# Trends in Toxic Releases over Time in US Industry

*Chase Foster*

*February 18, 2020*

## Introduction

In the United States, much attention is being paid to environmental issues in politics, industry, and people's everyday lives. Awareness of these issues has resulted in efforts to decrease the negative impact our society has on the environment through individual and legislative action. In 1970, the Environmental Protection Agency (EPA) was created under President Nixon to implement protective regulations to limit damages caused by various industries to the environment. One program the EPA uses to facilitate this goal is the Toxics Release Inventory (TRI), created in 1987 as a response to several chemical releases that killed thousands of people in the US and India. This program requires companies to submit records to the EPA detailing the amounts of chemicals released, stored, recycled, transported, and processed each year. These data are openly available to the public and can be accessed easily online through Data.gov and the EPA website.

For this analysis, I worked to uncover trends in toxic chemical releases over time. The TRI program provides a substantial amount of industrial data that allows the public to examine how policy decisions have impacted toxic releases across the US. The contribution of this program towards safer industry can be quantified by analyzing TRI data over several years and demonstrating changes in toxic release patterns over this period. This process can also be used to evaluate the efficacy of similar policy aimed at reducing environmental harm in industry.

In this report, I analyze TRI data to reveal trends in toxic chemical releases over time in the US. I use simple statistical methods to quantify change over time, and I provide easy-to-follow visualizations to illustrate these methods. I provide appropriate context to the decisions I make and the conclusions I draw, and I include a discussion of the limitations that these methods have for making predictions about future trends in industrial toxic releases. All the code I used to create this report is available in the attached GitHub directory.

## Data Collection and Cleaning

The TRI data used in this report are available for download from the EPA database, which I accessed online here. The relevant data for US-wide toxic releases are listed in the TRI Basic files at the federal level. These can be downloaded in CSV format for personal use. In this report, I use the TRI Basic data from 2003-2018, each downloaded separately. Historically, the EPA implemented the TRI Program in 1987 and has collected industrial toxic release information from that year to present. However, the TRI Program has undergone significant changes in policy and data classification over time. Major changes to their collection methodology occurred in 1991, 1998, and 2000 (page 6 shows these different milestones). In 2002, the EPA adopted the current classification regime used today in the TRI Basic data, and these changes were fully implemented by the 2003 collection year. In light of these changes, I restricted data collection to 2003 onward to match the consistency of the program. While this decision limits the amount of data used, it still provides a 16-year window of data on over 50,000 companies each year, allowing for statistically sound conclusions to be drawn from the population of companies.

The most pertinent TRI data for the purposes of this report are those containing information regarding the total amount of toxic releases into the environment reported by each company. These data are collected in the "On-Site Release Total" column of each file. Companies are also required to report the amounts of toxic chemicals processed in through other means (e.g. off-site recycling), but chemicals processed in these ways are not classified as fugitive releases into the environment. I processed the on-site data and created several complementary CSV files containing information pertaining to the mean, median, and total releases by US State, region, and the US at large. These files are included in the GitHub directory.

For region data, I grouped the States according to the US Census definitions of NE, S, MW, W. The TRI program also collects data on US territories, specifically Guam, Virgin Islands, District of Colombia, American Samoa, Northern Mariana Islands, and Puerto Rico. I excluded these locations from the region-level analysis, but I included them for US trends at large for completeness. US territories tend to have fewer companies than a given US state, so the amount of toxic releases therein is expected to be small relative to the US States.

Additionally, while most release data are reported in pounds, certain chemicals called dioxins (and dioxin-like compounds) are reported in grams. These chemical releases are usually much smaller than releases of metals and other compounds, and they have a different impact on the environment and human health than other releases. For example, Dioxin compounds have a proven carcinogenic effect in humans, so it is important to minimize releases of these compounds as a matter of public health. Nevertheless, this report focuses on toxic releases in aggregate, not at the granular level. So I converted all releases to pounds for consistency, but I acknowledge that this may gloss over important trends in dioxin releases beyond the scope of this report.

## Data Analysis

The first order of business is to see how the average amount of toxic releases in the US is changing over time. To quantify this pattern, the mean amount of toxic releases in the US each year is plotted in fig 1. Additionally, a simple least-squares linear model is superimposed in blue with a corresponding 95% confidence interval (CI) for the model shown in gray. We see immediately that the trend is very weak ($R^2 = 0.001$), and the width of the confidence interval indicates that there is high uncertainty in the model. Projecting these results forward would likely result in inaccurate predictions.

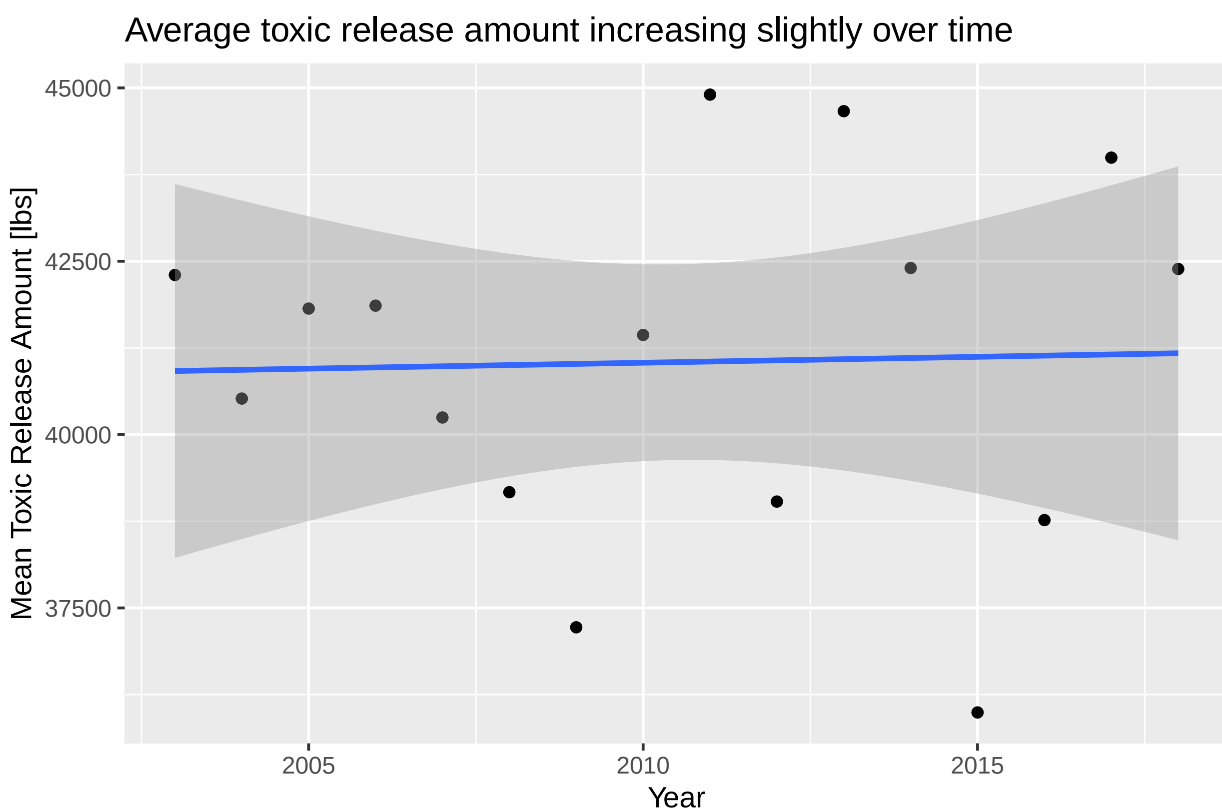


Fig 1

Digging deeper, I plotted the average releases per year by region in fig 2.

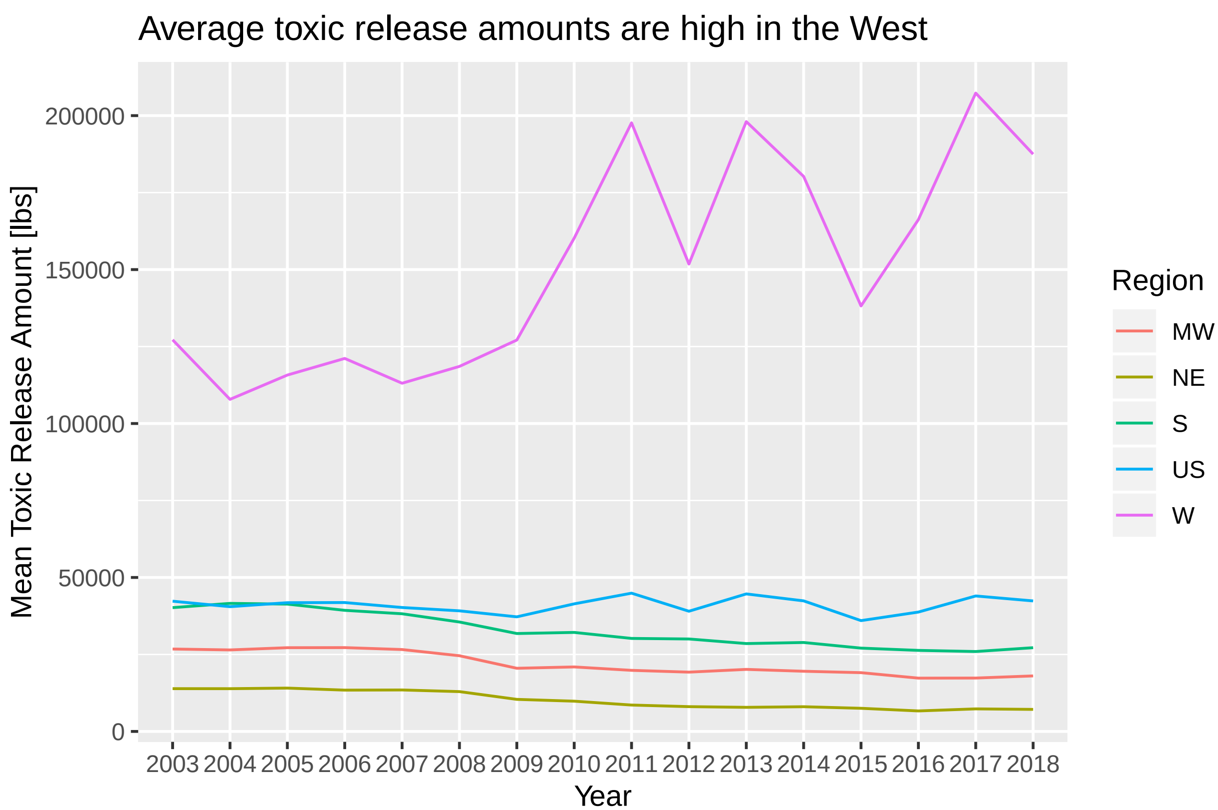


Fig 2

Here, we see that the average amount of toxic releases is considerably higher in the West each year than in the other regions. The same country-wide average release amount is shown for comparison. We can also observe the trend of average toxic releases in each region by separating each region out. The results of this process are shown in fig 3. Note that the vertical axes are different for each region because of differing ranges of toxic releases in each.

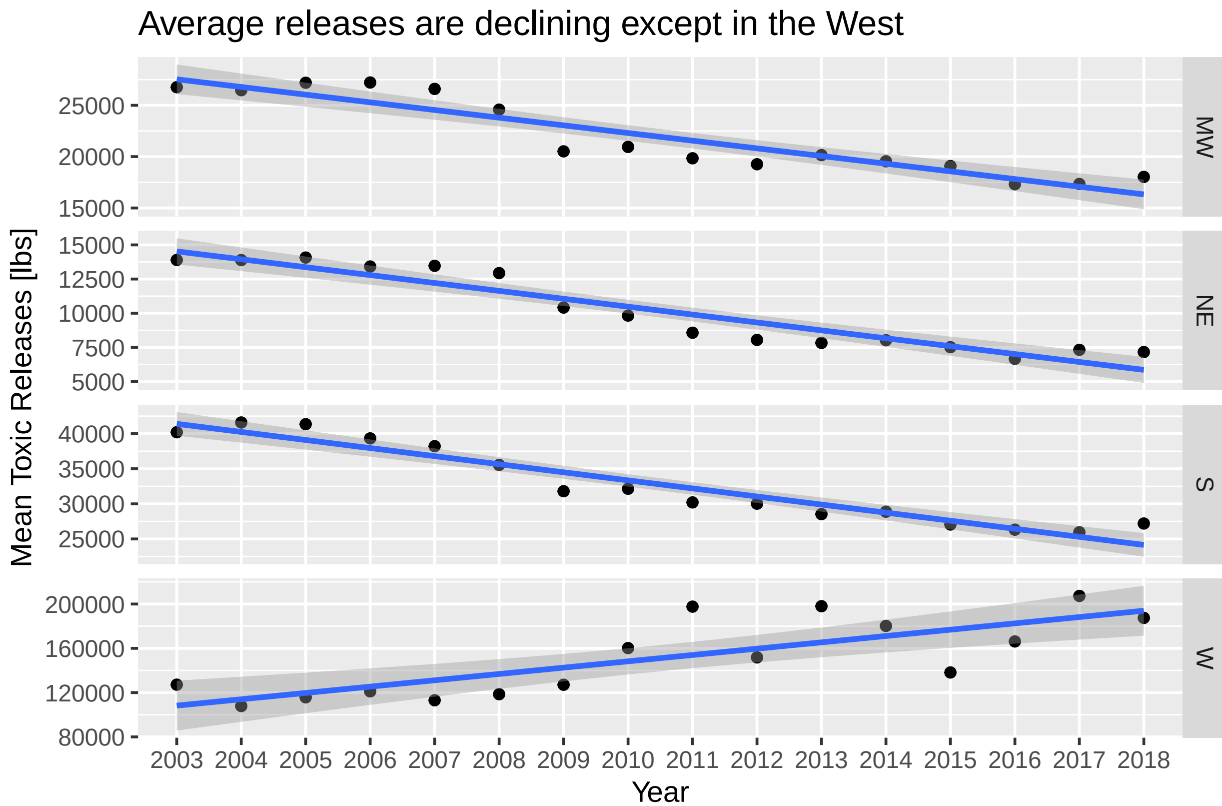


Fig 3

In each region except the West, the average amount of toxic releases each year is tending downward, while the same quantity is tending upward in the West. It is clear that some companies in the West are buoying up the average for the country as a whole. But in which States are these high-release companies present? To answer this, I plotted the 95th percentile of total toxic releases per year by State in fig 4.

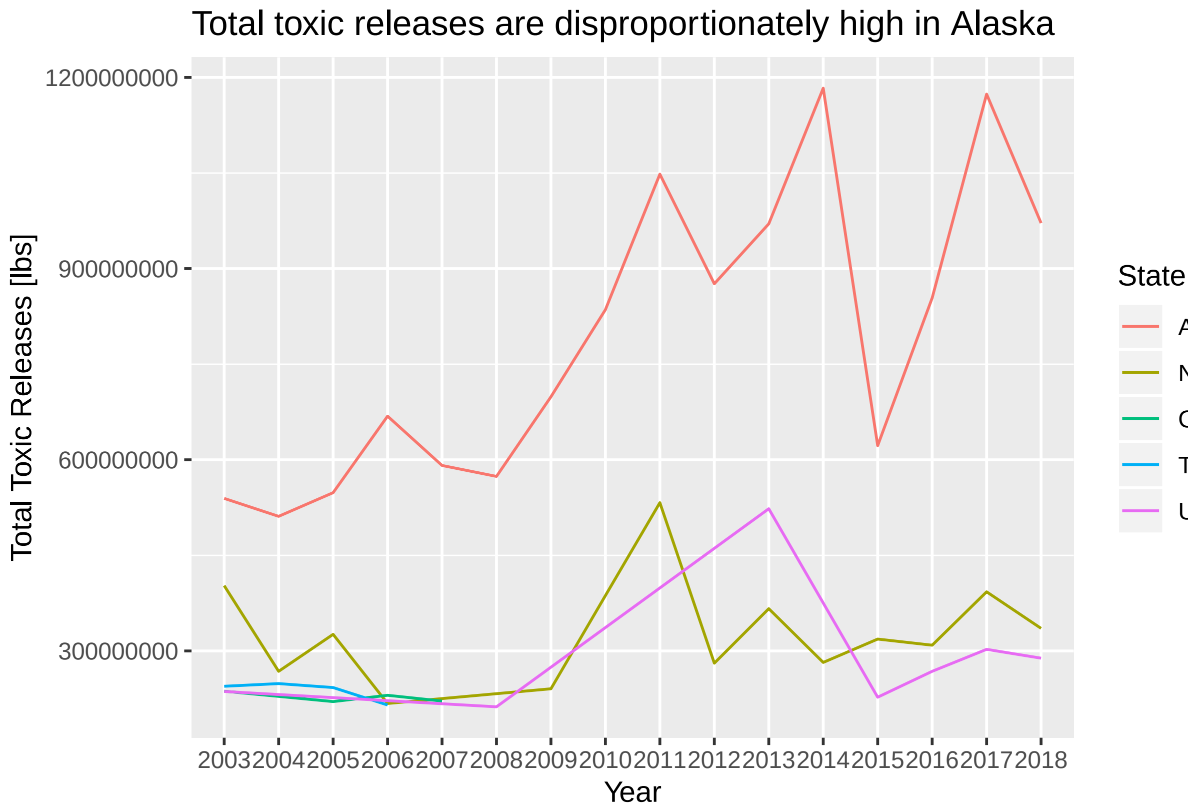


Fig 4

These data indicate that companies in Alaska are clearly responsible for the highest total amount of toxic releases in the county, despite brief peaks in total releases in Nevada in 2011 and Utah in 2013. In fact, the pattern of total toxic releases in Alaska is similar to the average amount of toxic releases in the West observed in fig 2.

So what is happening in Alaska? To answer this question, I looked at which companies in Alaska produced the most toxic releases each year. In every year in the dataset, a company listed as "Red Dog Operations" produced the majority of toxic releases in Alaska. This company was listed several times each year, meaning they produced several types of toxic releases annually. After combining the different toxic release amounts per year, I obtained Red Dog's total toxic releases each year. Because these numbers were quite large (often exceeding 500,000,000 pounds in each year), I also calculated the proportion of the total toxic releases in the US for which Red Dog were responsible each year. These results are shown in Table 1 below.

Table 1: Table 1

| Year | Red Dog Total [lbs] | US Total [lbs] | Proportion of Total |
|---|---|---|---|
| 2003 | 487365521 | 3985853610 | 12.2 |
| 2004 | 511271279 | 3733208310 | 13.7 |
| 2005 | 548412055 | 3842195214 | 14.3 |
| 2006 | 668510036 | 3801739821 | 17.6 |
| 2007 | 591136748 | 3580395926 | 16.5 |
| 2008 | 573954190 | 3401547743 | 16.9 |
| 2009 | 698891654 | 3038737510 | 23.0 |
| 2010 | 835635275 | 3401823411 | 24.6 |
| 2011 | 1048248269 | 3701652497 | 28.3 |
| 2012 | 876175659 | 3232492313 | 27.1 |
| 2013 | 970307667 | 3721523311 | 26.1 |

| Year | Red Dog Total [lbs] | US Total [lbs] | Proportion of Total |
|---|---|---|---|
| 2014 | 1183146990 | 3537256782 | 33.4 |
| 2015 | 622251692 | 2967684153 | 21.0 |
| 2016 | 853952446 | 3128756592 | 27.3 |
| 2017 | 1173882527 | 3536128617 | 33.2 |
| 2018 | 971511776 | 3369249603 | 28.8 |

It is immediately clear that Red Dog contributed a substantial proportion of all toxic releases in the US every year in the dataset. At most, in 2014, they were responsible for an astonishing 33.4% of all toxic releases in the US. These results are summarized graphically in fig 5.

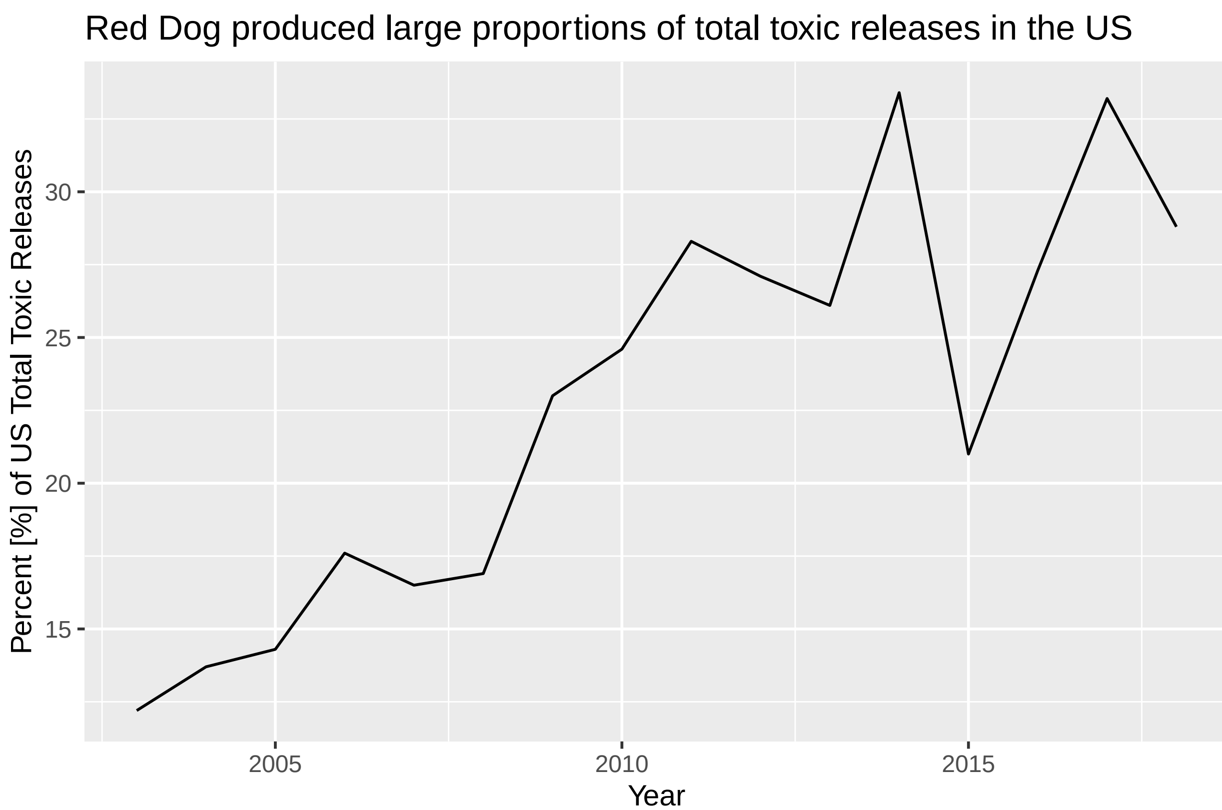


Fig 5

These results illustrate that toxic releases from Red Dog in Alaska are contributing disproportionately to the overall trend of toxic releases in not only the West but also the US at large. Indeed, if I remove Red Dog from the dataset altogether, then the overall trend in annual average toxic releases in the US becomes more obvious. This trend is shown in fig 6, where the average amount of toxic releases in US industry is now declining over time.

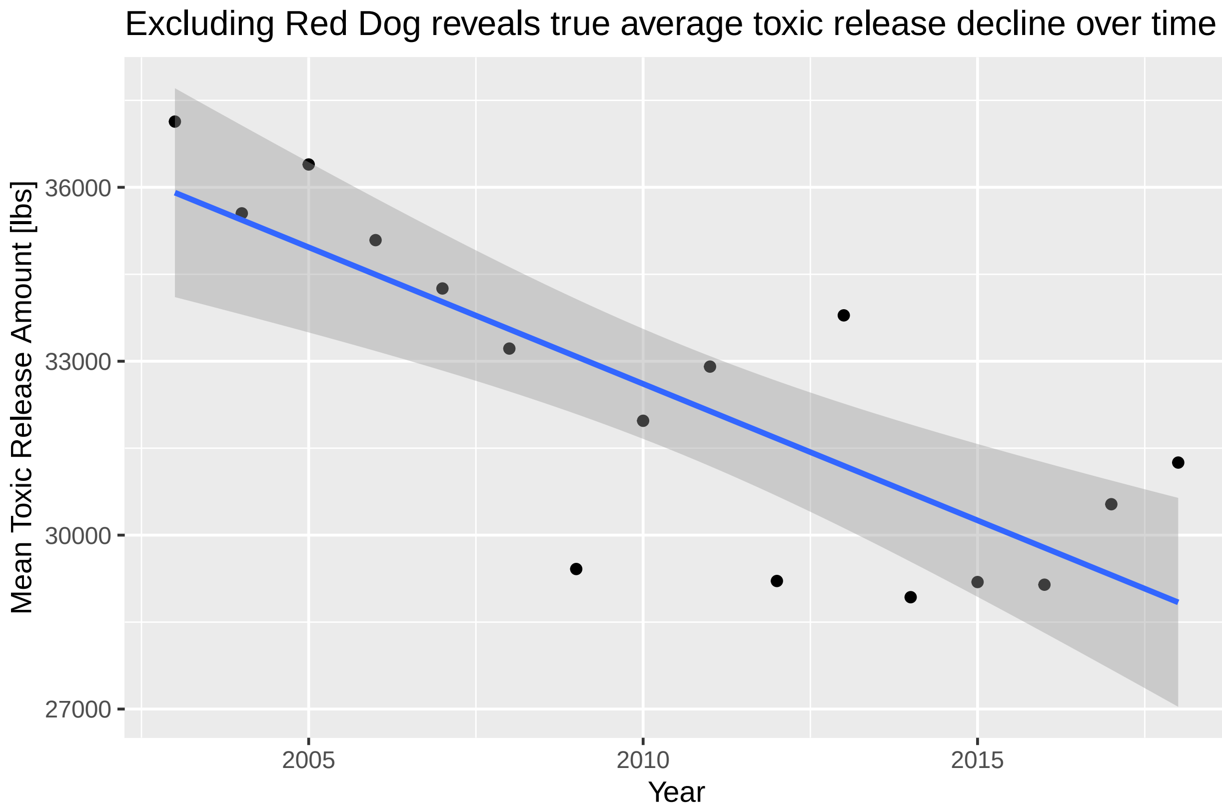


Fig 6

This model is a significantly better fit to the data ($R^2 = 0.635$), and the 95% CI is much tighter, meaning that the uncertainty about predictions is much lower.

Until now, I have studied changes in US toxic releases using annual mean amounts, but as we have seen, the mean is not robust to outlier companies that produce extremely high amounts of toxic releases. Another useful statistic for this analysis is the median amount of annual toxic releases because the median is less sensitive to outlier companies. Fig 7 shows the median amount of toxic releases in the US over the same 16-year time frame as before.

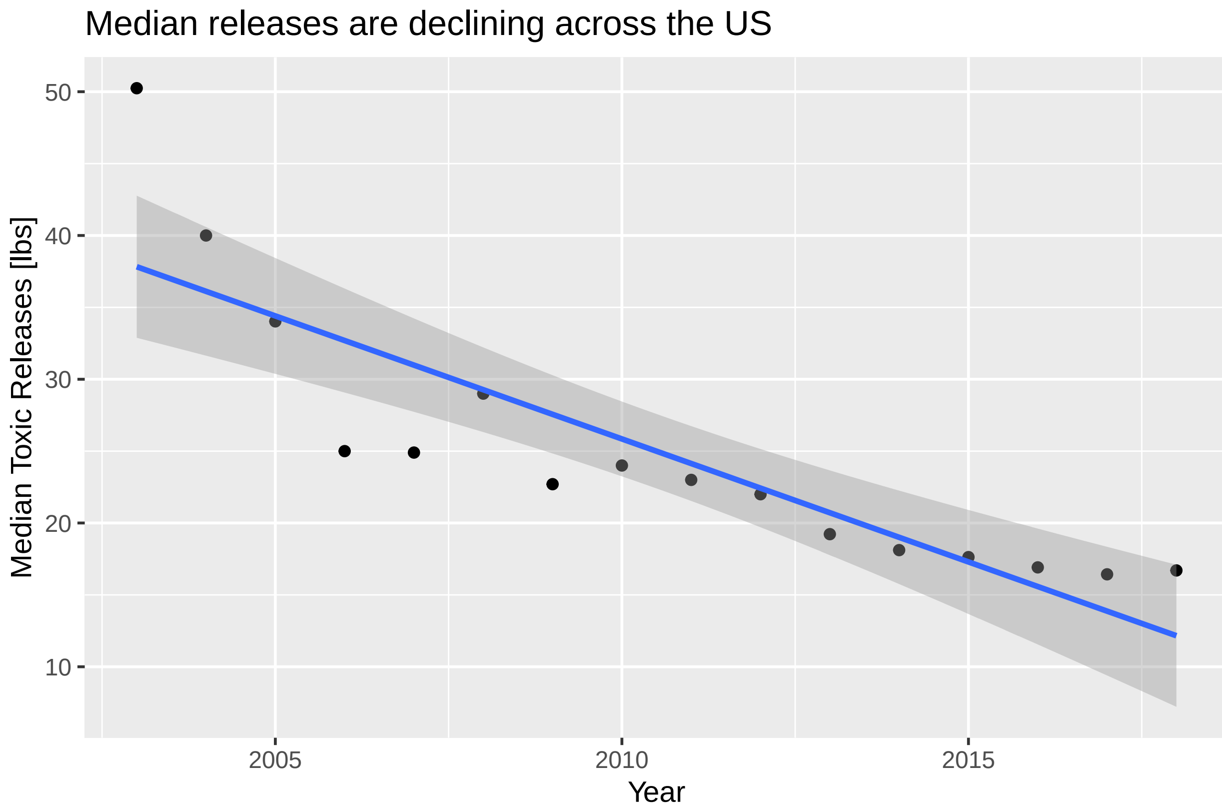


Fig 7

This model is fits the data fairly well ($R^2 = 0.753$), and the 95% CI shows that there is less uncertainty in the median model than in the mean model shown in fig 1. Another interesting piece of information is that the median toxic release amount annually is several orders of magnitude lower than the average amount each year, suggesting that at least 50% of companies produce relatively low amounts of toxic releases (under 100 pounds) per year. These patterns are reflected in the median toxic release amount each year by region, as shown in fig 8.

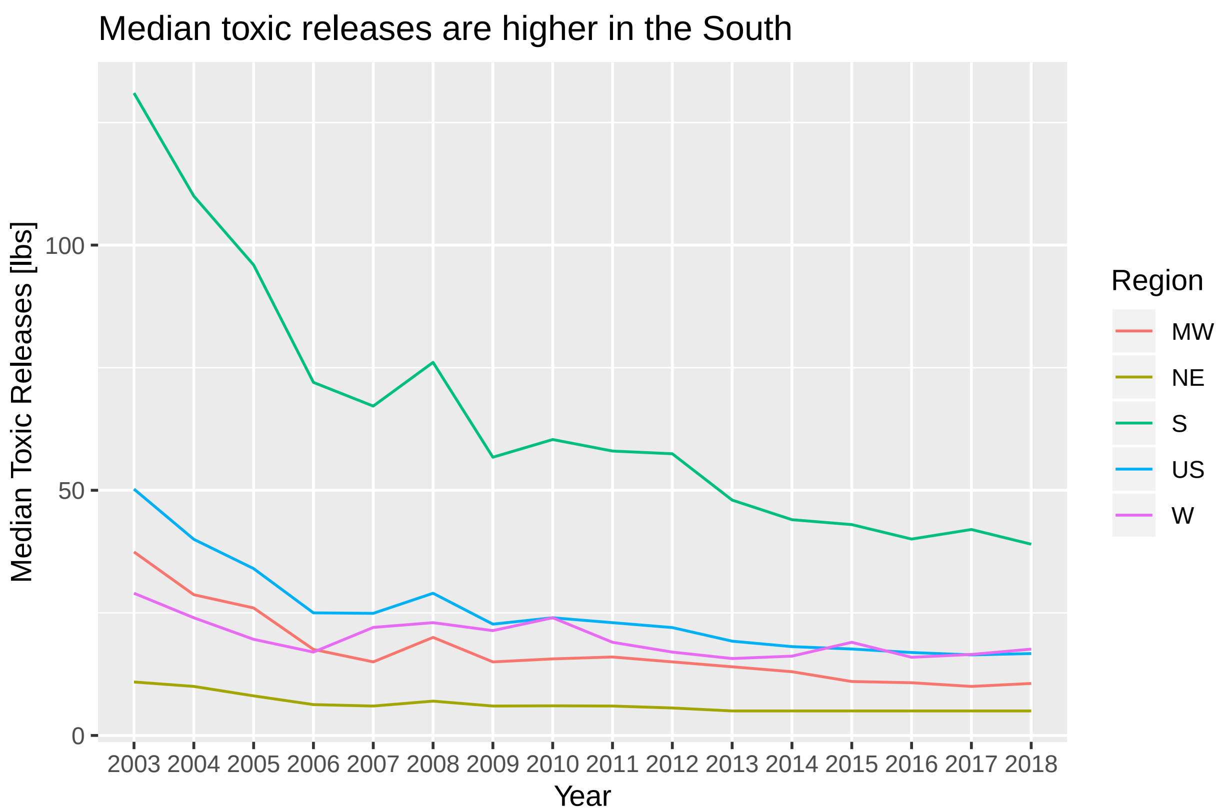


Fig 8

This time, we see that the median company in the South produces higher toxic releases than the median company in the US overall, contrasting with the results based on the average toxic release amount in fig 2, where companies in the West produced well above the average amount. As an aside, the median data indicate that companies in the South have a relatively larger impact on the environment through annual toxic releases, and these data would need to be considered in policy decision-making along with the data shown in fig 2 for completeness. Nevertheless, the median amount of toxic releases is declining in each region over time, as shown in fig 9. As before, note that the range of median toxic releases is different for each region.

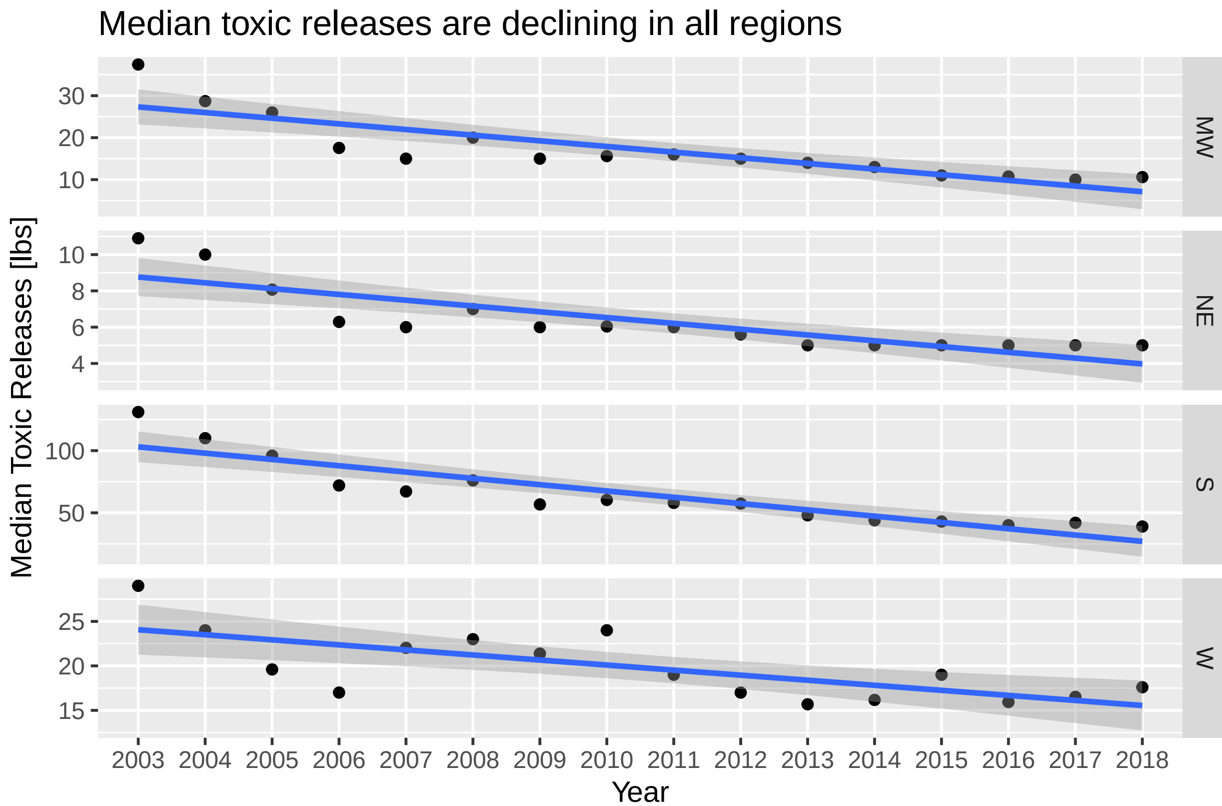


Fig 9

## Conclusions

With the removal of Red Dog from the mean data, both the mean and median amounts of toxic releases in the US are decreasing annually, showing that toxic releases in US industry have been lowered over time. These data provide evidence for the success of the EPA TRI program in reducing the amount of toxic releases by US companies into the environment. However, it is not possible to conclude that the aforementioned program is solely responsible for these trends. This study only demonstrates the association of the TRI program with lower toxic releases over time, but it does not attempt to isolate the exact causes of this decline. There may be other confounding factors that contribute to the lowering of toxic releases. For example, as technology advances, novel methods of toxic waste processing are developed and implemented in industry, and these methods also have an impact on the amount of toxic releases produced by companies. One such advancement is the recycling of plastics using sunlight.

The simple models used in this study are sufficient to understand general trends in toxic releases, but they are not necessarily predictive of future outcomes. Time is the main variable considered here, but while this study demonstrates high correlation between time and lower toxic releases, the passage of time alone is not an impetus for change in industrial trends. Still, there is no reason to suspect that these trends will suddenly reverse because EPA policy and technological advancement have been mostly consistent over time. While the exact amounts of toxic releases are unpredictable, they will probably continue to decline over time.