

## Supporting Information

### Trace element mass flow rates from U.S. coal fired power plants

Xiaodi Sun, Daniel B. Gingerich, Inês L. Azevedo, Meagan S. Mauter

41 pages

19 figures

5 tables

### Table of Contents

SI Section 1: Data and model repository .....	S2
SI Section 2: Coal-fired generation included and excluded from analysis .....	S3
SI Section 3: Variability of coal within each county .....	S5
SI Section 4: Map of eGRID sub-regions .....	S8
SI Section 5: Concentration of trace elements in coal blends .....	S10
SI Section 6: Effect of COALQUAL lower detection limit assumptions on the concentration of trace elements in coal blends .....	S14
SI Section 7: Coal purchases from preparation plants .....	S16
SI Section 8: Studies on partitioning of trace elements by air pollution control .....	S18
SI Section 9: Partitioning of trace elements to the solid, liquid, and gas phase at each coal boiler.....	S24
SI Section 10: Comparing partitioning fractions estimated matching individual air pollution control devices against matching combinations of air pollution control devices.....	S25
SI Section 11: Generation treated by different air pollution controls and control combinations.....	S27
SI Section 12: Comparing partitioning to gas from MATS against literature data.....	S28
SI Section 13: Generation normalized mass flow rates of trace elements at the boiler level in the U.S. and in different eGRID sub-regions.....	S31
SI Section 14: Temporal variability of trace element concentrations in coal .....	S32
SI Section 15: Comparison of concentrations of trace elements in coal using COALQUAL and EIA against MATS ICR reporting.....	S33
SI Section 16: Comparison of minimum model estimates of gas phase Hg mass flows against CEMS ..	S36
SI Section 17: Benchmarking estimates of gas-phase, generation-normalized mass flow rates against separate model results .....	S37
Supplementary Information References: .....	S40

## **SI Section 1: Data and model repository**

We uploaded all MATLAB code and data spreadsheets used in this analysis in the following GitHub repository: [https://github.com/xdansun/trace\\_elem\\_cfpp\\_model](https://github.com/xdansun/trace_elem_cfpp_model)

## SI Section 2: Coal-fired generation included and excluded from analysis

In the analysis presented in the main manuscript, we estimate Hg, Se, As, and Cl mass flow rates for 276 plants that generated 1190 TWh of electricity, which is 63% of all coal-fired power plants (CFPPs) and 89% of all coal electricity generation in 2015. For context there are 436 coal plants in the U.S. which generated 1330 TWh in 2015.<sup>1</sup>

We exclude some coal boilers in our analysis due to limitations from data on coal boilers collected using EIA-923, Schedule 3, EIA-860, Schedule 6, and COALQUAL. These data limitations include:

*Power plants with less than 1 MW capacity:* Power plants with a total nameplate capacity less than 1 MW are not required to report generation or environmental control data using EIA-923, Schedule 3 and EIA-860, Schedule 6. Therefore, we do not know what percentage of CFPPs are affected nor how much electricity these plants generated.

*Boilers connected to multiple generators and vice-versa:* Where boilers are connected to multiple generators and vice-versa, we are unable to attribute generation to a boiler or normalize mass flow rates based on electricity generation. About 24% of CFPPs and 4.6% of coal generation are removed from our analysis due to this limitation.

*Boilers without coal purchases:* Without the location of coal purchases and quantity of coal purchased, we cannot estimate the concentrations of trace elements in the coal blend. This data limitation removes an additional 6.7% of CFPPs and 3.1% of coal generation.

*Boilers that burn foreign and/or waste coal:* Our dataset for run-of-mill coal samples, COALQUAL, only includes domestic and non-waste coal samples, so we cannot estimate trace element concentrations in foreign or waste coal.<sup>2</sup> This data limitation removes an additional 5.3% of CFPPs and 2.8% of coal generation.

*Boilers that do not report fuel consumption:* Without knowing how much fuel is combusted in the boiler, we cannot estimate the mass flow rates of trace elements from the coal blend. This data limitation removes an additional 0.3% of CFPPs and 0.3% of coal generation.

The amount of generation excluded from our analysis due to the above limitations are listed in Table S1. Overall, at least 10.8% of coal generation is lost in the analysis, with 4.6% of total generation lost from boiler to generator linking issues and 6.2% of total generation lost from coal consumption and purchasing issues.

**Table S1:** Percent of total coal generation that is not included in our analysis due to data limitations in input data of CFPPs. For context there are 436 coal plants in the U.S. which generated 1330 TWh in 2015.<sup>1</sup>

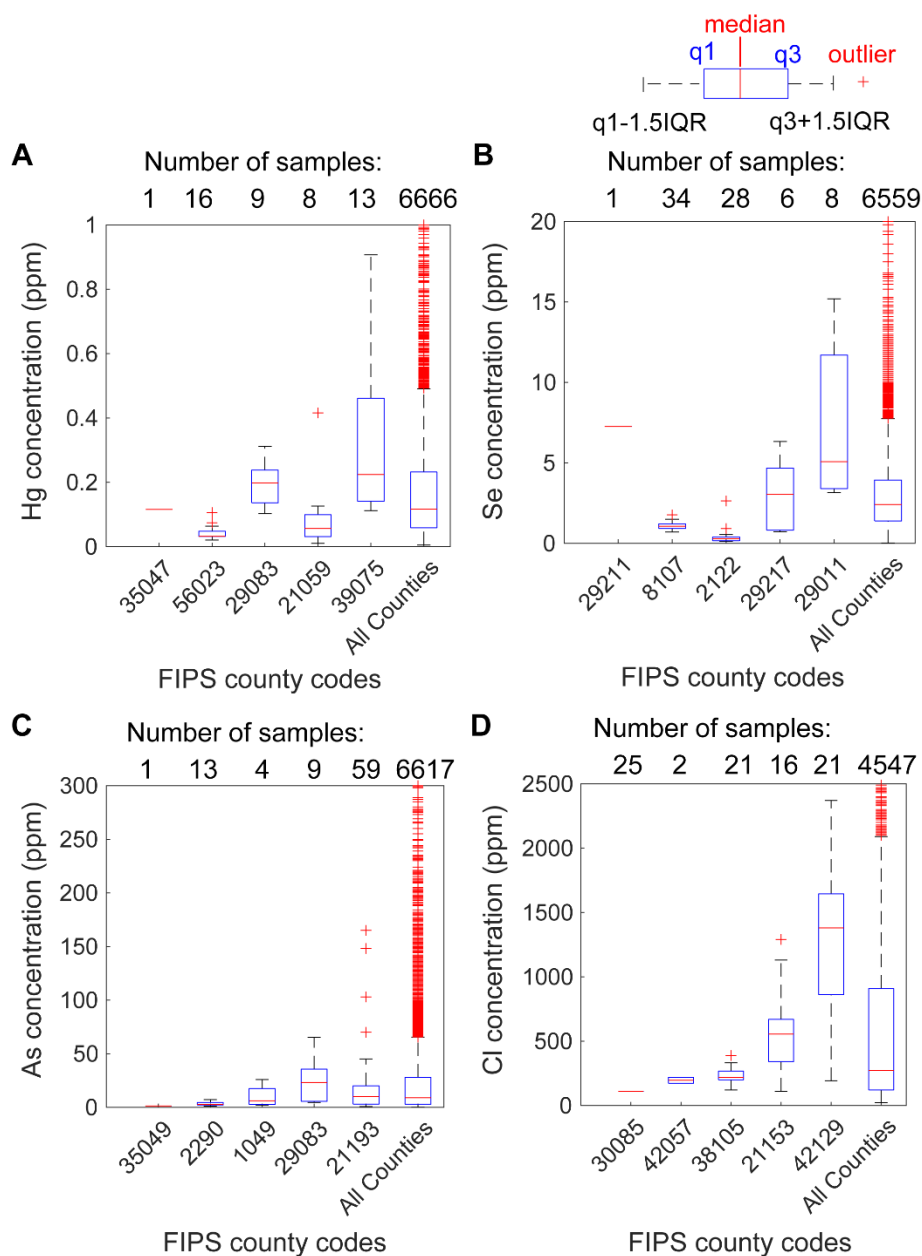
Data limitation	Plants excluded in analysis		Generation excluded in analysis	
	(plants)	(%)	(TWh)	(%)
Power plants with less than 1 MW capacity	Unknown	Unknown	Unknown	Unknown
Boilers connected to multiple generators and vice-versa	105	24.1	61	4.6
Power plants which did not report coal purchases	29	6.7	41	3.1
Boilers that burn foreign and/or waste coal	23	5.3	37	2.8
Boilers that do not report fuel consumption	1	0.2	4	0.3
<b>Total</b>	<b>158</b>	<b>36.3</b>	<b>144</b>	<b>10.8</b>

Table produced by the authors using data from COALQUAL, EIA-923, Schedule 3 and EIA-860, Schedule 6.<sup>1-3</sup>

### **SI Section 3: Variability of coal within each county**

COALQUAL reports trace element concentrations in coal samples across different counties in the U.S. For each county, the concentration of trace elements and the number of samples can vary. In Figure S1 we show the variability in the concentration of trace elements across five illustrative U.S. counties and for all samples reported in COALQUAL. We also report the number of samples reported in each county. We selected the five counties for each trace element by first, ordering the counties by range of trace element concentration in ascending order and second, choosing one county from each quintile of the ordered counties.

At the county-level, we find significant variability in the concentration of trace elements even if there are a few samples from the county. For example, in Figure S2B, county 29011 has 8 Se samples reported but the samples vary by a factor of 5 (lowest is 3.2 ppm and highest is 15.2 ppm). At the national level, there are many outliers in the COALQUAL database. About 2-8% of the reported data for each trace element are outliers, all of which are high values. We also find that the range of measurements from the 75<sup>th</sup> percentile to the maximum is much greater than the range from the minimum to the 25<sup>th</sup> percentile, which suggests that our data may have more outliers with high trace element concentrations that can lead to trace element concentrations in coal and subsequently, higher estimates of trace elements in the waste streams.



**Figure S1:** Boxplots of trace element concentrations in coal samples in COALQUAL at a few illustrative U.S. counties and all coal samples reported in COALQUAL for the trace elements A) Hg, B) Se, C) As, and D) Cl. The y-axis is cutoff for readability. The highest concentration reported is 3.3 ppm for Hg, 153 ppm for Se, 2200 ppm for As, and 8920 ppm for Cl. In the boxplot, the red line is the median, the blue lines are the first and third quartiles, the lower whisker is the first quartile minus 1.5 times the interquartile range, the upper whisker is the third quartile plus 1.5 times the interquartile range, and the red crosses are outliers. The figure is produced using COALQUAL data.<sup>2</sup>

Additionally, we calculate the correlation between the concentration of trace elements using the data in USGS COALQUAL. The correlation coefficients and significance levels of these correlations are reported in Table S2. We find that the As, Cl, and Hg concentrations in coal are all correlated with each other with a significance of at least 0.10 level. The only trace element concentration with a statistically significant correlation with Se is Hg, as As and Cl have p-scores above the 0.10 level (p-scores of 0.2913 and 0.1363, respectively).

**Table S2:** Coefficients of correlation between the concentrations of As, Cl, Hg, and Se in the USGS COALQUAL database.”

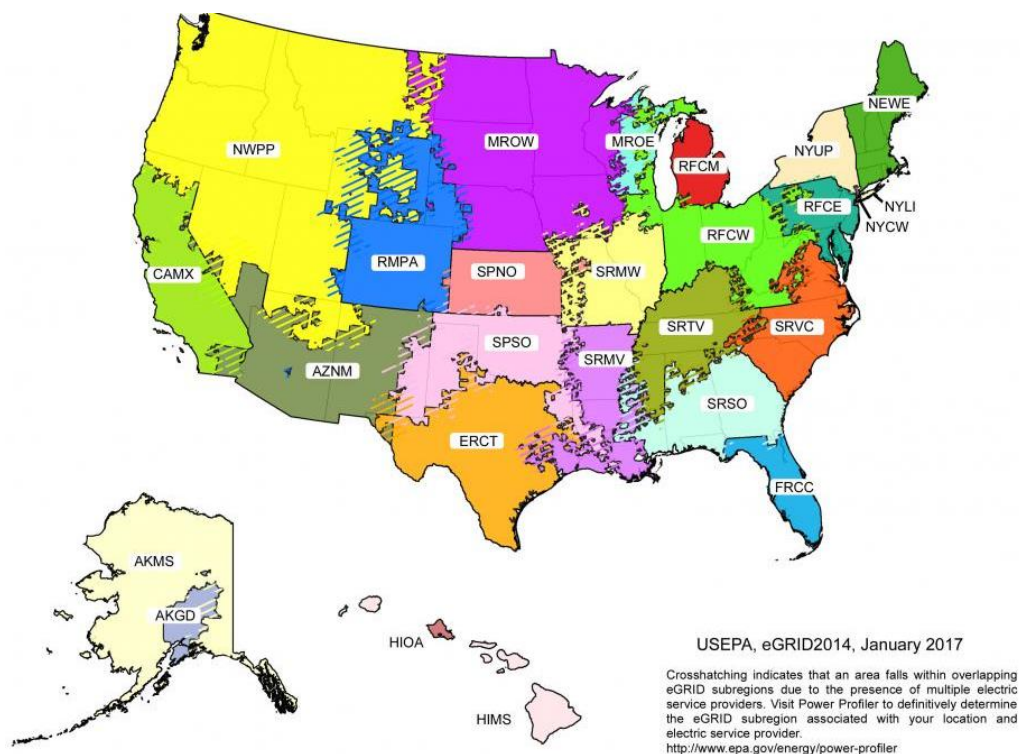
	As	Cl	Hg	Se
As	1			
Cl	-0.004 <sup>b</sup>	1		
Hg	0.294 <sup>b</sup>	0.077 <sup>a</sup>	1	
Se	0.067	0.068	0.170 <sup>b</sup>	1

<sup>a</sup> Significant at the 0.10 level.

<sup>b</sup> Significant at the 0.01 level.

## SI Section 4: Map of eGRID sub-regions

Figure S2 shows the 26 Emissions and Generation Resource Integrated Database (eGRID) sub-regions in the U.S. The U.S. Environmental Protection Agency (EPA) designed the eGRID sub-regions to mimic regional transmission organizations and independent system operators within North American Electric Reliability Corporation regions. In our study, we focus on the contiguous U.S.



**Figure S2:** Map of eGRID sub-regions (source: <https://www.epa.gov/energy/egrid-sub-region-representational-map>)

Table S3 shows the number of plants with at least one boiler with coal as its primary fuel, coal generation, total coal capacity, and average capacity factor for all coal electrical generating units in each eGRID sub-region within the contiguous U.S. included in our modeling dataset (SI Section 3 for more information on plant inclusion). The NYCW and NYLI sub-regions have no coal plants. The one coal plant in sub-region CAMX and the five coal plants in sub-region NEWWE are removed from our analyses for the reasons listed in SI Section 2. One plant, Roy S Nelson, does not have an eGRID sub-region reported and is therefore not included in any eGRID-level analysis in either the main manuscript or the SI.

In Table S3 we also show the percentage of plants and percent of coal generation modeled for each eGRID sub-region. We find that we model 45-89% of coal plants and 68-96% of coal powered electricity generation in each eGRID sub-region. See SI Section 3 for more details about coal plants included in the analysis.



**Table S3:** The number of coal plants, coal generation, coal capacity, and coal capacity factor in 2015 for all coal plants included in the model dataset in each eGRID sub-region.

eGRID sub-region <sup>a</sup>	Coal plants <sup>b</sup>	Generation (TWh)	Capacity (GW)	Average capacity factor	Coal plants modeled (%)	Coal generation modeled (%)
AZNM	7	24	4.6	0.59	71	96
ERCT	17	98	20.8	0.54	88	92
FRCC	11	41	10.2	0.46	45	56
MROE	12	15	3.5	0.49	50	94
MROW	63	111	21.6	0.59	54	96
NWPP	21	111	20.2	0.63	71	95
NYUP	7	2	1.9	0.14	57	85
RFCE	19	46	14.7	0.36	68	68
RFCM	24	52	11.5	0.52	50	78
RFCW	79	280	64.5	0.50	66	92
RMPA	17	46	7.9	0.67	76	99
SPNO	14	41	8.6	0.55	86	100
SPSO	18	67	14.6	0.53	89	92
SRMV	5	24	6.6	0.42	80	87
SRMW	29	107	22.2	0.55	62	88
SRSO	19	79	21.5	0.42	63	88
SRTV	28	103	26.3	0.45	71	91
SRVC	32	77	20.6	0.43	59	87
Roy S Nelson <sup>c</sup>	1	2.1	0.6	0.39	0	0

<sup>a</sup> We exclude sub-regions AKGD, AKMS, HIMS, and HIOA because they are not in the contiguous U.S. and sub-regions CAMX, NEWE, NYCW, and NYLI because there are no coal plants remaining in those sub-regions after applying the criteria mentioned in SI Section 3.

<sup>b</sup> The number of plants with at least one boiler using bituminous, subbituminous, or lignite coals as its primary fuel.

<sup>c</sup> Roy S. Nelson power plant is excluded in all analysis in the manuscript and SI.

Note: Table produced using data from EIA-923, Schedule 3 and eGRID2014.<sup>1,4</sup>

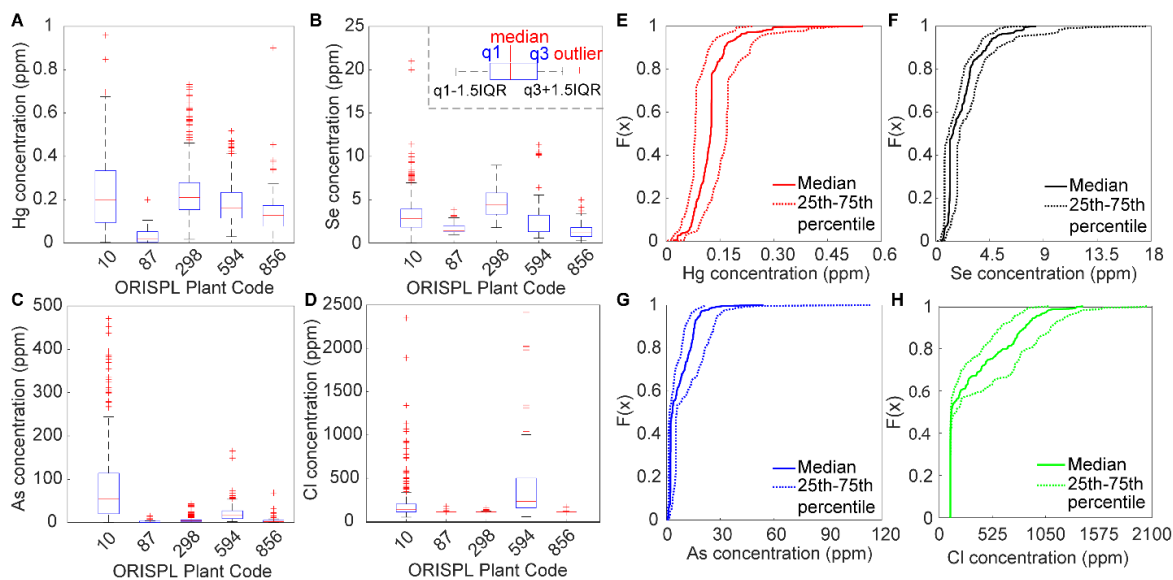
## SI Section 5: Concentration of trace elements in coal blends

We estimate a distribution for the concentration Hg, Se, As, and Cl entering each U.S. CFPP in 2015 as described in the Methods section of the main manuscript. In Figure S3A-D, we show an illustration of the results for a few CFPPs, and in Figure S3E-H we provide the distribution of median concentrations across our entire dataset. In Figure S4, we also show the distribution of median concentrations at each eGRID sub-region.

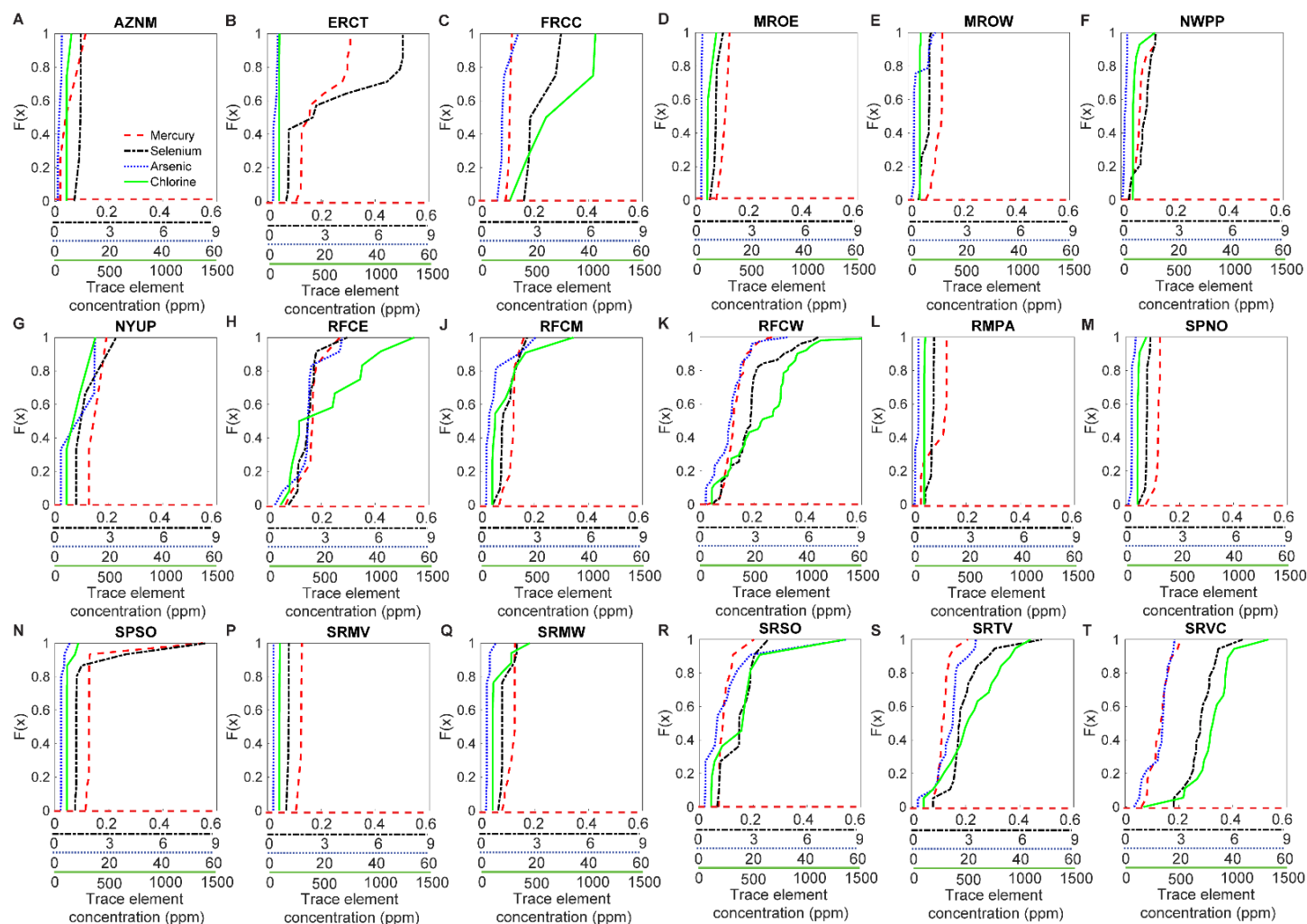
In Figure S3A-D, the five plants we chose are the Greene County plant in Alabama (ORISPL: 10), the Escalante plant in New Mexico (ORISPL: 87), the Limestone plant in Texas (ORISPL: 298), the Indian River Generating Station in Delaware (ORISPL: 594), and the E.D. Edwards plant in Illinois (ORISPL: 856). These 5 plants generated 2.2 TWh, 1.3 TWh, 9.2 TWh, 0.6 TWh, and 2.9 TWh in 2015, respectively.

As explained in the methods section, for each plant, we bootstrap 10,000 draws, and in Figure S3H the median distribution shows the cumulative distribution function of the median for each plant (and the same strategy is pursued for the 25<sup>th</sup> and 75<sup>th</sup> percentiles). We find that the median of the median concentrations of trace elements across the plants in our dataset is 0.13 ppm for Hg, 1.4 ppm for Se, 3.4 ppm for As, and 122 for Cl. These concentrations are within the same order of magnitude as the median of the median concentration of trace elements from the Mercury and Air Toxics Standards Information Collection Request (MATS ICR), which are 0.08 ppm for Hg, 1.0 ppm for Se, 2.0 ppm for As, and 200 for Cl (See SI Section 12 for more details).

There is significant variability in the concentrations of trace elements in combusted coal both within an individual plant and across the coal-fired fleet. For some CFPPs, variability associated with concentrations of trace element in coal is substantial (for illustration, plant 10 in Figure 3A has a Hg concentration in the coal blend ranging from 0.005 ppm to 1 ppm, representing a 200x difference between the lowest and highest estimated values), due to difference in the trace element in coal reported in the samples in COALQUAL and the variability in the monthly coal purchased from different counties. Across the fleet, the variability in trace element content is mostly determined by differences in the coal across U.S. regions, such as Wyoming Powder River Basin subbituminous coal versus eastern bituminous coal.<sup>5-8</sup> In SI Section 13, we also show these results for plants located in each eGRID sub-region, which highlights the regional diversity in the trace element content of coal combusted at CFPPs.



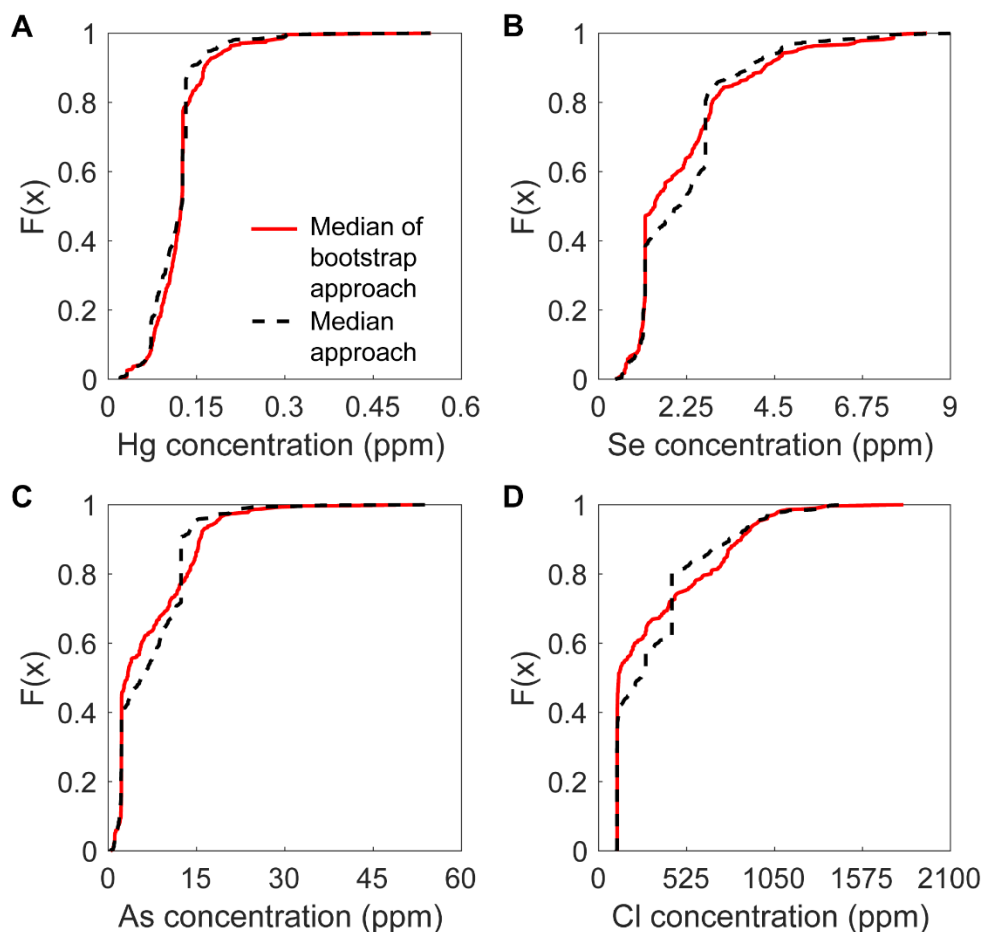
**Figure S3: (A-D)** Estimates of trace element concentration in coal entering a few illustrative U.S. coal-fired power plants. In the boxplot, the red line is the median, the blue lines are the first and third quartiles, the lower whisker is the first quartile minus 1.5 times the interquartile range, the upper whisker is the third quartile plus 1.5 times the interquartile range, and the red crosses are outliers. **(E-H)** Cumulative distribution functions of 25<sup>th</sup> percentile, median, and 75<sup>th</sup> percentile of the bootstrapped concentrations of trace elements in the coal blend across coal plants in the coal fleet. For the two panels, the following trace elements are represented: A/E) Hg, B/F) Se, C/G) As, and D/H) Cl



**Figure S4:** Cumulative distribution functions of the bootstrapped median Hg, Se, As, and Cl concentrations in the coal blend across coal plants in the eGRID sub-regions A) AZNM, B) ERCT, C) FRCC, D) MROE, E) MROW, F) NWPP, G) NYUP, H) RFCE, J) RFCM, K) RFCW, L) RMPA, M) SPNO, N) SPSO, P) SRMV, Q) SRMW, R) SRSO, S) SRTV, V) SRVC. We excluded sub-regions AKGD, AKMS, HIMS, and HIOA because they are not in the contiguous US and sub-regions CAMX, NEWE, NYCW, and NYLI because there are no coal plants remaining in those sub-regions after applying the criteria stated in SI Section 2. Figure produced by the authors using data from COALQUAL and EIA-923 Schedule 5.<sup>1,2</sup>

Lastly, we compare estimates of trace element concentrations by bootstrapping coal samples by location and by taking the median of all coal samples at the location. We refer to the former as the “bootstrap approach” and the latter as the “median approach.” The bootstrap approach is the approach utilized in the manuscript.

In Figure S5 we show the median concentration of trace elements using the bootstrap approach and the concentration of trace elements using the median approach. Although one may expect the median concentration of the bootstrap approach to nearly equal the median approach, they can produce different results at the plant-level because the median is not a linear operator. A plant that purchases coal from multiple counties will have a higher chance of bootstrapping a sample away from the median, which skews toward values much greater than the median. Hence, each trace element has plants where the two approaches estimate different concentrations. The median approach does not handle the variability of trace element concentrations at any level and can only generate point estimates. Thus, we use the bootstrap approach in our work.



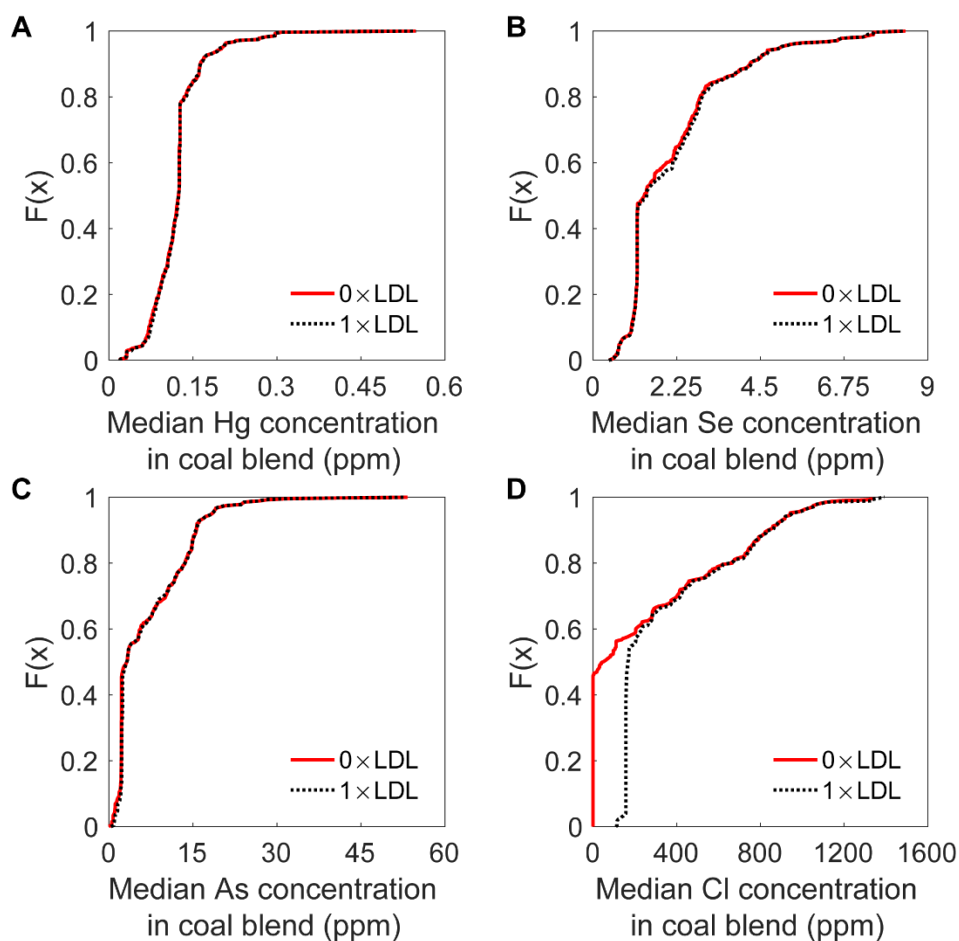
**Figure S5:** Cumulative distribution functions of concentrations of trace elements in the coal blend at the plant-level using bootstrap approach and median approach for trace elements A) Hg, B) Se, C) As, and D) Cl. For the bootstrap approach, the median of the bootstrapped trace element concentration is shown.

## **SI Section 6: Effect of COALQUAL lower detection limit assumptions on the concentration of trace elements in coal blends**

In COALQUAL, there are 6666 Hg samples, 6560 Se samples, 6618 As samples, and 4548 Cl samples. Of these samples, about 4% of Hg samples, 3% of Se samples, 1% of As samples, and 21% of Cl samples are below the detection limit. COALQUAL assumes that samples below the detection limit have concentrations equal to 0.7 times the detection limit.

In Figure S6, we plot a cumulative distribution function showing the median concentration of trace elements in the coal blend at the plant-level assuming samples below the detection limit have a concentration equal to either 0 or equal to the lower detection limit. We find that for Hg, Se, and As that the assumption of the lower detection limit has little effect on the concentration of trace elements at the plant-level. Indeed, the median of the difference in median concentrations for Hg, Se, and As is 0.04 ppm, 0.004 ppm, and 0.04 ppm respectively. These results are not surprising given the small percentage of samples in COALQUAL that are below the detection limit.

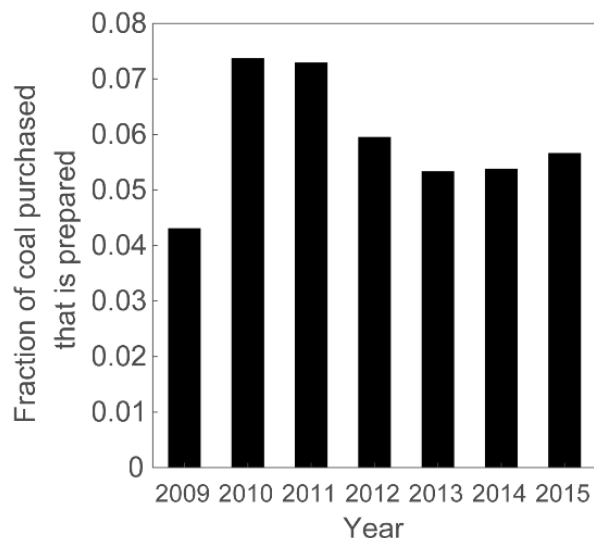
However, we find that Cl concentrations in the coal blend of half of the coal fleet can change by up to 130 ppm depending on whether we assume samples below the detection limit equal zero or equal the detection limit. Even though only 21% of Cl samples are below the detection limit, these results are still surprising because they affect half of the coal fleet. This result suggests that nearly half of the fleet purchase coal from counties where COALQUAL only has samples that are below the detection limit. Therefore, the detection limit issue highlights one additional problem with the COALQUAL data, especially when estimating Cl concentrations in coal blends.



**Figure S6:** Cumulative distribution functions of trace elements in the coal blend at the plant-level assuming all samples with concentrations lower than the detection limit have concentration equal to zero or equal to the detection limit for A) Hg, B) Se, C) As, and D) Cl. COALQUAL assumes all samples below the lower detection limit have concentration equal to 0.7 times the detection limit.

## SI Section 7: Coal purchases from preparation plants

Coal plants report coal purchases from mines and preparation plants in EIA-923, Schedule 5. In 2015, coal plants purchased nearly 780 million tons of coal, of which only 42 million tons were purchased from coal preparation plants. Thus, less than 6% of total coal purchases come from coal preparation plants.<sup>1</sup> This has been consistent since 2009, with about 4-8% of coal purchases from coal preparation plants, as shown in Figure S7. Additionally, less than 1% of trace metals are removed by on-site coal pulverization prior to combustion.<sup>9,10</sup> Because cleaned coal purchases account for a small percentage of total coal purchases, and on-site processing does not remove a significant fraction of trace elements, we assume coal cleaning does not affect our trace element input estimates.



**Figure S7:** Fraction of prepared coal purchased to total coal purchased made by CFPPs from 2009 to 2015. Figure produced using data from EIA-923, Schedule 5.<sup>1</sup>

Modeling coal cleaning is not simple, due to a lack of information about the processes that take place at different coal preparation plants. We perform a conservative analysis where we assume coal cleaning removes the highest concentration of trace elements observed in the literature for all coal samples in a county with a preparation plant. The highest removals we observed for Hg, Se, As, and Cl from coal cleaning are 78.3%, 80.3%, 84.6%, and 68%.<sup>11,12</sup> We rerun our model after adjusting for coal concentrations using the aforementioned coal cleaning removals and recreate Figure 4 from the manuscript.

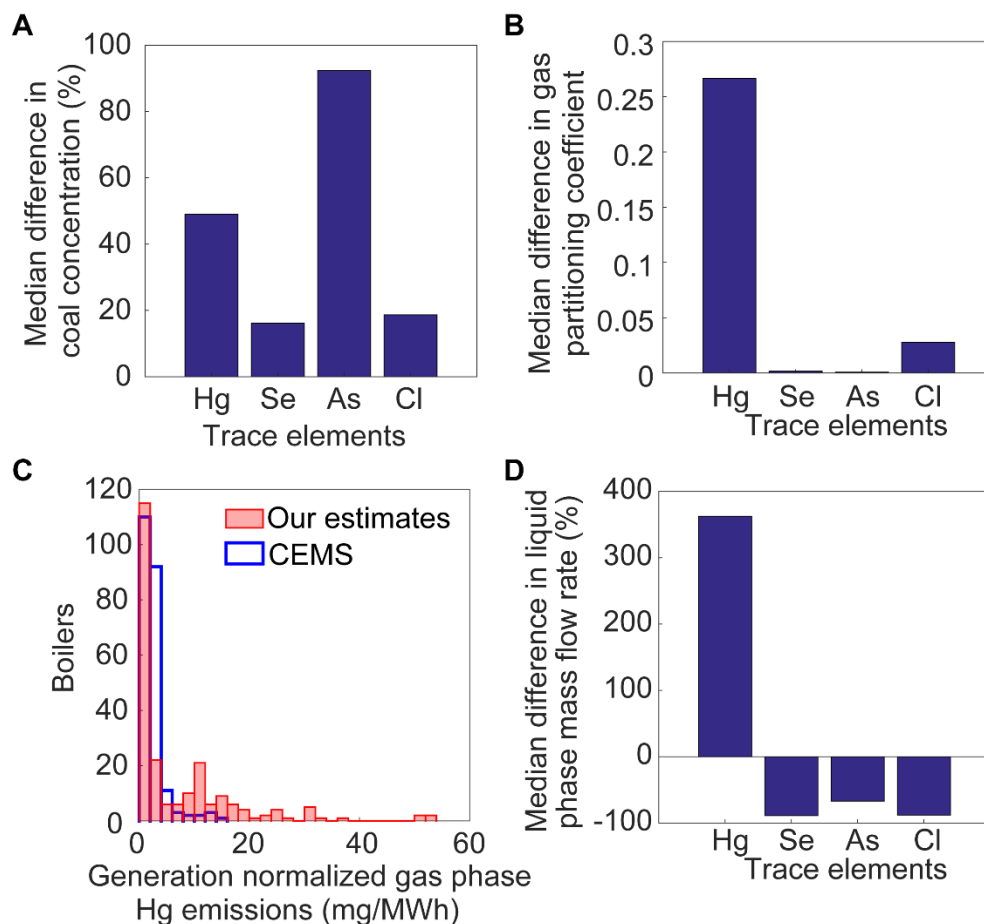
As shown in Figure S8, model accuracy changes only modestly. The notable changes are in the Hg CEMS result, where the median estimate is 1.8 mg/MWh, which is an improvement from the previous result of 2.8 mg/MWh. However, we still observe 81 high-rank coal boilers, which is close to a third of boilers in the Hg CEMS validation analysis, that exceed 5.9 mg/MWh, which is the regulated Hg emissions limit for high-rank coal units. Therefore, even with this conservative coal cleaning assumption, we still overestimate Hg emissions to the gas phase.

The other notable improvement is with the validation against the wet flue gas desulfurizer (FGD) mass flow rate observed by the EPA during the Effluent Limitation Guidelines for the Steam Electric Generating Sector (ELGs) rulemaking. We find that the Hg validation improves from a 500% difference



between our estimate and EPA's to 360% under the coal cleaning assumption. This improvement gets the model closer to the EPA estimated values, but still represents an overestimate of Hg mass flow rate.

Ultimately, while we find some improvements with a conservative coal cleaning assumption, we find that our results are the same. Data quality improvements and more data of lesser documented processes, such as bromine addition, are more important to bridging the gap between model estimates of trace element mass flow rates to the ground truth. Therefore, we believe coal cleaning is a second-order effect and that we can safely not account for it in our main analysis.



**Figure S8:** Summary of the comparison analyses under the assumption that all coal purchased from counties with coal preparation plants are cleaned to the maximum extent. A) Median percent difference in plant-level trace element composition of coal between model estimates and data reported in the MATS ICR; B) Median difference in boiler-level trace element partitioning to the gas phase between model estimates and data reported in MATS ICR; C) Histogram of generation normalized boiler level gas phase Hg emissions estimated from model and reported in the CEMS in 2015; D) Median percent difference in liquid phase mass flow rates estimated at the fleet level compared to average mass flow rates for the 88 plants sampled and reported in the Environmental Assessment of the Effluent Limitation Guidelines.

## SI Section 8: Studies on partitioning of trace elements by air pollution control

In Table S4, we list the removal fraction of trace elements by the boiler and different air pollution control devices (APCDs) for each study included in the model. For those studies that include several boilers or plants, we calculate the partitioning factors for each plant and use the average removal to represent that study. We only include trace element partitioning estimates from U.S. CFPPs. We assume all waste exiting PM controls and the gypsum stream of wet FGDs is in the solid phase, all waste exiting the Cl purge stream of wet FGDs is in the liquid phase, and all remaining wastes exit the plant in the gaseous phase. We further assume wet flue gas desulfurization is always the last APCD downstream of the boiler.

A significant fraction of trace metals are removed by particulate matter (PM) controls because of the presence of particulate bound trace metals in the flue gas and high fly ash removal rates of PM controls.<sup>13</sup> PM controls are more efficient at removing Se and As from the flue gas than Hg because Se and As are less volatile than Hg.<sup>13</sup> However, the addition of activated carbon injection (ACI) combined with a PM control can remove up to 80% of Hg in the flue gas, which is comparable to the removal of Se and As by PM controls.

A fraction of trace metals is also removed by SO<sub>2</sub> controls. There are two types of postcombustion SO<sub>2</sub> controls: wet FGDs and dry/semi-dry FGDs. Wet FGDs capture a significant fraction of the fine particles that are not captured by upstream PM controls, which leads to conversion of trace metals from gas phase to solid phase.<sup>14</sup> Specifically for Hg, the selective catalytic reactor (SCR) helps promote the oxidation of Hg<sup>0</sup> into Hg<sup>+2</sup> so that wet FGDs can direct more Hg into liquid waste streams, but it is not clear that SCRs have the same effect on other trace metals.<sup>15</sup> Dry and semi-dry FGDs serve a similar function as the wet FGD but produce zero liquid phase waste. Waste products from dry FGDs are entrained in the flue gas and removed by a downstream PM control, typically a fabric filter (FF) though some dry FGDs will have an additional waste for a fraction of the byproducts.<sup>16</sup>

Cl exists predominantly as HCl in flue gas, which is not affected by PM controls but easily removed by any of the SO<sub>2</sub> controls, including dry sorbent injection (DSI). Dry SO<sub>2</sub> controls will direct Cl into solids that are removed by downstream PM controls.<sup>17</sup> Wet FGDs will direct Cl into the liquid phase.

Within each APCD, the partitioning of trace elements varies across different studies due to differences in coal composition and operating conditions. For simplicity, we assume SCRs and ACI only affect Hg partitioning, DSI only affects Cl partitioning, and all APCDs not included in Table S4 do not affect trace element partitioning.

For studies with multiple samples, we calculate the partitioning factor for each sample and use the average partitioning to represent the study. However, we use the partitioning factor of each sample to represent the study if they sample APCDs that are not well studied so that we can increase the sample size in our bootstrapping analysis. Therefore, we use sample level data for Hg partitioning by FFs and combinations of ACI and PM controls. We also use sample level data for Cl partitioning by DSI. Those studies where we use sample level data are listed in Table S4.

Additionally, some combinations of APCDs are best modeled as a single unit. For example, the particulates created in a dry FGD are typically removed by a FF. We treat the following APCD combinations as a single device: DSI and any PM control device,<sup>17</sup> dry FGDs with any PM control device,<sup>18</sup> and SCRs with wet FGD for Hg.<sup>15</sup> Due to a lack of Cl partitioning data for hot-side electrostatic precipitators (hsESPs) and FFs, we assume those PM controls partition Cl in an identical manner as a

cold-side electrostatic precipitator (csESP). This assumption is not expected to significantly impact the study results because Cl is generally unaffected by PM controls.<sup>13</sup>

Several studies of plants with electrostatic precipitators (ESPs) installed do not report whether the technology is a csESP or a hsESP. We assume that all unspecified ESPs are csESPs, because they are more prevalent than hsESPs (58% of coal generation vs. 8% of coal generation as shown in Table S4 in SI Section 11). Swanson et al.<sup>19</sup> describe that fly ash from cold-side and hot-side ESPs are mixed before trace elements were sampled, and so we assume the partitioning factors for cold-side and hot-side ESPs are the same in that study.

Multiple studies report partitioning of trace elements by wet FGDs. However, only Cheng et al.<sup>13</sup> specify the partitioning of trace elements exiting the wet FGD into the liquid Cl purge and solid gypsum stream, and so we assume these same partitioning between liquid and solid phases hold for other studies. The ratio of liquid to solid trace element exiting the wet FGD based on Cheng et al. is 0.11 for Hg, 0.04 for Se,  $9\text{e-}4$  for As, and 0.98 for Cl.

**Table S4:** Removal fraction of trace elements listed by study, boiler or air pollution control, and pollutant for studies included in the model.

		Removal by boiler				Removal by PM control				Removal by SO <sub>2</sub> control				Source	Data type	Other
	APCDs	Hg	Se	As	Cl	Hg	Se	As	Cl	Hg	Se	As	Cl			
Brekke et al. (1995) <sup>9</sup>	csESP					0.3	0.2							Coal plants	Removal fractions	
	FF					0.6	0.65									
Brown et al. (1999) <sup>10 a, b</sup>	csESP, wFGD					0.3				0.411				Coal plants	Reported from another study	assume csESP
	FF					0.65										
						0.85										
						0.90										
						0.55										
						0.65										
						0.68										
						0.95										
						0.99										
						0.98										
						0.99										
						0.50										
						0.65										
						0.67										
Cheng et al. (2009) <sup>8</sup>	SCR, csESP, wFGD	0.003	0.069	0.064	0.004	0.174	0.604	0.946	0.024	0.946	0.86	0.997	0.988	Coal plants	Mass flows	assume csESP
Chu and Porcella (1995) <sup>11</sup>	csESP, wFGD					0.3				0.15				Coal plants		
	FF					0.3										
Devito et al. (2002) <sup>12</sup>	csESP, wFGD	0.029				0.161				0.444				Coal plants	Removal fractions	assume csESP
Felsvang et al. (1994) <sup>18</sup>	dFGD, FF									0.333	0.998	0.999		Coal plants	Removal fractions	

	ACI, dFGD, FF								0.990						
Flora et al. (2003) <sup>13, b</sup>	ACI, FF				0.122								Coal plants	Removal fractions	
					0.841										
					0.887										
					0.682										
					0.31										
					0.428										
					0.656										
					0.627										
					0.037										
					0.054										
					0.90										
					0.875										
					0.796										
					0.965										
					0.808										
					0.732										
					0.328										
					0.967										
					0.925										
					0.918										
					0.877										
					0.862										

						0.102										
Helble (2000) <sup>14</sup>	csESP					0.289	0.491	0.961						From EPRI and DoE		assume csESP
Karlsson (1984) <sup>20</sup>	dFGD, FF												0.97	Coal plants	Removal fractions	
Klein et al. (1975) <sup>15</sup>	csESP	0.019	0.163	0	0.006	0.038	0.038	0.805	0.818	0.011				Coal plants	Mass flows	
Laird et al. (2013) <sup>17, b</sup>	DSI												0.917	Coal plants	Mass flows	
													0.989			
													0.9996			
													0.969			
													0.969			
													0.71			
													0.78			
													0.64			
													0.84			
													0.77			
Laudal et al. (2000) <sup>16</sup>	csESP, wFGD					0.034				0.19				Coal plants	Mass flows	assume csESP
NRMRL (2005) <sup>17, b</sup>	ACI, csESP					0.65								Coal plants	Removal fractions	
						0.70										
						0.73										
						0.85										
						0.94										
						0.94										
	ACI, hsESP					0.40										
						0.80										
Ondov et al. (1979) <sup>18</sup>	csESP, wFGD						0.941	0.921	0		0.05	0.075	0.964	Coal plants	Removal fractions	
Pavlish et al. (2003) <sup>30 a</sup>	csESP, wFGD					0.233				0.403				From 1999 EPA ICR	Removal fractions	

	hsESP, wFGD					0.128				0.29						
	FF, wFGD					0.323				0.103						
Rubin (1999) <sup>20 c</sup>	csESP	0.008	0.015	0.014	0.001	0.254	0.252	0.595	0.96	0				From model	Removal fractions	assume csESP
Swanson et al. (2013) <sup>7</sup>	csESP	0	0	0.04		0.02	0.2	0.5						Coal plants	Removal fractions	
	hsESP	0	0	0.04		0.02	0.2	0.5								

<sup>a</sup> We incorporate data from only one of the studies cited within the review paper, because the other studies overlapped with another study already included in our table. We cite the review paper, because we were unable to obtain copies of the original paper cited in the review.

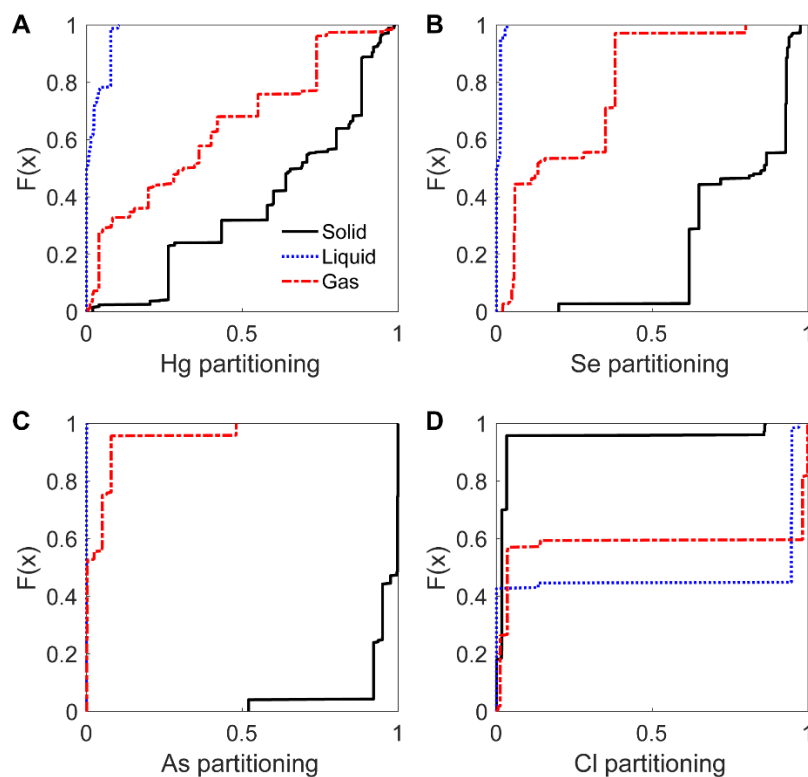
<sup>b</sup> To increase sample size for the bootstrapping analysis, we use sample-level data for this study.

<sup>c</sup> Rubin estimated partitioning of trace elements using data from a combination of empirical samples and literature data.<sup>21</sup> His results are summarized as point estimates, which we use for bootstrapping.

Notes: Studies either sampled data from CFPPs or perform secondary analysis using reported data from other studies. The partitioning of trace elements is reported in either fractions removed or in mass flows across different air pollution controls. Plants that did not specify if they were csESP or hsESP were assumed to be cold-side ESP. Only studies that included U.S. CFPPs are included. The controls are: ACI – activated carbon injection; csESP – cold-side electrostatic precipitator; dFGD – dry flue gas desulfurizer, DSI – dry sorbent injection, FF – fabric filter; hsESP – hot-side electrostatic precipitator; SCR – selective catalytic reactor; wFGD – wet flue gas desulfurizer.

## SI Section 9: Partitioning of trace elements to the solid, liquid, and gas phase at each coal boiler

In Figure S9, we show the cumulative distribution functions of bootstrapped partitioning fractions of trace elements into solid, liquid, and gas phases for each U.S. coal boiler in our dataset. As expected, a small fraction of trace metals and a large fraction of Cl will end up in liquid waste. For CFPPs with a wet FGD, trace element partitioning into the liquid phase ranges from 0.003 to 0.105 for Hg, 0.005 to 0.035 for Se,  $4 \times 10^{-5}$  to  $9 \times 10^{-4}$  for As, and 0.13 to 0.97 for Cl. Trace element partitioning into the solid phase ranges from 0.02 to 0.99 for Hg, 0.20 to 0.98 for Se, 0.52 to 1.0 for As, and 0.001 to 0.86 for Cl. The partitioning factors to solid phase observed in Figure S9 are consistent with the chemistry of these trace elements, as less volatile trace elements, such as Se and As, are more likely to condense out of the gas phase, while less volatile elements, such as Hg and Cl are less likely to condense out.<sup>19,22</sup>



**Figure S9:** Bootstrapped cumulative distribution functions of trace element partitioning into solid, liquid, and gas phases at each boiler across the coal fleet for A) Hg, B) Se, C) As, and D) Cl. Figure produced by the author using data from Table S4.



## SI Section 10: Comparing partitioning fractions estimated matching individual air pollution control devices against matching combinations of air pollution control devices

In the manuscript, we estimate trace element partitioning fractions at the boiler-level by estimating the partitioning fraction for each APCD downstream of the boiler and combining as shown in Equation S1.

$$x_{boiler,\phi,j}^{link} = x_{PM,\phi,j} + (x_{PM,\phi=gas,j} \times x_{SO_2,\phi,j}) \quad (S1)$$

Where  $x_{boiler,\phi,j}^{link}$  is the overall partitioning fraction for the boiler for a phase  $\phi$  and a trace element  $j$ . Also,  $x_{PM,\phi,j}$  is the partitioning fraction at the PM control and ACI and  $x_{SO_2,\phi,j}$  is the partitioning fraction at the SO<sub>2</sub> control and the SCR, both of which are derived from the literature. We call this method the “link” method.

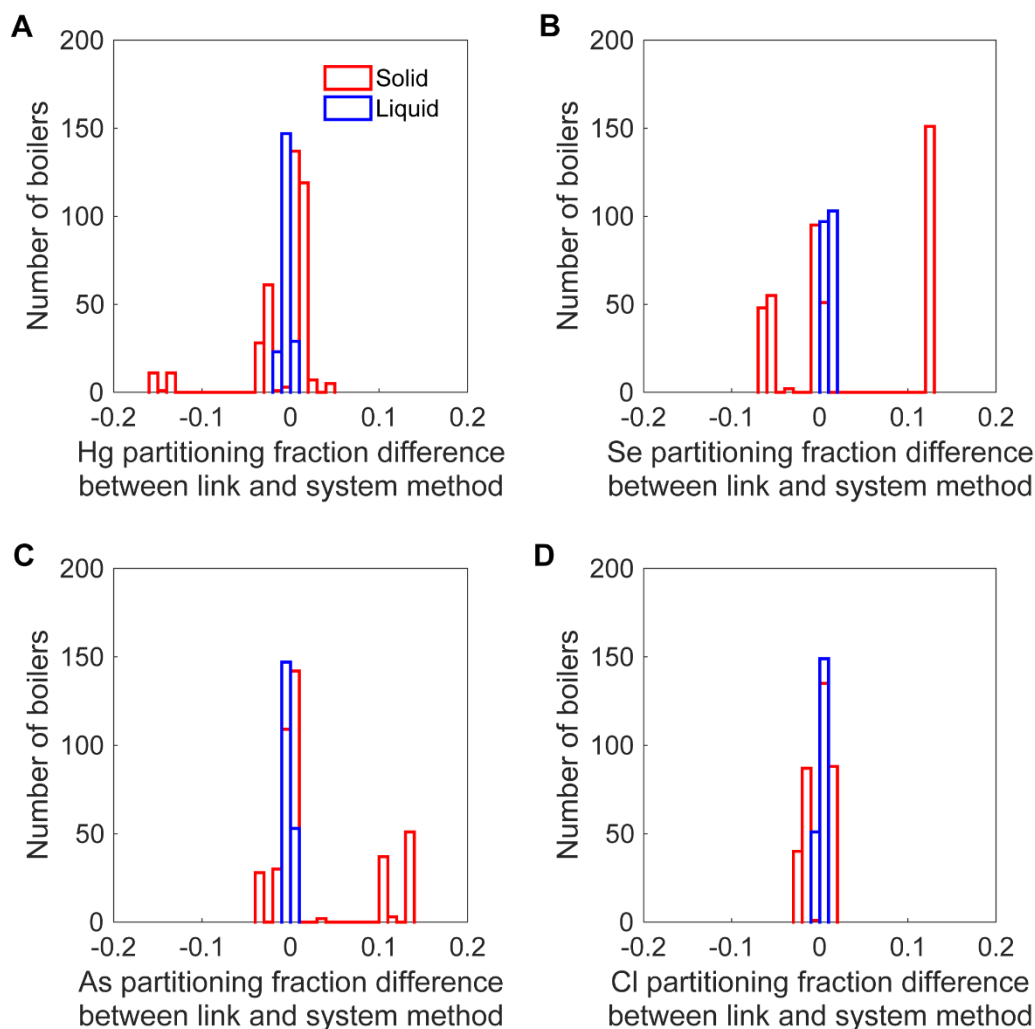
The method presented in Equation S1 assumes that partitioning is independent of the plant processes before or after the air pollution control. To test this assumption of independence, we estimate the partitioning fraction of trace elements at the boiler-level by matching the combination of APCDs downstream of the boiler against the exact combination of controls in the literature data. This new formulation is shown in Equation S2.

$$x_{boiler,\phi,j}^{system} = x_{PM,SO_2,\phi,j} \quad (S2)$$

Where  $x_{PM,SO_2,\phi,j}$  is the partitioning fraction of trace element  $j$  into phase  $\phi$  for the combination of PM and SO<sub>2</sub> controls. We call this method the “system” method.

We assumed partitioning by APCDs is independent to model mass flow rates from a greater fraction of the fleet. We can estimate mass flow rates for 89% of coal generation using the link method but only 67-89% of coal generation using the system method.

In Figure S10, we show a histogram of the difference between the partitioning fraction calculated using the link method and the system method. For all four trace elements, we observe differences of at most two percent in the partitioning fraction to liquid phase across the two methods. For partitioning to solids, we observe differences greater than 10% between the link and system method for Hg, Se, and As. The differences in partitioning are driven by the method that determines which set of literature studies included. Because of the variability in trace element partitioning observed across literature studies for each APCD, the median partitioning fraction will shift depending on which studies are included. Therefore, the assumption of independence across different APCDs will not affect partitioning estimates for most coal boilers, but will affect partitioning estimates of some specific trace elements and pollution controls.



**Figure S10:** Histograms illustrating differences of boiler-level estimates of trace element partitioning fractions derived from literature by matching literature data with individual APCD and matching literature data with combinations of air pollution control devices for the trace elements A) Hg, B) Se, C) As, and D) Cl.

## SI Section 11: Generation treated by different air pollution controls and control combinations

In Table S5, we show the amount of coal generation and the percent of total coal generation in 2015 for each of the postcombustion control options installed to treat Hg, NO<sub>x</sub>, PM, and SO<sub>2</sub>. We use data from EIA-860, Schedule 6 and electricity generation at the boiler level from EIA-923, Schedule 3.<sup>1,3</sup> EIA-923, Schedule 3 reports 1330 TWh generated from coal in 2015.<sup>1</sup>

The most widely used APCDs for Hg, NO<sub>x</sub>, PM, and SO<sub>2</sub> are ACI systems, SCRs, csESPs, and wet FGD systems, respectively. These systems respectively treat 33%, 45%, 58%, and 57% of total coal generation. About 57% of coal generation is not treated by primary Hg controls in 2015, 38% is not treated by NO<sub>x</sub> controls, and 18% is not treated by SO<sub>2</sub> controls. All generation is treated by a PM control. As explained in greater detail in SI Section 3, we do not account for roughly 10% of coal generation from boilers due to data limitations in our datasets. Although a large portion of the generation fleet does not use ACI to remove Hg, other APCDs, such as csESP, hsESP, FFs, and wet FGD, which are not listed by the EIA as Hg pollution controls, remove some Hg.<sup>15</sup>

**Table S5:** Coal electricity generation treated by different postcombustion APCDs (in TWh) and as a percent of total coal generation in 2015.

	Air pollution control	Generation treated (TWh)	Generation treated / total coal generation (%)
Hg	None	755	57%
	Activated carbon injection	435	33%
NO <sub>x</sub>	Selective catalytic reactor	603	45%
	None	510	38%
	Selective non-catalytic reactor	77	6%
	Other / Unnamed	0	0%
PM	Cold-side electrostatic precipitator	770	58%
	Fabric filter	382	29%
	Wet flue gas desulfurizer	178	13%
	Hot-side electrostatic precipitator	108	8%
	Cyclone	14	1%
	Other / Unnamed	26	2%
SO <sub>2</sub>	Wet flue gas desulfurizer	753	56%
	None	259	19%
	Dry flue gas desulfurizer	146	11%
	Dry sorbent injection	59	4%

Note: Table produced by the authors using data from EIA-860, Schedule 6 and EIA-923, Schedule 3.<sup>1,3</sup>

## SI Section 12: Comparing partitioning to gas from MATS against literature data

The MATS ICR reports concentrations of trace elements in coal, the heating value of coal, and the heat input normalized trace element mass flow rate exiting the stacks of 278 coal boilers. From that we can estimate the partitioning of trace elements to gas by APCDs and compare those to our bootstrapped estimates.

The MATS ICR sampled 278 coal boilers for trace element concentration in coal and trace element mass flows at the stack. For each boiler in the MATS ICR, we estimate total trace element solid and liquid removal by taking the difference between the mass of trace elements entering the boiler and the mass of trace elements exiting the stacks. After we merge all boilers from the mass flow rate analysis to boilers in the MATS ICR, we compare Hg, Se, As, and Cl removal for 107, 80, 66, and 61 boilers, respectively. Boilers are excluded because the MATS ICR does not sample all four trace elements at every boiler. The unique boilers in this analysis represent 22% of total coal generation in 2015.

We compare our estimated partitioning of trace elements to gas from the literature against the partitioning of trace elements calculated from the 2010 MATS ICR. The MATS ICR reports mass flow rates to gas but not solids or liquids. We estimate a partition fraction  $x_{gas}$  for each boiler in the MATS ICR dataset using the median trace element concentrations, median heat input normalized mass flow rate, and median heating values reported at the unit level as shown in Equation S3.

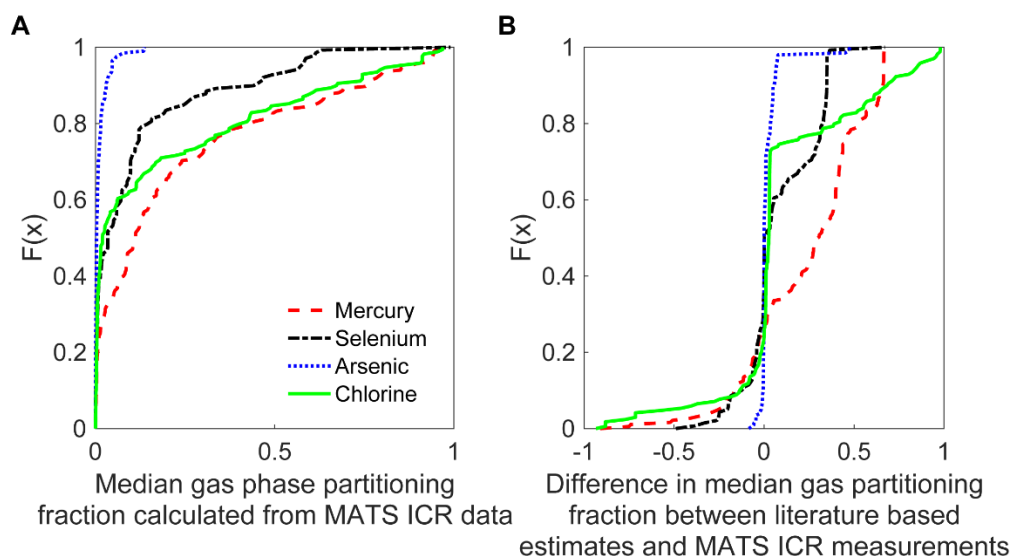
$$x_{gas}^{MATS} = \frac{\text{heat input normalized mass flow rate } [\frac{lb}{BTU}]}{\text{concentration of trace elements } [ppm] * 10^{-6} [ppm^{-1}] * \text{higher heating value } [\frac{BTU}{lb}]} \quad (S3)$$

Between our dataset and the MATS ICR dataset, there are 107, 80, 66, and 61 boilers with Hg, Se, As, and Cl removals, respectively. These boilers account for roughly 25% of total coal generation in 2010.

We find that our trace element partitioning estimates are comparable to those reported in the MATS ICR, with As and Cl showing better agreement than Hg and Se. In Figure S11A, we show the partitioning of trace elements to gas calculated from the MATS ICR dataset. We find that the median partitioning of trace elements to solids and liquids in the MATS ICR are 0.11 for Hg, 0.03 for Se, 0.01 for As, and 0.02 for Cl. For context, the median partitioning to gas estimated from the literature is 0.35 for Hg, 0.21 for Se, 0.05 for As, and 0.03 for Cl.

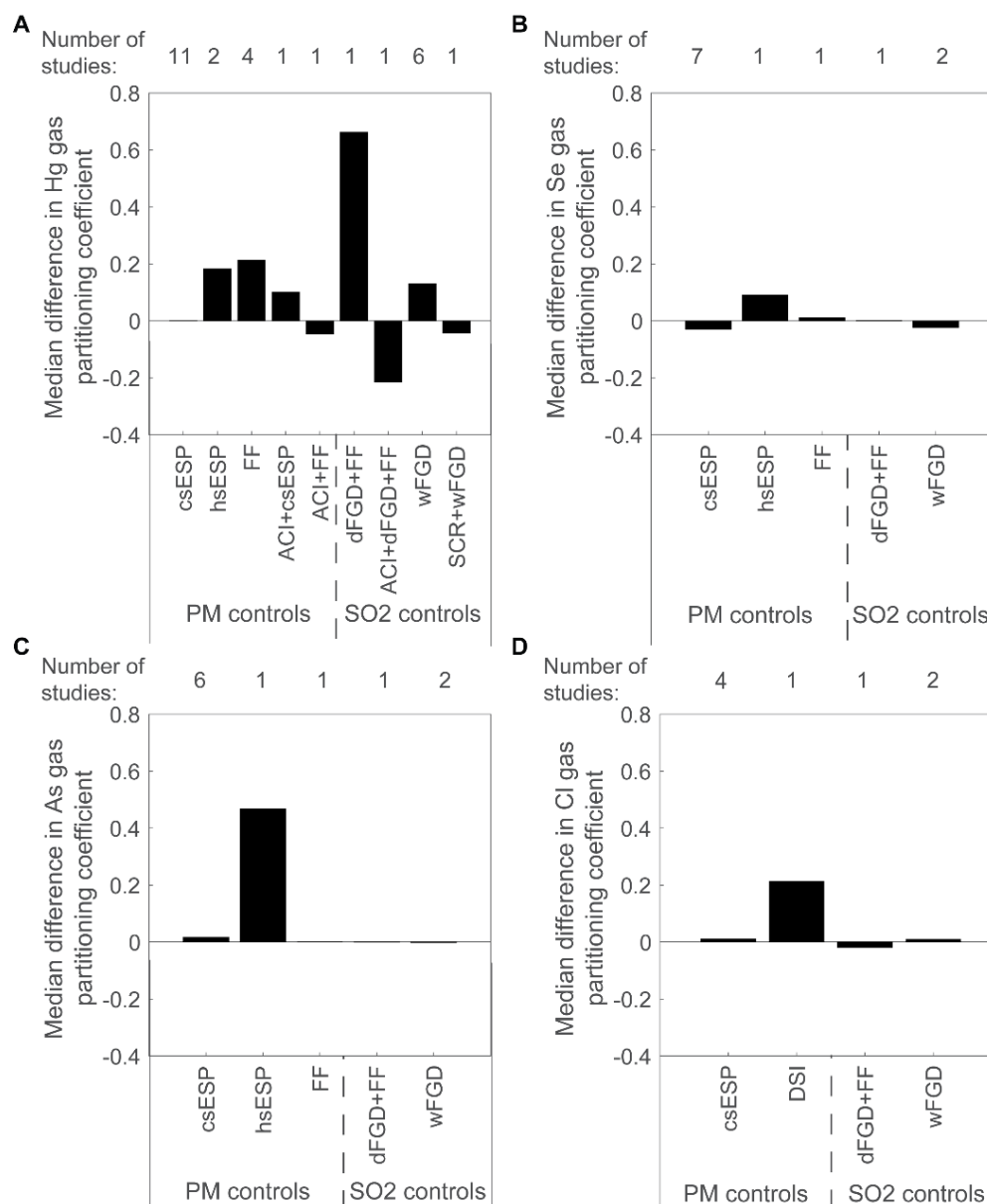
In Figure S11B, we show the difference between partitioning into solids and liquids from our dataset versus the MATS ICR dataset. The difference between median gas partitioning estimates from the literature and the MATS ICR is 30% for Hg, 0.18% for Se, 0.074% for As, and 2.9% for Cl. However, at some boilers we find significant errors in estimates of partitioning fraction for all trace elements. Therefore, the mean difference between the partitioning fraction to solids and liquids between the literature versus the MATS ICR is 24% for Hg, 8.5% for Se, 1.9% for As, and 11% for Cl.

We expect the differences in trace element concentrations to contribute more to the overall total estimate differences than differences in partitioning, because the partitioning estimates from the literature are in reasonable agreement with the partitioning from the MATS ICR.



**Figure S11:** A) Cumulative distribution functions of median trace element partitioning to gas calculated from the reported MATS ICR data. B) Cumulative distribution functions of differences between median removal of trace elements from bootstrapping the literature and median gas partitioning of trace elements calculated from reported MATS ICR data. Figure produced by the authors using data from Table S4, EIA-860, EIA-923 Schedule 5, and MATS ICR.<sup>1,3,23</sup>

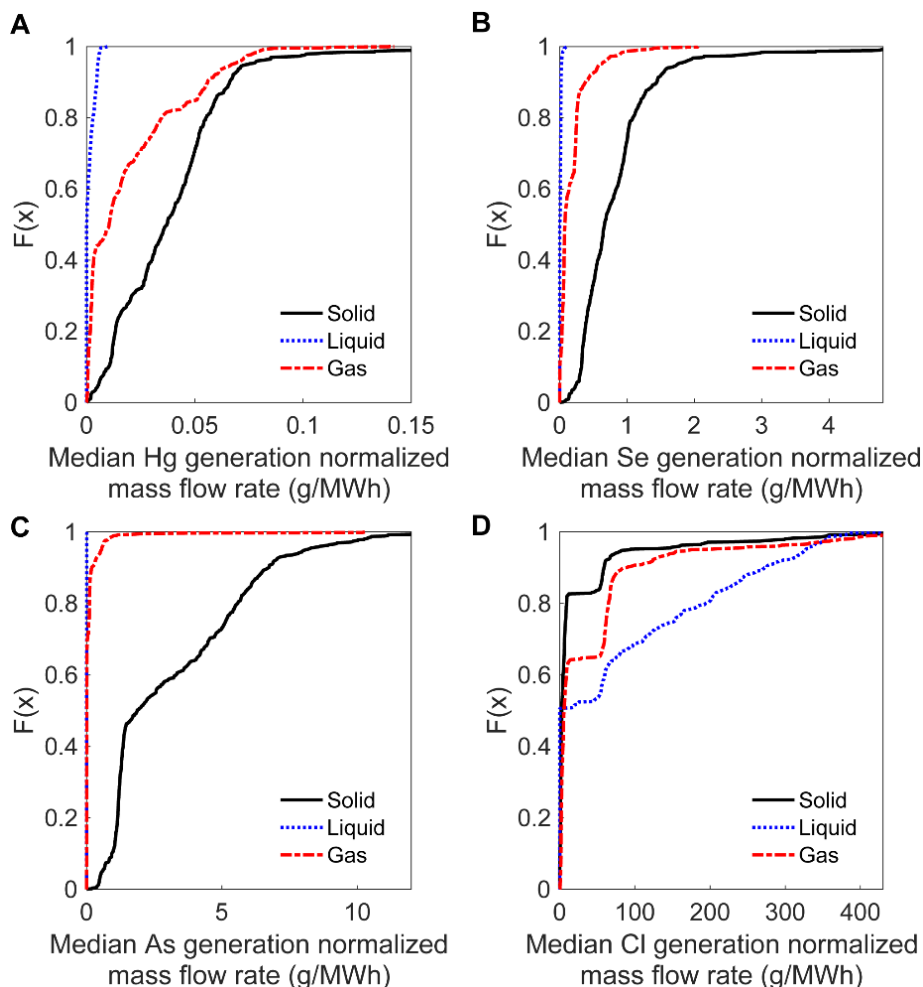
One explanation for our errors is that we have systematic biases in predicting trace element partitioning for certain APCDs. In Figure S12, we plot the median errors between our estimates and the values observed in the MATS ICR data set. The only APCD we find a systematic bias in is for hseSPs, leading us to potentially overestimate gas phase partitioning at the 8% of plants with a hseSP. For all other APCDs, we do not find a tendency to consistently over- or under-estimate the partitioning to the gas phase.



**Figure S12:** Median difference in gas partitioning coefficient between models estimates and MATS ICR estimates at the boiler-level for the trace elements A) Hg, B) Se, C) As, and D) Cl decomposed by PM and SO<sub>2</sub> controls. The number of literature studies used to estimate partitioning of each air pollution control device and device combination is included. While we have data for Hg partitioning by ACI+hsESP, there are no plants in the validation dataset that have that set of controls installed downstream. The air pollution control technologies listed are: ACI = activated carbon injection; csESP = cold side electrostatic precipitator; dFGD = dry flue gas desulfurizer, DSI = dry sorbent injection; FF = fabric filter; hsESP = hot side electrostatic precipitator; SCR = selective catalytic reactor; wFGD = wet flue gas desulfurizer.

### SI Section 13: Generation normalized mass flow rates of trace elements at the boiler level in the U.S. and in different eGRID sub-regions

In Figure S13, we show the median generation normalized mass flow rate of each trace element from coal boilers across the fleet.



**Figure S13:** Cumulative distribution functions of median generation normalized mass flow rate into solid, liquid, and gas phases from bootstrapped coal concentrations in coal at each plant and trace element partitioning fractions for air pollution controls at each boiler across the coal fleet for A) Hg, B) Se, C) As, and D) Cl. The x-axis on all figures is cut-off at the 99<sup>th</sup> percentile (0.15 g/MWh for Hg, 4.6 g/MWh for Se, 11 g/MWh for As, and 406 g/MWh for Cl) for readability. Figure produced by the author using data from COALQUAL, Table S4, EIA-860, and EIA-923.<sup>1-3</sup>

We have included eGRID subregion figures in our data repository here:

[https://github.com/xdansun/trace\\_elem\\_cfpp\\_model/tree/master/Figures/subrgn\\_emfs](https://github.com/xdansun/trace_elem_cfpp_model/tree/master/Figures/subrgn_emfs). We encourage the reader to explore the repository and perform their own additional analysis.

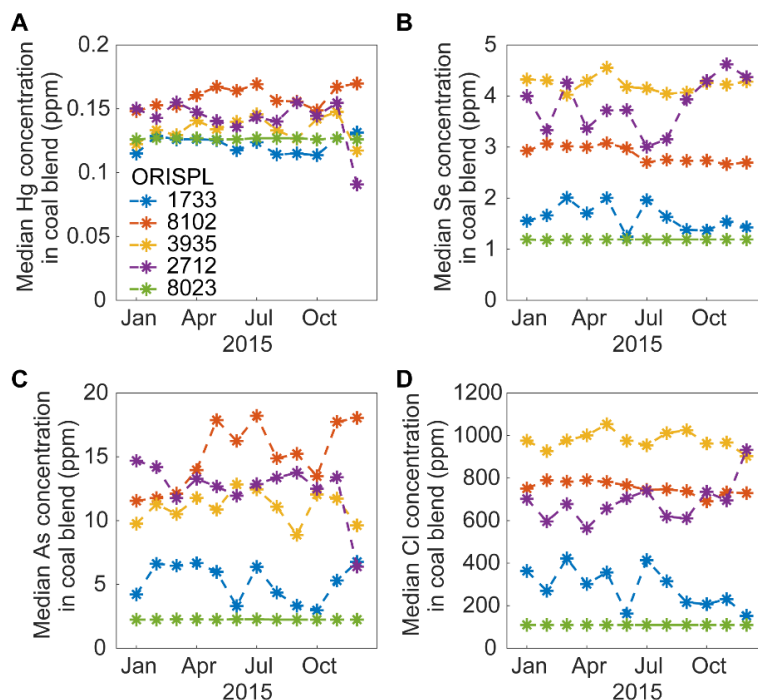
## SI Section 14: Temporal variability of trace element concentrations in coal

Coal plants purchase coal in different amounts across different counties from month to month. In Figure S14 we show the 2015 monthly median of the bootstrapped concentration of trace elements in coal at five example CFPPs. We estimate the concentration of trace elements for each month employing the same bootstrapping approach mentioned in the Methods section of the main manuscript using data from COALQUAL and coal purchase data at that month from EIA 923/5.<sup>1,2</sup>

The five plants we chose are the Monroe plant in Michigan (ORISPL: 1733), the Gavin plant in OH (ORISPL: 8102), the John E Amos plant in WV (ORISPL: 3935), the Roxboro plant in NC (ORISPL: 2712), and the Columbia plant in WI (ORISPL: 8023). These 5 plants generated 16 TWh, 16 TWh, 15 TWh, 4.9 TWh, and 4.9 TWh respectively.

We find that the concentration of trace elements varies temporally for a fraction of plants. In Figure S14, the concentration varies from month to month for four plants (ORISPL: 1733, 8102, 3935, and 2712). Concentrations in the coal blend can vary significantly for all four trace elements, such as the coal blend at Roxboro (2712). In other cases, the concentration varies only for one trace element, which is seen for the coal blend at Gavin (8102).

Lastly, the concentration will not vary across time for plants that purchase coal from a single county, such as the coal blend of Columbia (8023). There are 111 out of 276 CFPPs in our analysis that purchase from a single county. Those plants generated 370 TWh in 2015, which makes up 28% of total coal generation in 2015.



**Figure S14:** Monthly median bootstrapped concentrations of trace elements for five example plants. Data is produced from COALQUAL and EIA-923/5.<sup>1,2</sup>



## **SI Section 15: Comparison of concentrations of trace elements in coal using COALQUAL and EIA against MATS ICR reporting**

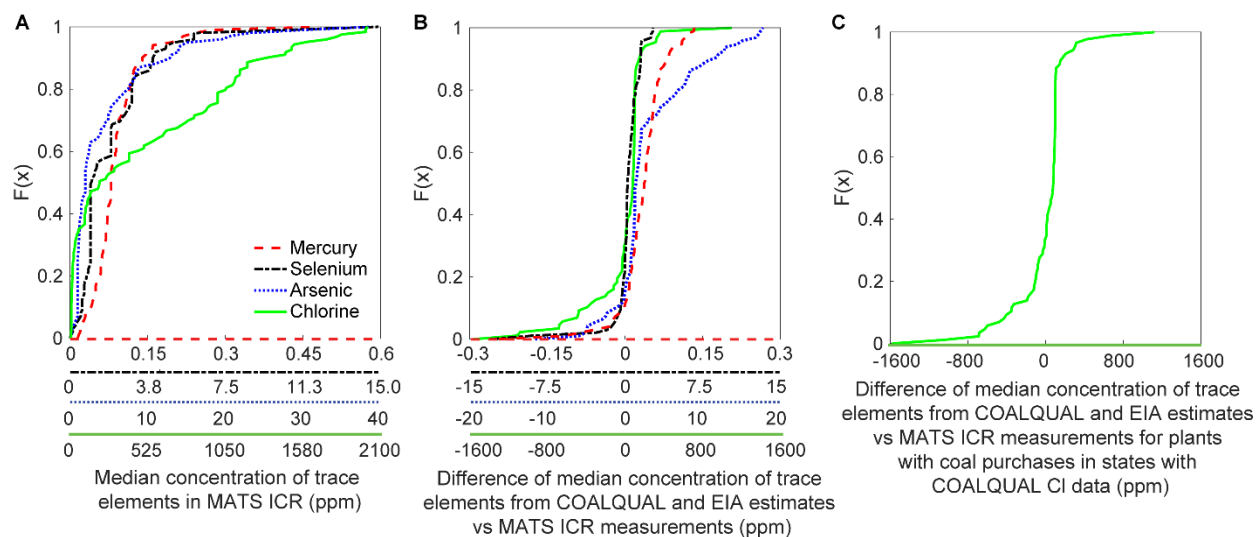
The 2010 MATS ICR is a data set resulting from an EPA program that sampled concentrations of trace elements in coal prior to combustion and sampled trace element emission rates from the stacks of CFPPs.<sup>23</sup> This dataset is useful for comparing our estimates of trace element concentrations in coal, since the MATS ICR samples coal after any coal cleaning processes might have taken place. The MATS ICR contains coal concentration data for 210 CFPPs, with anywhere from three to 54 samples per trace element per plant, but not every CFPP was sampled for every trace element. Of those 210 CFPPs, 167 plants have coal data for Hg, 161 for Se, 164 for As, and 201 for Cl. The total generation of the coal plants sampled in the MATS ICR was 573 TWh, or roughly 43% of total U.S. coal generation in 2010. After we merge all plants from the coal concentration analysis to the plants in MATS ICR, only 129 plants remain in the analysis.

Although the MATS ICR is a recent database, sampling by MATS ICR may not be representative of trace element concentrations combusted at plants year-round because sampling took place over a few weeks in 2010, and it will not capture changes in the mix of coal purchased by power plants over the course of a year (see SI Section 14 for additional information on the temporal variability of coal purchases from CFPPs).

Additionally, MATS ICR data was collected in 2010. Thus, to produce a viable comparison between our method and MATS ICR observations, we estimate concentrations of trace elements in coal using our method, but using coal purchase data from for the same months in 2010 as those in which MATS ICR data was collected.<sup>1</sup> Figure S15A shows the cumulative distribution functions of the median trace element concentrations from sampled coal concentrations in MATS ICR and Figure S15B provides the difference between bootstrapped median trace element concentrations and the median concentrations from sampled coal concentrations in MATS ICR. Our median estimates for trace element concentrations are 0.13 ppm for Hg, 1.4 for Se, 3.3 ppm for As, and 122 ppm for Cl and we find the median concentrations of coal in MATS ICR are 0.08 ppm for Hg, 1.0 ppm for Se, 2.0 ppm for As, and 200 ppm for Cl.

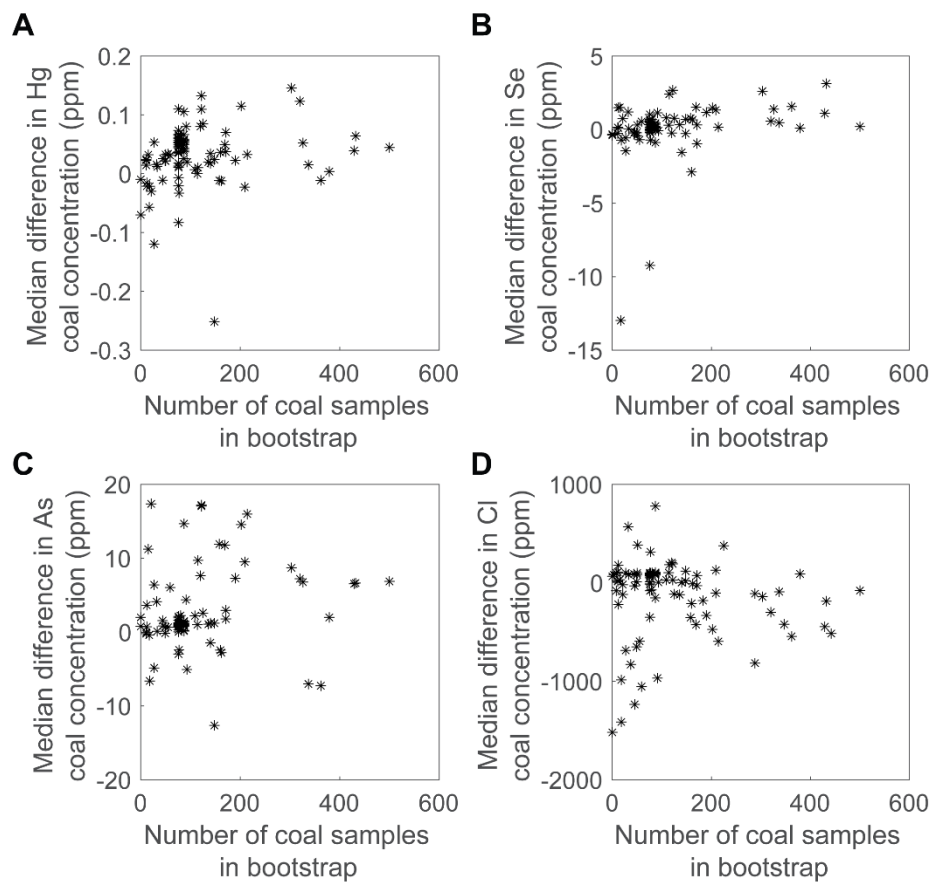
Of the 260 CFPPs in our dataset, there are 46 CFPPs that purchase coal from states without Cl coal samples. These plants represent 13% of all CFPPs and 22% of coal generation. For those states, we approximate Cl concentrations using COALQUAL samples at the basin level. The other trace elements have enough data at the state-level and do not require basin-level data. In Figure S15C, we re-create Figure S15B but only for CFPPs that purchase coal from counties and states with Cl coal data. The median Cl concentration in the coal blend for CFPPs with county- and state-level data is 113 ppm and the median Cl concentration for the same plants listed in the MATS ICR is 108 ppm. The Cl coal concentration estimates improve when we use this smaller dataset, because we interpolate the concentration from a more location specific area.

Our estimates are on the same order of magnitude as the concentrations reported in the MATS ICR, where we get the best agreement with Se and the least agreement with As. We note that our estimates tend to overestimate the concentrations of trace elements reported by the MATS ICR, which means that we are also likely to overestimate mass flow rates from power plants.



**Figure S15:** Cumulative distribution functions of A) median concentrations in coal combusted at the CFPPs reported in the MATS ICR data set, B) differences of median concentrations in coal combusted at the plant-level estimated from our model and reported from the MATS ICR, and C) median CI concentrations in coal combusted at the plant-level estimated from our model and reported from the MATS ICR only for plants with coal purchases from states with CI COALQUAL data. Figure produced by the authors using data from COALQUAL, EIA-923 Schedule 5, and MATS ICR.<sup>1,2,23</sup>

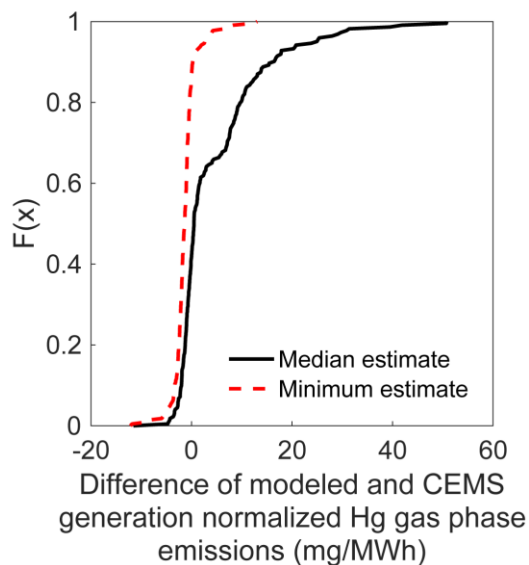
Additionally, the error in our estimates is not correlated to the number of coal samples used in the bootstrapping process, as shown in Figure S16, indicating that our results are not systematically biased based on the number of COALQUAL coal samples in the county.



**Figure S16:** Median difference of A) Hg, B) Se, C) As, and D) Cl at the plant-level between model estimates and MATS ICR measurements as a function of the number of total coal samples in COALQUAL used in the bootstrapping.

## SI Section 16: Comparison of minimum model estimates of gas phase Hg mass flows against CEMS

Figure S17 plots the CDF of the difference between Hg CEMS emissions and modeled boiler-level calculated emissions using the median Hg concentration in coal and Hg partitioning to the gas phase and the minimum Hg concentration in coal and Hg partitioning to the gas phase. As shown in Figure S17, we find we overestimate Hg emissions for 58% of 224 boilers in our Hg CEMS validation. When we select the minimum Hg concentrations and minimum gas-phase Hg partitioning coefficient, we still overestimate Hg CEMS emissions for 15% of boilers in our Hg CEMS validation. This suggests that there are fundamental issues with the quality of published data on the concentration of Hg in the coal and/or the partitioning coefficients for gas-phase Hg. For example, it is possible that COALQUAL has yet to sample low Hg concentration coals or that the coal mining and coal purchasing process can select for coal with lower ash content, which may lead to lower Hg content. We posit that our model would more closely match CEMS data if our input data quality more accurately reflected the Hg concentrations in coal that is combusted in plant boilers and/or removed in APCDs. In short, the inaccuracies should not be viewed as grounds to dismiss the model, but rather highlight the need for greater data fidelity.



**Figure S17:** Difference of modeled and CEMS Hg gas phase emissions where the modeled emissions are either the median result from the manuscript (Median estimate) or use the lowest reported Hg concentration in coal and the lowest gas phase Hg partitioning fraction (Minimum estimate).

## **SI Section 17: Benchmarking estimates of gas-phase, generation-normalized mass flow rates against separate model results**

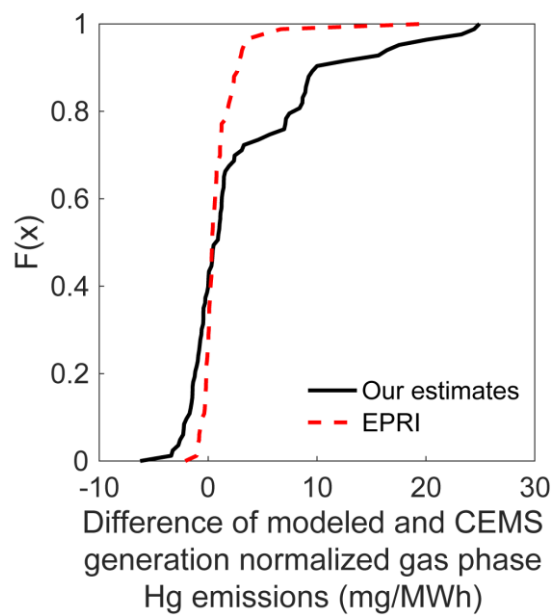
In 2018, the Electric Power Research Institute (EPRI) estimated the hazardous air pollutant emissions from U.S. coal-fired electric generating units for the baseline year 2017, but they do not perform a separate validation analysis.<sup>24</sup> Here we validate their results for Hg emissions and compare their results with our results for As, Se, and Cl emissions to evaluate the relative strengths and weaknesses of the two emissions models. We also validate both our work and EPRI's analysis against 2015 Hg emissions from CEMS, because EPRI used coal consumption and fuel purchase data from 2015.<sup>24</sup>

In Figure S18, we show the difference of modeled Hg gas phase emissions and CEMS reported Hg gas phase emissions, where the modeled emissions are generated by either our median bootstrapping analysis or EPRI's emissions model, provided in Appendix G of the EPRI report.<sup>24</sup> We note that the EPRI model estimates Hg emissions for 216 CFPPs, which generated 66% of coal-fired electricity in 2015. There are 84 plants in common between EPRI's 216 plants and the 113 plants in our Hg validation analysis. Because EPRI's emission numbers are at the plant-level, we assume that each boiler at the plant has the same generation normalized emissions as the plant. We find that EPRI's emissions model is generally a closer match with the CEMS emission data, but can still differ from CEMS emissions by up to 20 mg Hg/MWh. We find that the error of 68 of the 84 boilers in the EPRI analysis are within 2 mg Hg/MWh of CEMS emissions while 49 of the same 84 boilers are within 2 mg Hg/MWh.

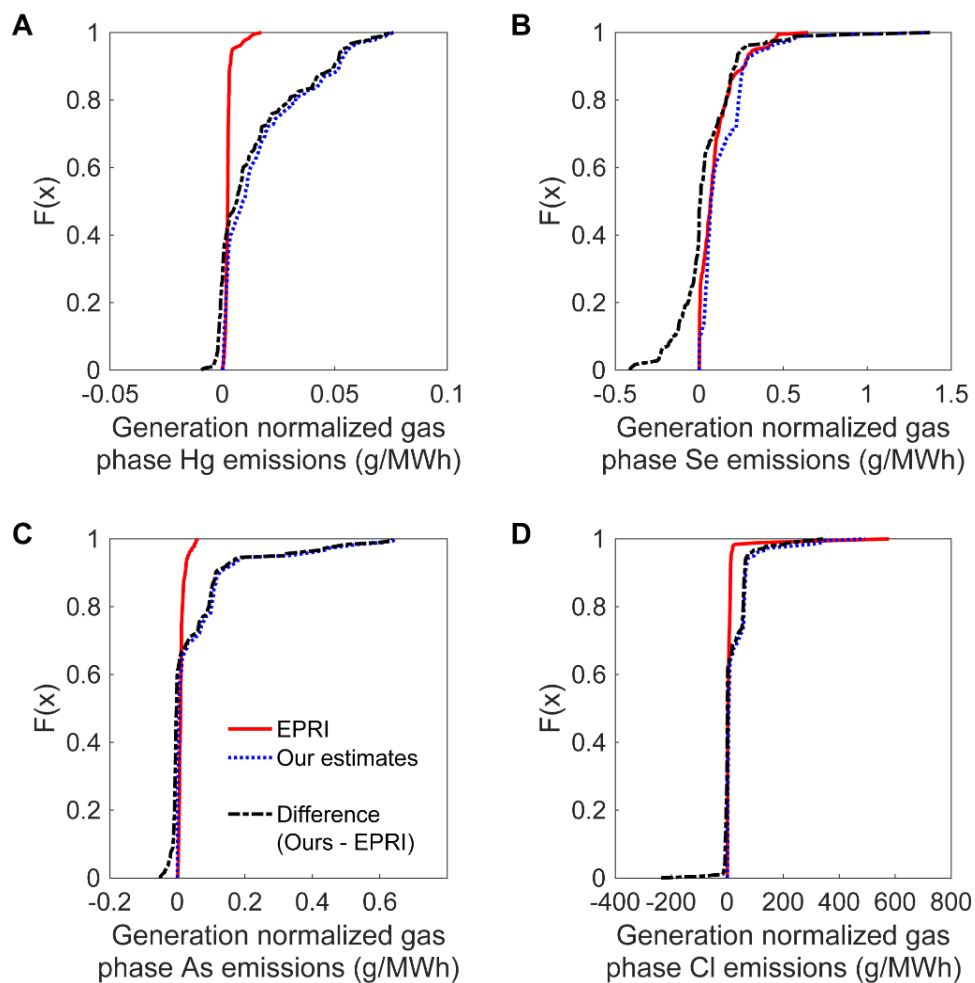
We hypothesize that EPRI's model better fits the CEMS emissions data for three reasons. First, EPRI uses the CEMS data to develop their Hg emission estimates, utilizing the boiler measurement data when available and the average CEMS heat-input normalized emissions (lb/trillion BTU) when the measurement data was not available. Because EPRI is only estimating trace element mass flows to the gaseous phase, they can use the MATS ICR dataset. Second, EPRI also has access to additional coal concentration and plant-level partitioning data obtained from their own sampling efforts. Between MATS ICR and EPRI's internally held datasets, EPRI reports that they have access to at least 524 Hg, 250 Se, 263 As, and 280 Cl sampling events.<sup>24</sup> Third, the larger dataset enables EPRI to use regression based approaches to derive heat-input normalized emissions for each APCD combination as a function of coal trace element content, ash content (for As), and sulfur content (for Se).

EPRI also modeled gas phase emissions for Se, As, and Cl from the same 216 CFPPs they used to model Hg emissions. We benchmark our model's median bootstrapped gas phase emissions against EPRI's gas phase emissions. We find 184 coal plants in common between our data sets, which generated 65% of 2015 coal generation, that are used in the benchmarking.

In Figure S19, we show the results of our benchmarking. The median difference between our results and EPRI's is 5.8 mg Se/MWh, -2.9 mg As/MWh, and 980 mg Cl/MWh. Based on these results, we tend to overestimate gaseous emissions at the plant-level for Hg, Se, and Cl compared to EPRI's estimates. We find our Se, As, and Cl estimates are within  $\pm 30\%$  of EPRI's estimates for 56%, 63%, and 44% of 184 plants, respectively. For Se and As, more than half of our CFPPs show comparable estimates to EPRI's model. We find these results encouraging, as it suggests that if input data quality improves, then the bootstrap model might be comparable or even more accurate than other plant-level estimates.



**Figure S18:** Difference of CEMS reported versus modeled generation normalized gas phase Hg emissions, where the modeled emissions are either the median bootstrapping result from the manuscript (our estimates) or the reported result from EPRI's modeling analysis.



**Figure S19:** Generation normalized gas phase emissions estimated by the EPRI model and by our model alongside the difference between our model and EPRI's for the trace elements A) Hg, B) Se, C) As, and D) Cl.

## Supplementary Information References:

- (1) U.S. Energy Information Administration. Annual Electric Utility Data <http://www.eia.gov/electricity/data/eia923/> (accessed Sep 1, 2017).
- (2) Palmer, C. A.; Oman, C. L.; Park, A. J.; Luppens, J. A. *The U.S. geological survey coal quality (COALQUAL) database version 3.0*; 2015.
- (3) U.S. Energy Information Administration. Annual Electric Generator Data <https://www.eia.gov/electricity/data/eia860/> (accessed Sep 1, 2017).
- (4) U.S. Environmental Protection Agency. Emissions & Generation Resource Integrated Database, eGRID2014 Version 2.0 (year 2014 data) <https://www.epa.gov/energy/emissions-generation-resource-integrated-database-egrid> (accessed Aug 1, 2017).
- (5) Yudovich, Y. E.; Ketris, M. P. Mercury in coal: A review. *Int. J. Coal Geol.* **2005**, 62 (3), 107–134.
- (6) Yudovich, Y. E.; Ketris, M. P. Selenium in coal: A review. *Int. J. Coal Geol.* **2006**, 67 (1–2), 112–126.
- (7) Yudovich, Y. E.; Ketris, M. P. Arsenic in coal: A review. *Int. J. Coal Geol.* **2005**, 61 (3–4), 141–196.
- (8) Yudovich, Y. E.; Ketris, M. P. Chlorine in coal: A review. *Int. J. Coal Geol.* **2006**, 67 (1–2), 112–126.
- (9) Capes, C. E.; McIlhinney, A. E.; Russell, D. S.; Sirianni, A. F. Rejection of trace metals from coal during beneficiation by agglomeration. *Environ. Sci. Technol.* **1974**, 8 (1), 35–38.
- (10) Dutta, A.; Basu, P.; Ghosh, A.; Chakraborty, P. Innovative solution to the problem of mill rejects in thermal power plants. In *18th International Conference on Fluidized Bed Combustion*; 2005; pp 1–6.
- (11) Akers, D. J. The Redistribution of Trace Elements during the Beneficiation of Coal. In *Environmental Aspects of Trace Elements in Coal*; Swaine, D. J., Goodarzi, F., Eds.; Springer Netherlands: Dordrecht, 1995; pp 93–110.
- (12) Chen, H. L.; Pagano, M. The Removal of Chlorine from Illinois Coal by High Temperature Leaching. *Fuel Process. Technol.* **1986**, 13, 261–269.
- (13) Cheng, C.-M.; Hack, P.; Chu, P.; Chang, Y.-N.; Lin, T.-Y.; Ko, C.-S.; Chiang, P.-H.; He, C.-C.; Lai, Y.-M.; Pan, W.-P. Partitioning of mercury, arsenic, selenium, boron, and chloride in a full-scale coal combustion process equipped with selective catalytic reduction, electrostatic precipitation, and flue gas desulfurization systems. *Energy & Fuels* **2009**, 23 (10), 4805–4816.
- (14) Córdoba, P. Status of Flue Gas Desulphurisation (FGD) systems from coal-fired power plants: Overview of the physico-chemical control processes of wet limestone FGDs. *Fuel* **2015**, 144, 274–286.
- (15) National Risk Management Research Laboratory Office. *Control of Mercury Emissions from Coal Fired Electric Utility Boilers: An Update*; Research Triangle Park, NC, 2005.
- (16) Aunela-Tapola, L.; Hatanpää, E.; Hoffren, H.; Laitinen, T.; Larjava, K.; Rasila, P.; Tolvanen, M. A study of trace element behaviour in two modern coal-fired power plants. *Fuel Process. Technol.* **1998**, 55 (1), 13–34.
- (17) Laird, C. P.; Smith, K. J.; Looney, B. Results of Dry Sorbent Injection Testing to Reduce HCl. In *J. Air Waste Manag. Assoc. Power Plant Air Pollutant Control “mega” Symposium 1*; 2012; pp



136–147.

- (18) Felsvang, K.; Gleiser, R.; Juip, G.; Nielsen, K. K. Activated carbon injection in spray dryer/ESP/FF for mercury and toxics control. *Fuel Process. Technol.* **1994**, *39* (1–3), 417–430.
- (19) Swanson, S. M.; Engle, M. A.; Ruppert, L. F.; Affolter, R. H.; Jones, K. B. Partitioning of selected trace elements in coal combustion products from two coal-burning power plants in the united states. *Int. J. Coal Geol.* **2013**, *113*, 116–126.
- (20) Karlsson, H. T. Spray dry scrubbing of secondary pollutants from coal burning. In *Proceedings of the third annual Pittsburgh Coal Conference, MEMS*; 1986; pp 237–252.
- (21) Rubin, E. S. Toxic releases from power plants. *Environ. Sci. Technol.* **1999**, *33* (18), 3062–3067.
- (22) Otero-Rey, J. R.; López-Vilariño, J. M.; Moreda-Piñeiro, J.; Alonso-Rodríguez, E.; Muniategui-Lorenzo, S.; López-Mahía, P.; Prada-Rodríguez, D. Arsenic, mercury, and selenium flue gas sampling in a coal-fired power plant and their fate during coal combustion. *Environ. Sci. Technol.* **2003**, *37* (22), 5262–5267.
- (23) U.S. Environmental Protection Agency. Mercury Air Toxic Standards Information Collection Request <https://www3.epa.gov/airtoxics/utility/utilitypg.html> (accessed Sep 1, 2017).
- (24) Electric Power Research Institute. Hazardous Air Pollutants (HAPs) Emission Estimates and Inhalation Human Health Risk Assessment for U.S. Coal-Fired Electric Generating Units; 2018.