Lead and Livestock: Estimating India's Bovine Lead Exposure*

Greg Ferraro[1], Bret Ericson[2], Andrew M. Simons[1], Emily Nash[3], and Mohammed Kabir[3]

[1]Dept. of Economics, Fordham University
[2]Dept. of Earth and Environmental Sciences, Macquarie University
[3]Pure Earth

November 17, 2020

Abstract: We provide an annual bovine lead exposure estimate for India utilizing random forest modeling, PureEarth's Toxic Sites Identification Program database, and the FAO's Gridded Livestock dataset. In aggregate, India suffered 3,273 bovine fatalities, \$ 21,437,575 of economic damages, and 23.1 km^2 of lethally contaminated land across 222 used lead acid battery recycling sites. This implies per-site averages of 14.7 fatalities, \$ 96,566 in economic damages, and 0.10 km^2 of lethally contaminated land. Without remediation efforts, damages are likely to repeat every year. While this is a conservative estimate, the identified per-site damages indicate an unrecognized severe burden on the rural poor.

---

[*] Lead author and contact: Greg Ferraro, ghferraro@gmail.com. We would like to thank Dr. Johnathon Cooper, who provided a veterinarian's review of the paper's assumptions on livestock toxicology. Financial support was provided by the United States Agency for International Development (USAID) for the Toxic Sites Identification Program (TSIP). The opinions expressed herein are those of the authors and do not necessarily reflect the views of the study sponsors. The funders had no role in study design, data collection, analysis or data interpretation. We thank Fordham University's Graduate School of Arts and Sciences for student travel funding to visit associated used lead acid battery recycling sites in India.

1 Introduction

Lead exposure outcome assessments generally focus on human health, but livestock lead exposures may also represent a considerable, and previously under examined, negative environmental externality. We estimate bovine mammal (BM) lead poisoning at used lead acid battery (ULAB) recycling sites in rural India.

Lead is a widely used industrial input with an annual global demand exceeding 10 million tons, more than half of which is met through secondary smelting (i.e., recycling) (International Lead Association, 2014). Approximately 85 percent of the lead used worldwide goes into the production of lead-acid batteries (International Lead Association, 2014). These batteries are used in traditional and electric vehicles, back-up power supplies, critical systems such as hospitals and telecommunications, and for green technologies such as photovoltaic and wind turbine energy storage (World Health Organization, 2017). While this recycling takes place in regulated and monitored facilities across the US and Europe, it occurs frequently in informal and unregulated settings in low- and middle-income countries (LMICs). A 2016 study suggested anywhere between 10,599 to 29,241 informal ULAB recycling sites existed across 90 LMICs (Ericson et al., 2016). Demand for ULAB recycling activity will likely continue to grow as, for example, the number of new vehicles sold in LMICs more than tripled between 2000 and 2018 (Organisation Internationale des Constructeurs d'Automobiles (OICA), 2016).

While the demand for lead recycling is likely to remain high, informal ULAB recycling is a major source of environmental contamination and human lead exposure (World Health Organization, 2017; UNICEF and PureEarth, 2020). Recycling operations are often located in backyards, where unprotected workers break open batteries with hand tools and remove the lead plates. These are smelted in open-air pits that spread lead-laden fumes and particulate matter over wide swaths of surrounding neighborhoods (World Health Organization, 2017). Lead-laced acid from the batteries is often drained onto the bare ground or dumped directly into waterways (World Health Organization, 2017). Documented soil lead contamination ranges from 2,500 mg/kg (Daniell et al., 2015) to 302,000 mg/kg (Haefliger et al., 2009) in residential areas in some LMICs. We focus on India as a significant portion of worldwide ULAB recycling occurs in India, and, despite substantial reductions in exposure from lead-paint, ULAB recycling continues to contribute to elevated blood lead levels among the Indian population (Ericson et al., 2018; Belliniger et al., 2005; Chatham-Stephens et al., 2013; Ericson et al., 2013, 2018; Sharma et al., 2005). In this paper, we quantify the environmental externalities of ULAB recycling related to the under examined relationship of lead and livestock in India.

Livestock production is estimated to contribute 4% of India's GDP and as much as 70% of rural employment (Roy and Singh, 2013). Additionally, poorer households in India depend more on livestock than richer households. For example, farmers holding less than 0.01 hectares of land earned 26% of their household income from animal husbandry while farmers holding over 10 hectares only earned 6% of household income from animal husbandry (Chakravorty et al., 2019). Furthermore, for non-migrating families, livestock ownership was an important source of livelihood diversification for households in the lower half of the income distribution (Deshingkar et al., 2020).

Once emitted, lead particles are highly immobile and persistent in soil, tending to remain near the surface for prolonged periods. This poses a continuous potential risk to grazing cattle, which ingest from 1% to nearly 18% of their dry matter intake as soil (Thornton and Abrahams, 1983). Accordingly, the soil intake pathway represents a significant source of bovine lead exposure (Mayland et al., 1975; Alloway, 2012; McDowell, 2003; Sharpe, 2004). General symptoms of BM lead poisoning include blindness, ataxia, cramps, muscle tremors, convulsions, aggression, teeth grinding, anorexia, diarrhea, constipation, and respiratory failure, among others (Blakley, 1984; Zmudski et al., 1983; Bates and Payne, 2017). One of the earliest studies found bovine death occurred from chronic low-level exposure within 20 days (Hammond and Aronson, 1964), and more recent observations suggest that relatively low soil lead concentrations can result in severely adverse health outcomes in bovines (Abrahams and Thornton, 1994; Aslani et al., 2014; Cowan and Blakley, 1998; Ikenaka et al., 2012; Krametter-Froetscher et al., 2007; Thornton and Abrahams, 1983; Zadnik, 2010). Unfortunately, treatment may be ineffective due to the rapid progression of nervous system disease. Thus, euthanization is often the most practical intervention (Cowan and Blakley, 1998). These effects are consistent across breeds and sexes (Cowan and Blakley, 1998).

Despite toxicological evidence, documented cases, and the potential for harmful human livelihood impacts, few studies have attempted to quantify mortality and associated economic costs of bovine lead exposure in India. To fill this gap, we first model potential bovine lead exposures from soil using FAO data on animal density in combination with soil pollution mapping conducted by the NGO PureEarth. Second, we identify the dose response relationship between lead exposure and bovine death to obtain estimated moralities attributable to lead poisoning. Finally, we estimate the net present value of economic damages from these bovine mortalities.

2 Methods

2.1 Attributing livestock densities at ULAB sites

To determine the number of livestock grazing on lead contaminated land in India, we overlaid the FAO gridded livestock data (Robinson et al., 2014) with soil lead concentrations collected by the PureEarth. As part of their Toxic Sites Identification Program (TSIP), a USAID funded project, (PureEarth, 2020), PureEarth recorded geo-located soil lead concentrations surrounding 222 informal used lead acid battery (ULAB) recycling sites in India. PureEarth staff utilized in-situ X-Ray Fluorescence (XRF) spectrometry to quantity surface soil lead concentrations at these sites. The FAO gridded livestock data are available at a spatial resolution of 3 minutes of arc (about 5×5 km at the equator). This data is initially based on nationally reported livestock statistics and observed livestock densities, then expanded with statistical modeling and adjusted according to corroborating datasets from FAOSTAT and elsewhere (Robinson et al., 2014). The bovine densities for India are overlaid with the TSIP ULAB lead contamination sites (Fig. 1). Additional information on GIS methods used are available in the appendices.

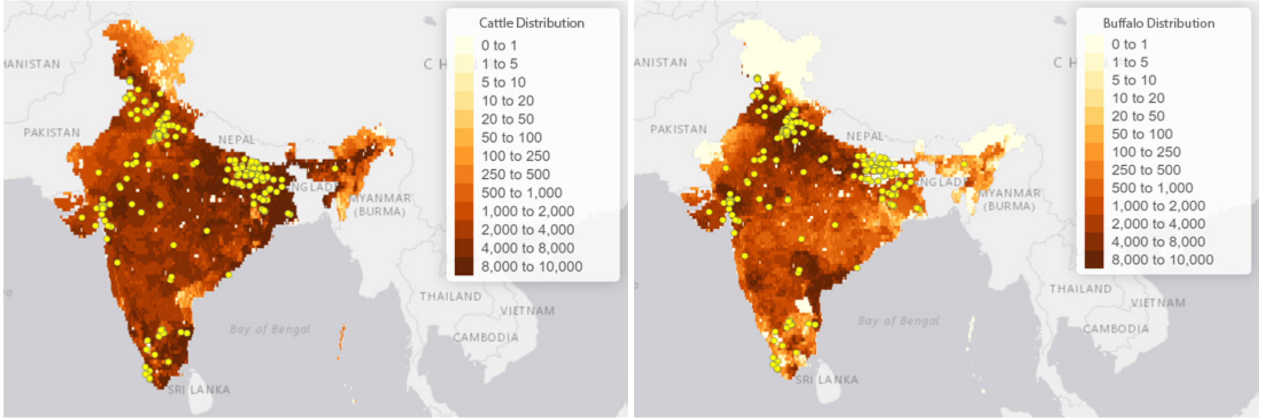

FIGURE 1: FAO gridded bovine densities for India overlaid with the 222 lead contaminated sites used in the analysis

2.2 Modeling soil lead levels and area of exposure

The PureEarth TSIP database does not necessarily map a comprehensive characterization of soil contamination at each of the 222 sites. To interpolate surface soil lead levels at each site, we modeled the likely spatial attenuation of soil lead concentrations (understanding how lead pollution levels decay as one moves away from the centroid of the pollution site). We examined whether wind influenced the direction of migration of lead particles, but found that it did not (Appendix B).

2.2.1 Spatial attenuation

Next, we examined how lead pollution levels decay as one moves away from the source point of the pollution site. To do this, we used PureEarth's global soil lead contamination database which includes 4587 soil samples taken at 690 ULAB sites around the world. We call the soil sample at each site with the highest lead concentrations the "source" point for pollution at that local site. Ultimately, we wanted to map pollution levels in 10 m concentric circles radiating out from the source point up to a radius of 500 m away.

The "portion of source point lead value" was defined as:

$$r_d = \frac{S_a}{S_o} \tag{1}$$

where S_o is the source soil point lead value ($\frac{mg}{kg}$), S_a is the sample soil point lead value ($\frac{mg}{kg}$) and r_d is the portion of the S_o lead value. Note that r_d is indexed by the distance S_a is away from S_o. For example, if the portion of source point lead value at 50 m from the source point was 0.25, then the soil lead levels were estimated to have fallen to 25% of the source point's value at 50 m away from the source. Figure 2 provides a graphic depiction of this process.

We ran a random forest predictive model and two ordinary least squares (OLS) regressions to predict equation 1 from a training sample (data summaries and model comparisons can be found in the appendices). We used the model with the most predictive capacity to estimate the share of source point lead values at 10 m concentric intervals radiating out from the source point at each of India's 222 ULAB sites. We used the outputs of this prediction process as inputs for the BM exposure modeling. However, we recognize that these predictions carry their own error margins and so include a detailed description in the appendix (appendix D). These bounds include a 90%-confidence interval for the random forest model using an out-of-bag interval method described by Zhang et al. (2019).

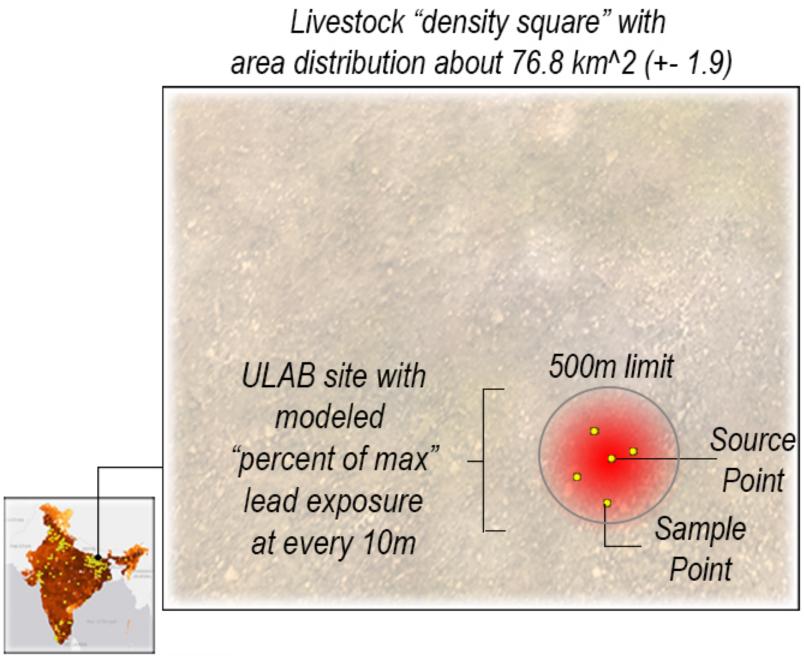

FIGURE 2: Overview of spatial attenuation modeling process. Having generated a random forest predictive model for soil lead contamination from all globally known 690 ULAB sites (4587 total samples), we predicted each of India's 222 sites' soil lead contamination up to 500 m from the source point at 10 m intervals. Given that FAO's gridded livestock population density is statistically modeled as uniform (Robinson et al., 2014), we set the probability of exposure equal to the percent of the "density square" covered by contaminated lead. Actual exposure dosage is represented by equation 2.

2.3 Calculating lead dose

We calculated the amount of lead ingested each day by ruminants based on the mean soil lead concentrations in each of the 10 m concentric circles described in the spatial attenuation section and several assumptions about bovine body mass and diet. These inputs were used in the following calculation based on Johnsen and Aaneby (2019):

$$D = \frac{S * F * S_i}{B_w} * G \tag{2}$$

where D is the lead dose per day ($\frac{mg}{kg}$ body weight per day), S is the soil lead concentration ($\frac{mg}{kg}$), F is the amount of fodder ingested per day (kg of dry weight), S_i is the daily soil ingestion rate, B_w is the body weight (kg), and G is the duration of exposure.

We predict lead concentration values S using random forest predictive modeling for concentric circles emanating out from the ULAB site's source point of contamination. Bovine mammals ingest fodder (F) in relationship to their body weight (B_w). A bovine mammal ingests approximately 3% of its bodyweight in fodder each day (Birthal and Dikshit, 2010; Department of Primary Industries and Regional Development, 2020).

BM soil intake (S_i) ranges between 1% and 18% of total dry matter intake (Thornton and Abrahams, 1983). Because India features heavy monsoon rains and more sparse grazing conditions that have been shown to increase soil intake, we use a conservative value of 8% for S_i (Thornton and Abrahams, 1983). The 8% is broadly consistent with values noted elsewhere, including Siberia (Mamontova et al., 2007) and the French West Indies (Collas et al., 2019).

The final piece of information we need to calculate the lead dosage per day that BMs face is the duration of exposure (G). This is complicated by unknown BM behavior and location (free ranging, in a pen, or some mixture). To resolve this challenge, we opt for the most conservative approach by assuming animals are penned up and that lead exposure occurs exclusively in their stalls. We multiply soil lead levels by 0.7 (G), which can serve as a soil-to-structure-dust transfer variable taken from the US Environmental Protection Agency's IEUBK model (US Environmental Protection Agency, 1998). Transfer of lead from owners to the holding pens is realistic considering farmers interact with their livestock daily (Birthal and Dikshit, 2010; Sharma et al., 2019) and indoor lead dust from ULAB recycling has been illustrated as a major source of exposure in South Asian countries (Chatham-Stephens et al., 2013). While this is likely an underestimate (because in practice many BMs are not exclusively penned up), it resolves the challenge of modeling herding behavior while providing a conservative estimate.

2.4 Estimating the number of lethally exposed BMs

The number of lethally exposed BMs is represented by Equation 3. If contaminated lead soil provides a daily dose above the fatal threshold values, 6 mg/kg and 5 mg/kg for adults and calves respectively (Zmudski et al., 1983), then it is considered an area of fatal exposure (A_E). These threshold values were also used by Johnsen and Aaneby (2019) in their ruminant soil lead exposure assessment.

$$BMFatalities_i = \frac{A_{Ei}(S_o, r_{di})}{A_{Ti}} * D_{ti} \tag{3}$$

The source point of pollution at each site is based on that site's highest recorded soil sample (S_o) and the site-specific portion of source point lead value (r_d) was calculated via Equation 1. The density square area (A_{Ti}) and the density of BMs in the density square (D_{Ti}) are provided by the FAO's gridded livestock dataset (Robinson et al., 2014). We estimated the number of BMs with fatal exposure as the number of BMs in the density area containing leaded soil above the threshold values. The percentage of adult and calf BMs, of the total BM population, were based on the most recent Indian livestock census (Department of Animal Husbandry and Dairyings, 2012).

2.5 Estimating the monetary value of BMs

We estimated the net present value (NPV) for current and future BM income streams, where the income streams are based on Khan and Iqubal's study of rural farmers' household economies (Khan and Iqubal, 2010). We used an exchange rate of 70.49 Rupees to 1 USD (World Bank, 2020). We used a 5% discount rate, which has previously been used for cattle cost-benefit analysis in India (Singh et al., 2018).

2.5.1 Value of BMs in the household economy

In order to calculate the value of BM mortalities we incorporated the total time horizon of all income streams, yearly costs, yearly benefits, and a discount rate (5%). The productive lifespan of a typical Indian BM is 20 years and they produce on average 4 calves (Dhillon, 2018; Sharma et al., 2019). We adjusted household cattle and buffalo values from Khan and Iqubal (2010) from 2008 values to 2019 values using a 2.24 consumer price index change (CPI) (World Bank, 2020). In the NPV calculation, the first year of ownership includes the cost of purchase and the net income from inputs and outputs. Subsequent years included only the value of inputs and outputs. We later refer to the household value of a BM as V_{HH} (value to the household).

In our modeling, female bovines produce offspring only after age 4 and then every other year according to the interval estimates of offspring production in Sahiwal cows (Ahmad and Sivarajasingam, 1998) up to a maximum of 4 calves. Concerning calves, we assumed that they were produced from adults and not purchased (meaning no purchase price was included in the NPV). Also, calves were assumed to only incur input costs (no output values) in their first year of existence. Because our BM fatality estimates assume that calves perish within their first year of existence, a perished calf always incurs the loss of its entire potential lifespan. However, because BM fatality estimations were designated between calves ($<= 1$ year old) and adults, and an adult could be any age when it came into contact with a fatal lead dose, we use 10-years (the mid-point of the length of the productive adult lifespan) to account for the unknown age at which a BM was fatally exposed to lead. We assume BMs stop being productive at the start of their 21st year. As buffalo are not usually used for meat at the household level (Naveena and Kiran, 2014) and cattle are generally not slaughtered (Harris, 1992), the value of their slaughter products were not included in income estimates.

To capture productivity gains from draft power, we adapted Okello et al.'s study (2015) from Uganda. They find the time requirement for plowing 1 acre using a pair of draft cattle was 2.2 days, while farmers required 12 days to plow 1 acre using a hand held hoe. Cattle can begin plowing at 2 years of age and continue until 11 years of age. A single draft animal provided a 272% increase in time efficiency (Okello et al., 2015). We assume similar productivity gains for using draft animals in India. The average land holdings from 2001 to 2016 from India's Agriculture Census was 2.94 acres (1.19 ha) (CEIC, 2016). The average annual income for agricultural farmers of all land holdings from non-livestock activity was 70,428 Rs in 2013 (the closest census report) (Chakravorty et al., 2019). Using this figure suggests an opportunity cost from non-livestock daily income of 193 Rs per day. This estimate is similar in magnitude to the International Labor Organization's findings that average rural wages were 175 Rs per day in 2011-2012 in India (International Labor Organization, 2018). We use the amount of plow time per acre, 2.2, and take the difference from the hand plow time per acre ($12 - 2.2 = 9.8$). Then a pair of draft animals save 28.8 days per year (9.8 days saved per acre * 2.94 acres), this equates to an average of 5,560 Rs in productivity gains per year (193 Rs * 28.8 days) from a pair of bovine during their drafting years (9 years). Divided by two, attributing these gains across both animals, is 2,780 Rs, or \$ 2.39 per day (40.3 Rs-USD 2010 exchange rate). Khan and Iqubal (2010) found 45% of Indian farmers plowed land with draft animals, therefore we add 1,251 Rs (2,780 Rs * 0.45) to our NPV calculations for the value of draft animals.

2.6 Estimating damages

Damage incurred from exposure at any given site were calculated as follows:

$$Damage_i = BMFatalities_i * V_{HH} \tag{4}$$

where i is the site index, $BMFatalities_i$ is the number of BM fatalities, and V_{HH} is the net present value of one BM.

3 Results

3.1 Random forest model output and BM exposure estimates

The random forest method was more accurate than the OLS models (random forest: $R^2 = 0.75$ and SSE $= 34.12$). More details on model comparison can be found in the appendices. The random forest model provided outputs for soil lead predictions at 10 m intervals for each of the 222 ULAB sites in India. The prediction distribution is provided in Figure 3.

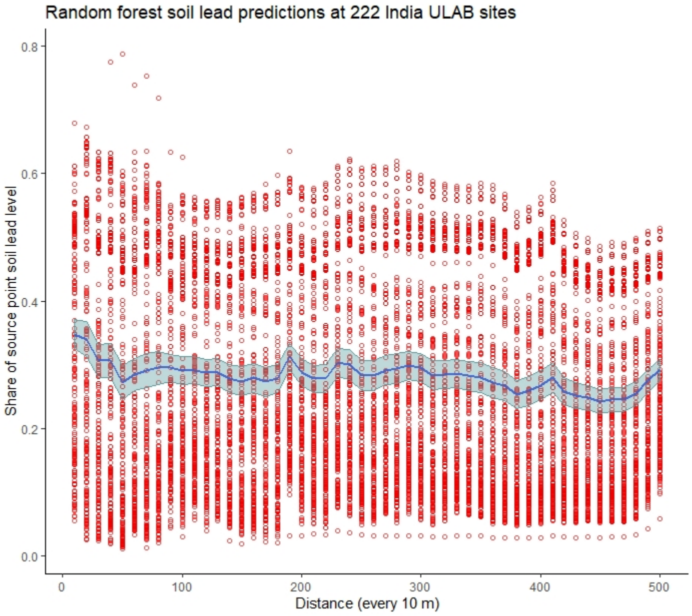

Figure 3: Share of source point soil-lead estimates, with 95% CI bounds. We described each known soil lead sample by its distance from the site center and by its share of the site's highest recorded soil lead value (the "source point," which was taken to be the site's center). Then we used random forest modeling to predict the decay rates above. That is, each point represents the predicted share of the source point's lead value by 10 m intervals for each site. We added a trend line to visualize the trend.

The results from the random forest exercise indicated that the average share of the maximum lead value remained between 0.2 and 0.4 when moving away from the source point at each site. The average lead concentration was 1809 mg Pb/kg (1734 - 1884 95 % CI) in soil over the 500-meter distance. The exact soil-level estimates, for each 10 m interval at every site, can be found in the supplemental materials (excel file).

3.2 Number of annual BM fatalities and area of exposure

Our subsequent estimates for annual BM fatalities are provided in Table 1. We find an aggregate of 3,273 BM fatalities across the 222 sites, with a per site average of 14.74 (95% CI, [7.28, 22.21]).[1] The total BM population estimated to be found within a 500 m radius of each ULAB site was 22,385, with an average of 101 per site (95% CI [87 - 114]). Therefore, we find the mortality rate for bovine within 500 m of a ULAB site to be 14.62%. The aggregate area of soil contamination contributing to lethal levels of lead exposure was 23.06 km^2, with a per site average of 0.104 km^2 (95% CI, [0.071, 0.137]). The total area of any soil lead contamination above 0 mg/kg was 120.89 km^2, with a per site average of 0.545 km^2 (95% CI, [0.478, 0.592]). The maximum soil lead level contributing to BM fatalities was 62,846 mg/kg, and the minimum contributing to fatalities was 8,992 mg/kg.

BM Type	**Estimated Fatalities**	**Average Per Site**
Adult Cattle	1224.60	5.51
Calf Cattle	427.39	1.93
Adult Buffalo	1,137	5.51
Calf Buffalo	483.64	11.58
Total	3,273.0	14.74

TABLE 1: Estimated total number of lethally exposed BMs from the 222 documented sites (daily intake threshold of above 5 mg/kg per day for calves and above 6 mg/kg for adults).

3.3 Cost of BM fatalities

The calculations for the cost of BM fatalities are detailed in Table 2. The total economic cost of lost output due to estimated annual bovine fatalities in these 222 sites is \$22,110,086 (in 2019 USD), with a per site average of \$ 96,566 (95%CI, [\$ 46,149, \$ 147,005]). Without any effective lead mitigation measures, we would expect these contaminated sites to cause this amount of additional damage each year.

---

[1] All per-site confidence intervals were calculated with nonparametric bootstrapping of the estimated per-site mean by resampling observations (with replacement) 5000 times.

Value of BMs to the Household Economy - 2008 Rs per annum converted to 2019 USD

Adult Purchase						
buffalo	56014					
cattle	22406					
Inputs	fodder	feed grain	oil cake	labor	total input	
buffalo	20165	32085	12682	13443	78375	
cattle	17925	32085	11203	13443	74656	
Outputs	milk	dung cake	manures	offspring	draft power	total output
buffalo	118751	30113	4481	5601	2803	163286
cattle	80661	22406	4481	4481	2803	116368
Net income stream	Total	(no offspring)	(no draft power)			
buffalo	84911	79309	80571			
cattle	41712	37231	37373			

Cost of BM fatalities

Adult NPV Total 5 % discount	Rs	USD	BM Quant.	Total Value	Site Average
buffalo adult total NPV	1,029,516 Rs	$14,620			
buffalo adult life average	514,758 Rs	$ 7,310	1137	$ 8,311,275	$ 37,438
cattle adult total NPV	493,610 Rs	$ 7,010			
cattle adult life average	246,805 Rs	$ 3,505	1225	$ 4,293,259	$ 19,339
buffalo calf total NPV	891,782 Rs	$ 12,664	484	$ 6,129,259	$ 27,609
cattle calf total NPV	445,891 Rs	$ 6,332	427	$ 2,703,713	$ 12,178
			Total	**$ 21,437,575**	**$ 96,566**

TABLE 2: Estimated cost (discounted net present value in 2019 USD) of BM fatalities to the household economy (based on Khan and Iqubal (2010)).

3.4 Distribution of BM fatalities

In addition to the aggregate annual deaths and economic damages presented above, there is another dimension of the burden of bovine death worth highlighting. The distribution of deaths is highly skewed, meaning that sites that have both a high density of livestock and with high soil lead concentrations are where the majority of deaths are expected to occur. According to Figure 4, approximately 80% of the 222 sites register zero estimated deaths, while the remaining 20% of sites account for all the estimated deaths. This distribution of damages implies important considerations should be taken when deciding which sites to prioritize for possible mitigation actions.

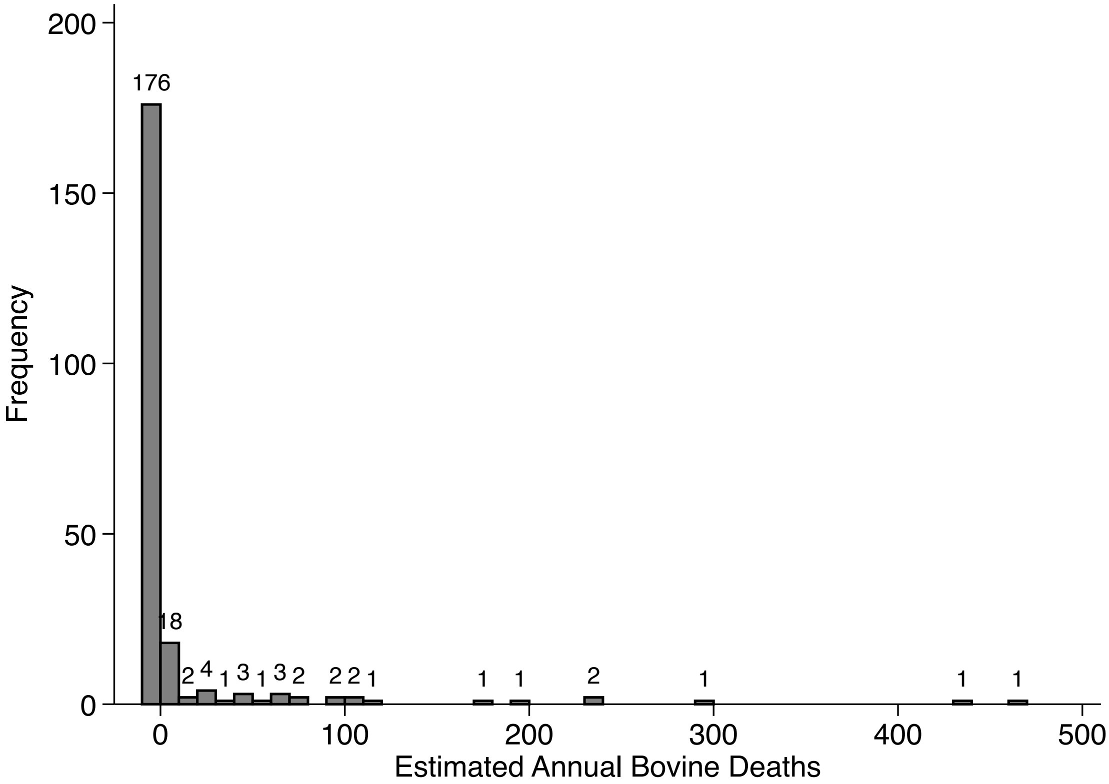

FIGURE 4: Histogram of the estimated annual bovine deaths across the 222 TSIP sites. All 176 sites with zero estimated annual deaths are placed in the left most bar. The number above each bar represents how many TSIP sites fall within that bin of the histogram.

4 Discussion

At approximately 15 BM fatalities (costing \$ 96,566) per site per year, for a total of \$ 21,437,575, our estimated economic damages and loss of animal life further justify proactive measures to reduce lead dust exposure. Previous literature suggests our estimates are both realistic and conservative. In their study of ruminants grazing lead-contaminated soil, Johnsen et al. (2019) use lower parameters for soil uptake and on-site grazing time yet find no risks for cattle or sheep grazing on soil contaminated with up to 3,700 mg Pb/kg. They report that 3,700 mg Pb/kg amount is above the Norwegian Veterinary Institute's suggested soil lead level for ruminant grazing (300 mg Pb/kg). Our estimates suggest zero BM fatalities up to 8,992 mg Pb/kg. Previously documented cases of livestock lead exposure in India further suggest similarities in the number of bovine fatalities, even at lower soil lead levels. Lemos et al. (2004) investigated lead exposure in a herd of 120 Nelore cows from a lead battery recycling plant in Brazil. Thirty-five cattle died within 45 days with clinical signs of cortical neurological disturbances. Soil lead concentrations in the pasture area were 147–431 mg/kg (Lemos et al., 2004). In the Thane district of India, over 300 cattle were found to have died from environmental exposure to lead, cadmium, and chromium attributable to a lead-zinc smelter in 1991 (Dogra et al., 1996).

Studies of Indian livestock morbidity and mortality in Karnal (Prasad et al., 2004), Maharashtra (Bangar et al., 2013), Himachal Pradesh (Chaudhary et al., 2013), and Haryana (Pal et al., 2018) have suggested all-cause bovine mortality rates of 14.17%, 4.42%, 9.14%, 2.56%, respectively. Our model suggests that lead poisoning provides an additional bovine mortality burden of 14.62 % to farmers living within 500 m of a ULAB site (0.785 km^2 or 78.5 ha). Given an 1.2 ha average farm land holding in India for 2010 (Hazell, 2015), the resulting bovine mortality burden would extend across approximately 65 farmers per site. Because a ULAB recycling site implies a higher bovine mortality rate, perhaps up to 28 %, without likely providing any economic benefit from the ULAB site's activities, it is possible that these farmers will subsequently not receive adequate utility from bovine husbandry to continue the practice. This is an area for further investigation. Note, our estimates only assume death if a bovine mammal encounters a fatal lead dose in a given year. We do not account for gradual lead exposure over multiple years at smaller doses that eventually reach a fatal threshold of total cumulative exposure. Therefore, the 14.62 % mortality rate should be considered the rate from more immediate lead poisoning (i.e., BMs that ingest enough lead to cause death in the same year).

Implementation of protective measures is ultimately left to policy makers and environmental specialists. However, we would like to emphasize two further points for consideration.

First, the TSIP database is largely sampled focusing on sites falling in the poorest regions of India (like Bihar). Thus, the $ 96,566 of economic damages estimated at each ULAB site, even spread among many small farmers, would be a tremendous burden to a rural agriculturalist. Second, it is likely that the 222 known sites do not even constitute a large share of ULAB sites in India. Chatham-Stephens et al. (2013) suggest that the TSIP database only captures around 10% of ULAB sites in South Asia, so it is quite possible that our mortality estimate only captures about one tenth of the true value of BM mortality due to lead exposure in India.

Unfortunately, there appears no easy policy solution for ULAB lead exposure on livestock. At any given ULAB site, eliminating soil-lead exposure requires active remediation by engineers. Lead's persistence in soil implies that shutting down ULAB sites will not resolve the soil-lead exposure problem. Additionally, because of low barriers to entry and the low-level of capital necessary to smelt lead, new sites can open relatively quickly. Closing current ULAB sites may promote their re-opening elsewhere and an increase in the total area of exposed soil. As the area of exposure increases, the area on which farmers could forage for fodder or graze livestock shrinks. Policy makers might consider designing incentives to register ULAB sites and protective regulations to contain site exposure areas. Yet, if the affected farmers have little political influence and the overall contribution to the total BM mortality rate is perceived as low, policy makers may not be driven to act at all.

This study has several limitations. First, it only models the costs related to animal mortality, but not morbidity. Modeling based on lethal daily dose largely precludes measuring acute exposures or non-lethal negative health outcomes (lost milk productivity, etc.). Cowan and Blakley (1998) found euthanization was the most effective option for lead-poisoned cattle in Canada given the recovery rates, product contamination, and medical costs. This suggests that the non-fatality related health outcomes are likely large and important. Second, because the FAO livestock density maps are not available past 2010, it is difficult to make year-to-year estimates up to the present date. Therefore, we suggest caution before using our estimates for forecasting. If the geographic distribution of livestock densities have not changed drastically in India since 2010, this is less of an issue, if however, there has been much change, this is an important limitation.

There remain many avenues for future study. First, similar concerns for lead exposure in other ruminants (sheep, goat, etc) have been documented. Expanding the study to include estimates of the number of fatalities for these species would be useful, especially as the poorest farmers are more likely to own sheep or goats rather than cattle or buffalos. Second, as humans consume livestock and livestock products, there is reason to investigate livestock products as a potential lead exposure pathway. To the degree to which these livestock products (milk, etc.) are consumed locally represents an additional (and unequal) burden for the rural poor due to the externalities of recycling lead acid batteries. Third, a broader economic analysis of ULAB recycling's market size would indicate the total value of externalities per battery produced. This could help policy makers determine appropriate taxes, permits, compensations, or other pollution reduction strategies. Fourth, because 20% of the sites cause the vast majority of fatalities, geographic targeting of mitigation activities is necessary. Modeling exercises like the one performed in this study could rank sites by expected mortality in order to prioritize mitigation investments so they could focus on the sites with the largest externalities. Finally, in the process of providing BM exposure estimates, we produced an empirically derived model of ULAB soil lead contamination that could serve as a framework for modeling other ULAB pollutants and damages.

5 Conflict of Interest

None

6 Financial Disclosure

Financial support was provided by the United States Agency for International Development (USAID) for the Toxic Sites Identification Program (TSIP). The opinions expressed herein are those of the authors and do not necessarily reflect the views of the study sponsors. The funders had no role in study design, data collection, analysis or data interpretation. We thank Fordham University's Graduate School of Arts and Sciences for student travel funding to visit associated used lead acid battery recycling sites in India.

Bibliography

Abrahams, P. and I. Thornton (1994). The contamination of agricultural land in the metalliferous province of southwest England: implications to livestock. *Agriculture, ecosystems & environment 48*(2), 125–137.

Ahmad, M. and S. Sivarajasingam (1998). ANALYSIS ON THE PRODUCTIVE AND REPRODUCTIVE TRAITS IN SAHIWAL. pp. 4.

Alloway, B. J. (2012). *Heavy metals in soils: trace metals and metalloids in soils and their bioavailability*, Volume 22. Springer Science & Business Media.

Aslani, M. R., M. Heidarpour, V. Najarnezhad, and M. Mostafavi (2014). Lead poisoning in cattle associated with batteries recycling: High lead levels in milk of nonsymptomatic exposed cattle. *4*(1), 6.

Bangar, Y., T. Khan, A. Dohare, D. Kolekar, N. Wakchaure, and B. Singh (2013). Analysis of morbidity and mortality rate in cattle in village areas of Pune division in the Maharashtra state. *Veterinary World 6*(8), 512.

Bates, N. and J. Payne (2017, July). Lead poisoning in cattle. *Livestock 22*(4), 192–197. Publisher: Mark Allen Group.

Belliniger, D. C., H. Hu, K. Kalaniti, N. Thomas, P. Rajan, S. Sambandam, P. Ramaswamy, and K. Balakrishnan (2005). A pilot study of blood lead levels and neurobehavioral function in children living in Chennai, India. *International journal of occupational and environmental health 11*(2), 138–143.

Birthal, P. and A. Dikshit (2010, January). India's Livestock Feed Demand: Estimates and Projections. *Agricultural Economics Research Review 23*.

Blakley, B. (1984, December). A retrospective study of lead poisoning in cattle. *Veterinary and human toxicology 26*(6), 505–507.

CEIC (2016). India Agriculture Census: Average Size of Operational Land Holdings: Size Group: All Holdings. Technical report.

Chakravorty, S., S. Chandrasekhar, and K. Naraparaju (2019, November). Land Distribution, Income Generation and Inequality in India's Agricultural Sector. *Review of Income and Wealth 65*(S1).

Chatham-Stephens, K., J. Caravanos, B. Ericson, J. Sunga-Amparo, B. Susilorini, P. Sharma, P. J. Landrigan, and R. Fuller (2013, July). Burden of Disease from Toxic Waste Sites in India, Indonesia, and the Philippines in 2010. *Environmental Health Perspectives 121*(7), 791–796.

Chaudhary, J., B. Singh, S. Prasad, and M. Verma (2013). Analysis of morbidity and mortality rates in bovine in Himachal Pradesh. *Veterinary World 6*(9), 614.

Collas, C., M. Mahieu, A. Tricheur, N. Crini, P.-M. Badot, H. Archimède, G. Rychen, C. Feidt, and S. Jurjanz (2019, June). Cattle exposure to chlordecone through soil intake. The case-study of tropical grazing practices in the French West Indies. *Science of The Total Environment 668*, 161–170.

Cowan, V. and B. Blakley (1998). Acute lead poisoning in western Canadian cattle — A 16-year retrospective study of diagnostic case records. *57*, 6.

Daniell, W. E., L. Van Tung, R. M. Wallace, D. J. Havens, C. J. Karr, N. Bich Diep, G. A. Croteau, N. J. Beaudet, and N. Duy Bao (2015). Childhood Lead Exposure from Battery Recycling in Vietnam. *BioMed Research International 2015*, 1–10.

Department of Animal Husbandry and Dairyings (2012). 19th Livestock Census 2012 All India Report. Technical report.

Department of Primary Industries and Regional Development (2020). Preventing residues when using stockfeed from organochlorine (OC) contaminated land. Agriculture, Government of Western Australia.

Deshingkar, P., J. Farrington, L. Rao, S. Akter, P. Sharma, A. Freeman, and J. Reddy (2020). Livestock and poverty reduction in India: Findings from the ODI Livelihood Options Project.

Dhillon, A. (2018). 'I ward them off with a cricket bat': Indian farmers despair over marauding cows. *The Guardian*.

Dogra, R., R. Murthy, A. Srivastava, J. Gaur, L. Shukla, and B. Varmani (1996). Cattle mortality in the Thane district, India: a study of cause/effect relationships. *Archives of environmental contamination and toxicology 30*(2), 292–297.

Ericson, B., J. Caravanos, K. Chatham-Stephens, P. Landrigan, and R. Fuller (2013, February). Approaches to systematic assessment of environmental exposures posed at hazardous waste sites in the developing world: the Toxic Sites Identification Program. *Environmental Monitoring and Assessment 185*(2), 1755–1766.

Ericson, B., R. Dowling, S. Dey, J. Caravanos, N. Mishra, S. Fisher, M. Ramirez, P. Sharma, A. McCartor, P. Guin, M. P. Taylor, and R. Fuller (2018, December). A meta-analysis of blood lead levels in India and the attributable burden of disease. *Environment International 121*, 461–470.

Ericson, B., P. Landrigan, M. P. Taylor, J. Frostad, and J. Caravanos (2016, March). The Global Burden of Lead Toxicity Attributable to Informal Used Lead-Acid Battery Sites. *Annals of Global Health 82*(5), 686.

Haefliger, P., M. Mathieu-Nolf, S. Lociciro, C. Ndiaye, M. Coly, A. Diouf, A. L. Faye, A. Sow, J. Tempowski, J. Pronczuk, A. P. F. Junior, R. Bertollini, and M. Neira (2009, October). Mass Lead Intoxication from Informal Used Lead-Acid Battery Recycling in Dakar, Senegal. *Environmental Health Perspectives 117*(10), 1535–1540.

Hammond, P. B. and A. L. Aronson (1964, April). LEAD POISONING IN CATTLE AND HORSES IN THE VICINITY OF A SMELTER*†. *Annals of the New York Academy of Sciences 111*(2), 595–611.

Harris, M. (1992, February). The Cultural Ecology of India's Sacred Cattle. *Current Anthropology 33*(S1), 261–276.

Hazell (2015). Is Small Farm-Led Development Still a Relevant Strategy for Africa and Asia?

The Fight Against Hunger and Malnutrition.

Hijmans, R. J. (2019). *raster: Geographic Data Analysis and Modeling.* R package version 3.0-7.

Ikenaka, Y., S. M. Nakayama, T. Muroya, J. Yabe, S. Konnai, W. S. Darwish, K. Muzandu, K. Choongo, G. Mainda, H. Teraoka, T. Umemura, and M. Ishizuka (2012, October). Effects of environmental lead contamination on cattle in a lead/zinc mining area: Changes in cattle immune systems on exposure to lead in vivo and in vitro. *Environmental Toxicology and Chemistry 31*(10), 2300–2305.

International Labor Organization (2018). India Wage Report: Wage policies for decent work and inclusive growth. Technical report, ILO.

International Lead Association (2014). Lead Recycling: Sustainability in Action. Technical report.

Johnsen, I. V. and J. Aaneby (2019, October). Soil intake in ruminants grazing on heavy-metal contaminated shooting ranges. *Science of The Total Environment 687*, 41–49.

Johnsen, I. V., E. Mariussen, and Voie (2019). Assessment of intake of copper and lead by sheep grazing on a shooting range for small arms: a case study. *Environmental Science and Pollution Research 26*(8), 7337–7346.

Khan, N. and A. Iqubal (2010). LIVESTOCK REVOLUTION AND ITS IMPACTS ON THE SUSTAINABILITY OF MARGINAL AND SMALL FARMERS IN INDIA: A CASE STUDY. pp. 14.

Krametter-Froetscher, R., F. Tataruch, S. Hauser, M. Leschnik, A. Url, and W. Baumgartner (2007). Toxic effects seen in a herd of beef cattle following exposure to ash residues contaminated by lead and mercury. *The Veterinary Journal 174*(1), 99–105.

Lemos, R., D. Driemeier, E. B. Guimarães, I. S. Dutra, A. E. Mori, and C. Barros (2004). Lead poisoning in cattle grazing pasture contaminated by industrial waste. *Veterinary and human toxicology 46*(6), 326–328.

Liaw, A. and M. Wiener (2002). Classification and regression by randomforest. *R News 2*(3), 18–22.

Mamontova, E., E. Tarasova, A. Mamontov, M. Kuzmin, M. McLachlan, and M. Khomutova (2007, April). The influence of soil contamination on the concentrations of PCBs in milk in Siberia. *Chemosphere 67*(9), S71–S78.

Mayland, H. F., A. R. Florence, R. C. Rosenau, V. A. Lazar, and H. A. Turner (1975, November). Soil Ingestion by Cattle on Semiarid Range as Reflected by Titanium Analysis of Feces. *Journal of Range Management 28*(6), 448.

McDowell, L. R. (2003). *Minerals in animal and human nutrition.* Amsterdam: Elsevier Science B.V. Issue: Ed.2.

Naveena, B. and M. Kiran (2014, October). Buffalo meat quality, composition, and processing characteristics: Contribution to the global economy and nutritional security. *Animal Frontiers 4*(4), 18–24.

Okello, W. O., D. Muhanguzi, E. T. MacLeod, S. C. Welburn, C. Waiswa, and A. P. Shaw (2015, December). Contribution of draft cattle to rural livelihoods in a district of south-eastern Uganda endemic for bovine parasitic diseases: an economic evaluation. *Parasites & Vectors 8*(1), 571.

Organisation Internationale des Constructeurs d'Automobiles (OICA) (2016). Vehicles in use. Technical report.

Pal, R., S. Kumar, D. Bardhan, J. Goyal, A. G, C. M, and S. Gulab (2018). Quantifying morbidity and mortality losses in bovine due to Haemorrhagic septicaemia (hs) in Haryana. *The Pharma Innovation 7(7)*, 802 – 806.

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal 10*(1), 439–446.

Prasad, S., N. Ramachandran, and S. Raju (2004). Mortality Patterns in Dairy Animals under Organized Herd Management Conditions at Karnal India. *Tropical Animal Health* *and Production 36*(7), 645–654.

PureEarth (2020). Toxic Site Identification Program.

Robinson, T. P., G. R. W. Wint, G. Conchedda, T. P. Van Boeckel, V. Ercoli, E. Palamara, G. Cinardi, L. D'Aietti, S. I. Hay, and M. Gilbert (2014, May). Mapping the Global Distribution of Livestock. *PLOS ONE 9*(5), e96084.

Roy, A. K. and J. P. Singh (2013). Grasslands in India: Problems and perspectives for sustaining livestock and rural livelihoods. *Tropical Grasslands - Forrajes Tropicales 1*(2), 240.

Sharma, Kennedy, Schuetze, and Phillips (2019, April). The Welfare of Cows in Indian Shelters. *Animals 9*(4), 172.

Sharma, M., M. Maheshwari, and S. Morisawa (2005). Dietary and inhalation intake of lead and estimation of blood lead levels in adults and children in Kanpur, India. *Risk Analysis: An International Journal 25*(6), 1573–1588.

Sharpe, R. (2004). An overview of lead poisoning in cattle. *Cattle practice 12*, 199–203.

Singh, B. B., P. Kostoulas, J. P. S. Gill, and N. K. Dhand (2018, May). Cost-benefit analysis of intervention policies for prevention and control of brucellosis in India. *PLOS Neglected Tropical Diseases 12*(5), e0006488.

Sumner, M. D. (2019). *spex: Spatial Extent Tools.* R package version 0.6.0.

Tennekes, M. (2018). tmap: Thematic maps in R. *Journal of Statistical Software 84*(6), 1–39.

Thornton, I. and P. Abrahams (1983, January). Soil ingestion — a major pathway of heavy metals into livestock grazing contaminated land. *Science of The Total Environment 28*(1), 287–294.

UNICEF and PureEarth (2020). *The Toxic truth: Children's Exposure to Lead Pollution Undermines a Generation of Future Potential.*

US Environmental Protection Agency (1998). IEUBK Model Mass Fraction of Soil in Indoor

Dust MSD Variable. Technical report.

World Bank (2020). World Bank Open Data.

World Health Organization (2017). Recycling used lead-acid batteries: health considerations. Technical report.

Zadnik, T. (2010). Monitoring of Lead in Topsoil, Forage, Blood, Liver, and Kidneys in Cows in a Lead-Polluted Area in Slovenia (1975–2002) and a Case of Lead Poisoning (1993). *International Journal of Chemical Engineering 2010*, 1–9.

Zhang, H. (2019). *rfinterval: Predictive Inference for Random Forests.* R package version 1.0.0.

Zhang, H., J. Zimmerman, D. Nettleton, and D. J. Nordman (2019, June). Random Forest Prediction Intervals. *The American Statistician*, 1–15.

Zmudski, J., G. Bratton, C. Womac, and L. Rowe (1983). Lead poisoning in cattle: reassessment of the minimum toxic oral dose. *Bulletin of environmental contamination and toxicology 30*(1), 435–441.

APPENDICES

For Online Publication

Lead and Livestock: Estimating India's Bovine Lead Exposure

Greg Ferraro, Bret Ericson, Andrew M. Simons, Emily Nash, Mohammed Kabir

November 17, 2020

Additional GIS method information

We used the coordinate reference system (CRS): "+proj=utm +zone=44 +datum=WGS84 +units=m +no_defs" for all GIS data. The TSIP dataset was provided a CRS with the sf package (Pebesma, 2018). The cattle and buffalo gridded livestock raster shapefile (Robinson et al., 2014) was imported with the raster package (Hijmans, 2019) and was vectorized using the spex (Sumner, 2019) package. The sf package was also used to get the size of each density area (in squared kilometers). The sf package attributed each TSIP ULAB site with its respective cattle or buffalo density. All maps were generated with the tmap package (Tennekes, 2018).

Additional wind direction method information

Because lead exposure from ULAB recycling is at least partially airborne, we felt it prudent to test wind direction on soil lead level distributions. However, we found no statistically significant evidence that wind direction influenced the distribution of lead in the soil. Three sites in the TSIP database permitted radial testing because they had been sampled in a near 360-degree radius around several localized concentrations within each site (Fig. A.1). As a ULAB site would feature multiple lead furnaces or battery breaking areas, we identified 1) localized sources of exposure (called localized source points) based on their elevated soil-lead levels, and 2) sample points, which are the lower lead-level sample points adjacent to source points. Each of these source points were provided a 100m or 150m circle of exposure zone, based on proximity to other source points, and the adjacent points located within these circles were collected into separate data frames using R as before. Thus, we had 8 localized source points between the sites with adjacent sample points to test for radial distributions.

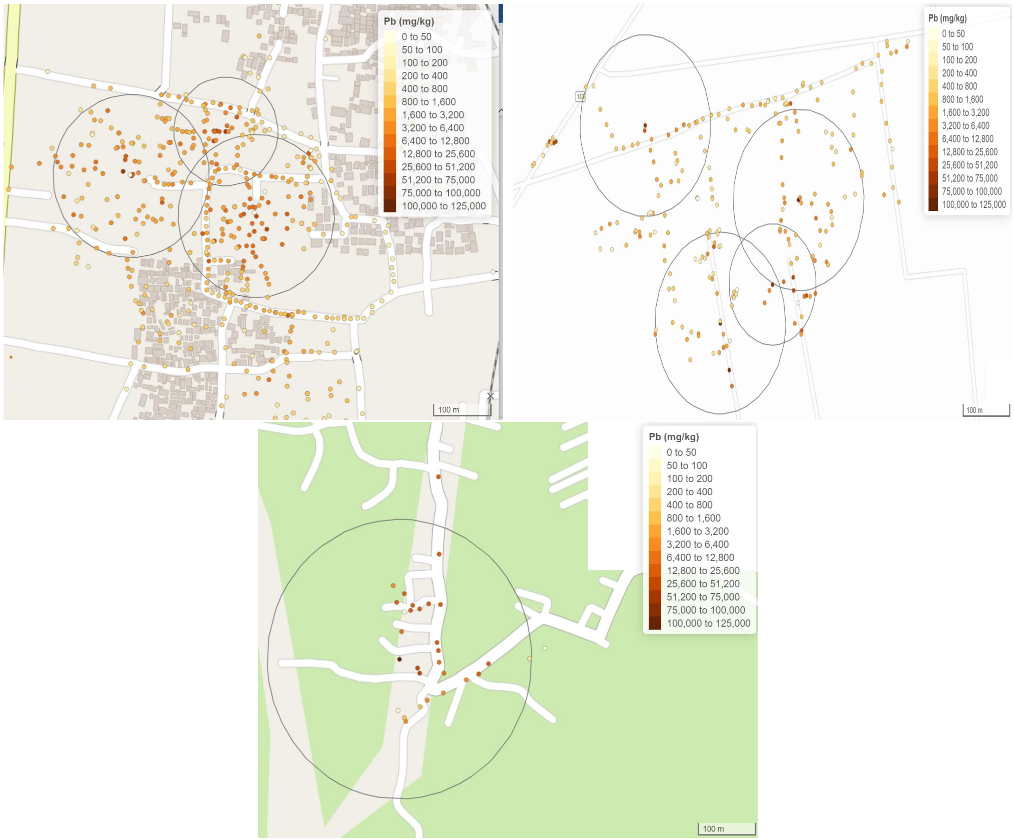

Figure A.1: Surface soil lead concentrations at targeted sites with projected circles of 100m and 150m radii used to determine the relative significance of prevailing wind on lead particle migration.

We measured each sample point's distance from its respective localized source point. Then, using the GPS coordinate of each point and the center point and obtained a slope for each sample point. We converted our results to both degree and radial measures around the localized source point, adjusted according the quadrant location. Figure A.2 is a hypothetical example, where the sample point (located at latitude 22 and longitude 108) is measured radially against the location of the source point at the center of the circle (located at latitude 20 and longitude 106).

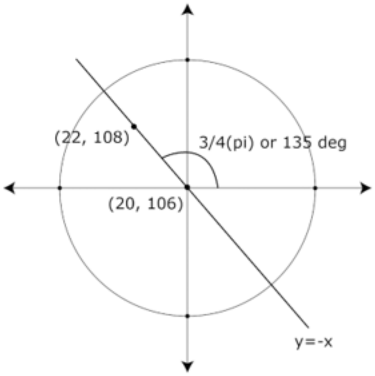

FIGURE A.2: Hypothetical Demonstration of the Radial coordinate assignment process. Center point is the localized source point and the point at (22,108) is the sample point.

We prepared a total of eight different localized source points and assumed that if the distribution soil of lead followed a wind direction, then the mean soil lead level would be greater in one hemisphere at any given source point. To test this, we performed t-tests at each of the eight local source points, among the three sites, to compare the mean soil lead values between hemispheres. As seen in Table A.1, all hemispheres tests were insignificantly different from one another at the 95 % CI level. Based on this, we excluded wind direction for our main analysis.

Wind direction t-test results

Dong Mai

Source Point 63					
North-South Comparison:	t = -0.30,	p = 0.76	CI 95 % = (-10578.42, 7912.02),	Avg north = 5506.6,	Avg south 6839.8
West-East Comparison:	t = -2.30,	p = 0.03	CI 95 % = (-13145.15, -599.14),	Avg west = 1358.57,	Avg east 8230.722
Source Point 224					
North-South Comparison:	t = -0.36,	p = 0.43	CI 95 % = (-2894.89, 2020.61),	Avg north = 1761.09,	Avg south 2198.23
West-East Comparison:	t = 0.83,	p = 0.03	CI 95 % = (-13145.15, -599.14),	Avg west = 1358.57,	Avg east 8230.722
Source Point 128					
North-South Comparison:	t = 0.55,	p = 0.60	CI 95 % = (-1322.18, 2208.02),	Avg north = 2576.27,	Avg south 2133.35
West-East Comparison:	t = -0.81,	p = 0.45	CI 95 % = (-3039.57, 1543.53),	Avg west = 2173.78,	Avg east 2921.80
Source Point 25					
North-South Comparison:	t = -1.09,	p = 0.28	CI 95 % = (-16425.33, 5001.27),	Avg north = 3058.23,	Avg south 8770.26
West-East Comparison:	t = -1.10,	p = 0.06	CI 95 % = (-18671.53, 711.46),	Avg north = 1079.74,	Avg south 10059.77

Tegal

Source Point 156					
North-South Comparison:	t = -1.54,	p = 0.13	CI 95 % = (-2821.83, 352.58),	Avg north = 3721.42,	Avg south 4956.04
West-East Comparison:	t = 0.60,	p = 0.55	CI 95 % = (-1220.96, 2258.83),	Avg north = 4494.89,	Avg south 3975.96
Source Point 410					
North-South Comparison:	t = -1.36,	p = 0.18	CI 95 % = (-7612.29, 1494.83),	Avg north = 4059.27,	Avg south 7118.00
West-East Comparison:	t = 1.00,	p = 0.32	CI 95 % = (-2175.71, 6499.34),	Avg west = 6497.06,	Avg east = 4335.24
Source Point 1 450					
North-South Comparison:	t = 1.98,	p = 0.05	CI 95 % = (-8.00, 5461.26),	Avg north = 5441.20,	Avg south 2714.58
West-East Comparison:	t = -0.04,	p = 0.97	CI 95 % = (-3620.43, 3468.62),	Avg north = 4554.39,	Avg south 4630.30

Cinangka

Source Point 156					
North-South Comparison:	t = 0.51,	p = 0.61	CI 95 % = (-6646.48, 10898.95),	Avg north = 13646.00,	Avg south 11519.77
West-East Comparison:	t = -0.40,	p = 0.75	CI 95 % = (-96813.28, 89364.53),	Avg north = 9145.00,	Avg south 12869.38

TABLE A.1: Results of T-tests comparing the soil lead distributions between North-South and West-East hemispheres around 8 localized source points among 3 different ULAB sites. Only the West-East comparison of source point 63 was found as significantly different, which was not enough to justify including wind direction in the spatial attenuation modeling.

Additional spatial attenuation method information

It was anticipated that the distribution of soil lead levels would remain relatively constant for the first few meters before quickly dropping to significantly lower values and decay to zero at a slower rate. That is, the soil-lead levels would remain high for a short distance from the source before dropping off quickly and flattening towards zero. A scatterplot of soil lead values by distance supported this.

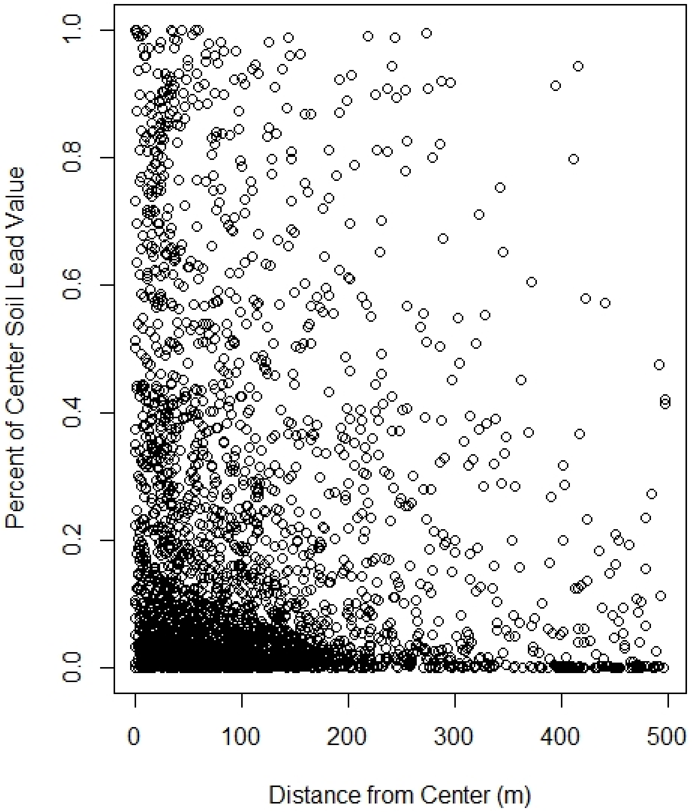

FIGURE A.3: Scatter Plot of soil level by distance for TSIP Site Lead Samples

The first proposed modeling approach, based on the apparent pattern of data points and the expected soil-lead distribution at each site, was to test 3rd- order and 5th-order linear regressions because they would emulate the hypothetical distribution of a rapid decay rate before leveling out. However, a linear model, in this case, would require researchers to assume factors such as site topography, site size, battery part stockpiling patterns, and human spread of contaminated dust were consistent. These are unlikely assumptions, but a well-fit linear regression could support better a priori extrapolation of exposure if necessary. In our second modelling approach, we used a random forest model, which makes value predictions based on information related to each site by averaging decision tree outputs. Importantly, the random forest process could utilize the unique site identification number to account for site heterogeneity in some regard. In all models, the percent of the sample point soil lead value (here called "percent of max") would be the dependent variable. To compare the fit of the three models, the TSIP database of 4587 was randomly assigned to training and testing groups in a $\frac{3}{4}$:$\frac{1}{4}$ ratio respectively. Then the models would be compared against one another through their ability to use the training group's values to predict the testing group's values, as indicated by the R^2 and Sum of Squares Error values:

SSE = (predicted dependent values – testing dependent values)^2

For each of the 4587 soil-lead samples among the global 690 ULAB sites (presented in the scatterplot below) we added the percent of max attribute so that each point had the following attributes:

- site id: Unique site identifier
- test result: sample soil-lead value
- distance: Distance from source point (m)
- max: The source point soil lead value (highest soil-lead value at site)
- percent of max: Ratio of sample soil lead value to max soil lead value from the source point

The linear regression models had the following regression equations and outputs:

FIGURE A.4:
3rd order OLS: Yp = [5.821e-01] – [9.193e-03]*D + [4.595e-05]*D2 – [6.246e-08]*D3
R2 = .32
SSE = 86.40138

FIGURE A.5:
5th order OLS: Yp = [7.460e-01] − [2.481e-02]*D + [2.976e-04]*D2 + [-1.455e-06]*D3 + [3.089e-09]*D4 + [-2.372e-12]*D5
R2 = .4507
SSE = 67.44117

Researchers created a random forest model from the randomforest package (Liaw and Wiener, 2002) with the same training data (using the site id, max, and distance variables as the independent variables). The outputs were as follows:

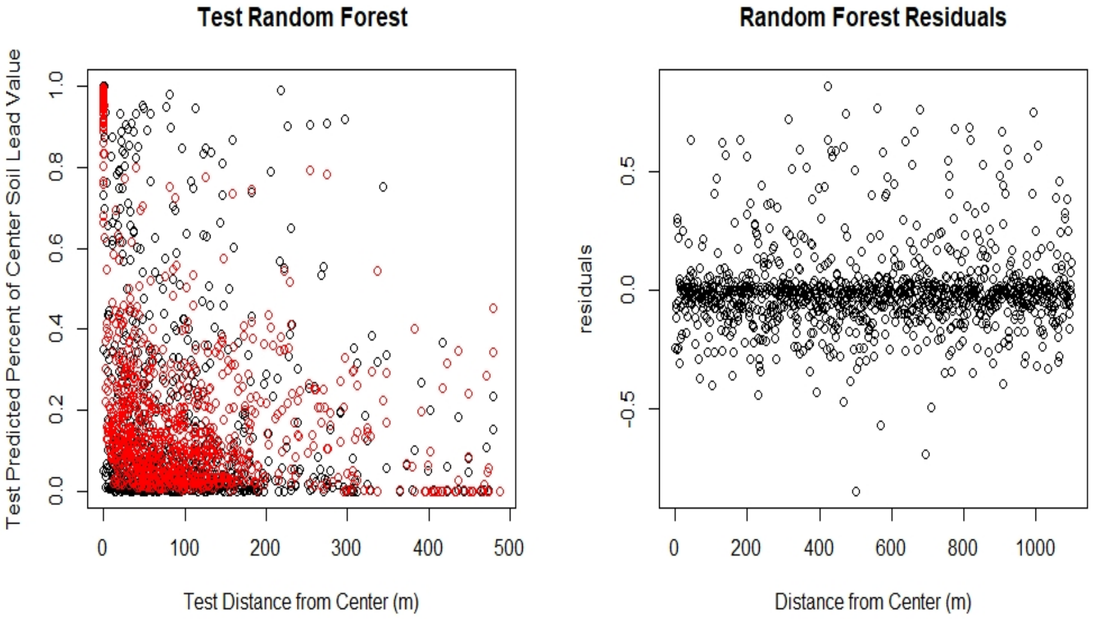

FIGURE A.6:
Random forest output:
SSE = 34.12064
R2 = .75024

Considering the substantially lower SSE value and considerably high R-squared value, the random forest plot was considered both the best model and a quality fit for the data, capable of predicting soil-lead values for any distance under 500m from a given center soil-lead sample point. We then used the rfinterval (Zhang, 2019) package to create upper and lower confidence bounds.

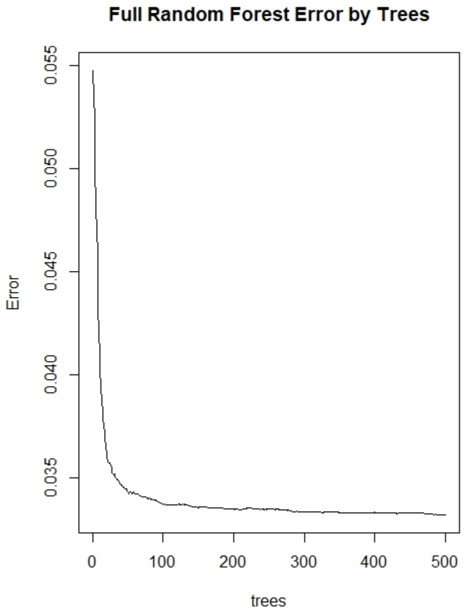

Figure A.7:
Graph of Decision Tree Error Amounts
site id: 57.54326
distance: 356.83788
max: 75.01989

Information on the Random Forest Confidence Interval

The random forest model possessed a distribution of predictions over the course of iterations, which can be described using Zhang et al. (2019)'s out-of-bag confidence interval method. We implemented their method with their rfinterval package in R (Zhang, 2019). A.8 allows for visual comparison of the lower, predicted, and upper estimated share of source point soil lead levels.

The resulting BM fatalities and associated costs can be found in A.2. We did not incorporate these values in our primary analysis because they represent cases in which every one of the 222 ULAB sites have values at or beyond the 90 % CI bounds. These cases would be extremely unlikely to occur, and, therefore, our estimation of the true value is most certainly captured between these bounds.

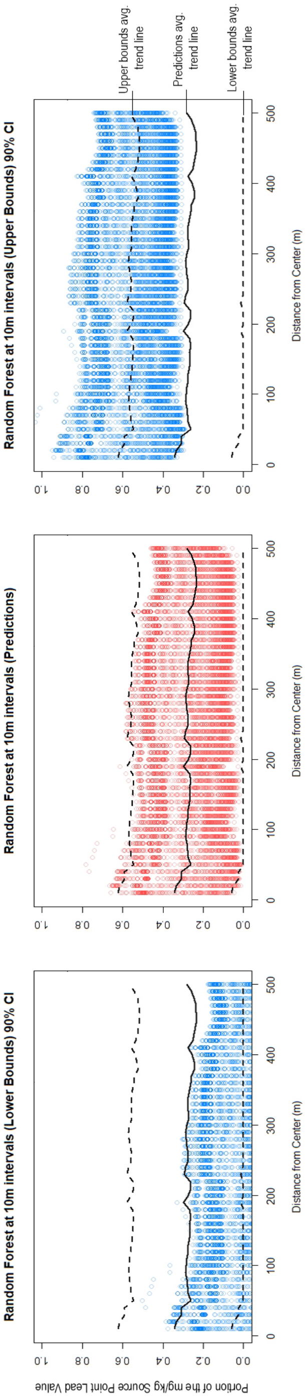

FIGURE A.8: Lower, predicted, and upper estimated share of source point soil lead distributions. Zhang et al. (2019)'s method.

Type	**Low/High Est. Fatalities**	**Low/High Cost**
Adult Cattle	10.8 – 2,880.7	$ 135,851 - $ 12,875,740
Calf Cattle	5.8 – 998.4	$ 40,416 - $ 10,780,320
Adult Buffalo	17.9 – 1,706.1	$ 178,256 - $ 9,386,051
Calf Buffalo	13.9 – 731.9	$ 37,190 - 6,401,853
Total	48.4 – 6,317.1	$ 391,715 - $ 39,443,966

TABLE A.2: lower and upper 90 % CI bounds values, Zhang et al. (2019)'s method