from fire seasons. However, the smallest granularity available for WQI datasets is quarters. Therefore, we opted to attribute weights to different quarters, group the data by year to simulate fire season data, and subsequently merge it with other datasets.

## 6.4 the duration of compounds in the water

To assess how wildfires have altered the Water Quality Index (WQI) for the year, it is crucial to consider potential lagging effects of chemicals in the water post each fire season. While determining the precise duration that compounds linger in the water system proves challenging, our current focus is on the concentrations of chemical compounds during each fire season. Ideally, our analysis should extend beyond just the WQI during the fire seasons, covering the entire year from the commencement of the current fire season to perhaps the onset of the subsequent fire season. Evaluating time spans exceeding one year may also be necessary to investigate the duration of compound persistence. Additionally, as we assess the

WQI for this period, it is crucial to attribute a portion of the concentration to the impact of last year's wildfires.

It should be noted that, the reason why this analysis chooses not to promptly look into the possibilities are mainly due to the lack of useful data and the lack of evidence to support the duration of the chemicals in water systems in recent research (Pennino et al.). However, it should also be stressed that it might be useful to look into cities close to Alamogordo, NM to see if those locations have enough data to support further analysis.

## 6.5 how to update the estimators for each case

Upon examining the built models, we observed that certain correlations between Water Quality Indices (WQIs) and the features of the fires are relatively stronger than others. While we anticipate that different chemical compounds enter the water system through distinct pathways, we also expect the impact of fires on WQIs for chemical compounds to vary based on their entry mechanisms. Currently, we maintain the same estimators for both models. However, for enhanced performance and specificity for each chemical compound, separate estimators should be developed to accommodate their unique characteristics.

Simultaneously, it is essential to incorporate additional data on other chemical compounds into the analysis. This allows us to examine whether, when entering the water system in the same manner, the concentrations of these compounds behave similarly. Only through inclusive cross-referencing with more data can we generalize the analysis to broader categories of compounds. With an expanded dataset, we aim to apply the same models to chemical compounds within the same categories, enabling us to make generalized predictions as more data become available.

# 7 Conclusion

In conclusion, this analysis found no urgent need to impose any action at this time, but further analysis should be carried forward to see if there exists any toxic chemical compound that requires immediate treatment since the count of fire and gaseous AQI are impactful for both models and both exhibit growing trends. Only after those could we decide whether the city should be placing more purification in water systems now to prevent future long-term exposure to toxic chemicals and to plan for future fire seasons. Further, considering that the AQI data comes from nearby cities, the city council should consider implementing a data collection center in the city for future investigation. Last but not least, more frequent sampling of water quality should be performed for more finer granularity and more concrete analysis in the future.

# 8 References

Haloacetic Acids ARD-EHP-36 - NH Department of Environmental Services, New Hampshire Department of Environmental Services, 2018, www.des.nh.gov/sites/g/files/ehbemt341/files/documents/ard-ehp-36.pdf.

Nitrate, National Cancer Institute, progressreport.cancer.gov/prevention/nitrate. Accessed 7 Dec. 2023.

Pennino, Michael J, et al. “Wildfires Can Increase Regulated Nitrate, Arsenic, and Disinfection Byproduct Violations and Concentrations in Public Drinking Water Supplies.” The Science of the Total Environment, U.S. National Library of Medicine, 15 Jan. 2022, www.ncbi.nlm.nih.gov/pmc/articles/PMC10084414/.

Peeples, L. (2020, December 21). The surprising connection between West Coast fires and the Volatile Chemicals Tainting America’s drinking water. Ensia. https://ensia.com/features/volatile-chemicals-vocs-drinking-water/

United States Environmental Protection Agency. (n.d.). *National Primary Drinking Water Regulations | US EPA*. National Primary Drinking Water Regulations. https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations

## 9 Data Sources

Health, Department of. “Index of Queryable Environmental and Health Measures.” NM, nmtracking.doh.nm.gov/dataportal/query/Index.html. Accessed 15 Nov. 2023.

USGS Water Data for the Nation, nwis.waterdata.usgs.gov/nwis. Accessed 15 Nov. 2023.