6 Distribution of heavy metals polluting the soil near an
7 abandoned mine in Northwestern Mexico

8 B. González-Méndez[1], R. Webster[2], R. Loredo-Portales[1], F. Molina-Freaner[3] and R.

9 Djellouli[4]

10 Correspondence B. González-Méndez

11   blancagm@protonmail.com

12 [1]CONACYT·Estación Regional del Noroeste, IGL, Universidad Nacional Autónoma

13 de México, Colosio y Madrid, 83000, Hermosillo, Sonora, Mexico

14 [2]Rothamsted Research, Harpenden AL5 2JQ, UK

15 [3]Instituto de Ecología, Universidad Nacional Autónoma de México, Hermosillo,

16 Sonora, 83000, Mexico

17 [4]Djilali Liabes University, Laboratoire Synthèse de l'information environnementale,

18 Route de Tlemcen, Sidi Bel Abbes, Algeria, 22000

19 ABSTRACT

20 Abandoned mine wastes pollute the soil in their vicinities and threaten the health of

21 livestock and human beings. This is the situation around San Felipe de Jesús in

22 northwestern Mexico. We surveyed 900 ha of agricultural land to assess and map the

23 concentrations of toxic elements in the topsoil to discover where pollution is serious,

1

what its source might be and to decide whether remediation is needed. The total concentrations of Pb, As, Zn, Cu and Mn plus Fe and Ca were analysed by X-ray fluorescence spectrometry. We found that all of the first five elements listed were concentrated near the tailings pile as a 'hot spot' and where the concentrations of Pb, As and Zn exceed national and international standards. Iron and Ca, in contrast, are evenly spread throughout the region. The elements Pb, Zn and Mn gradually decrease in concentration from the tailings pile toward the Sonora River, probably because they have spread in dust or by water during storms. Arsenic and Cu also decrease in concentration from the the tailings pile towards the centre of the region, but they increase again in the soil on the river flood plain, most likely as the result of spills in the catchment north of the river basin. These results will serve to assess the risks incurred in the use of the land for agriculture and to define policies for that use and possible remediation.

1. Introduction

Mining has caused severe damage to the environment. In particular, metal mining has led to widespread pollution of soil with potentially toxic elements; its legacy goes back centuries (Dudka and Adriano 1997; UNEP 2001). There are thought to be several million abandoned mines around the world. A rough estimate puts the number at more than 600 000 abandoned mines in South Africa, Australia, UK, USA and Canada alone (IIED 2002). In Mexico 585 abandoned mines and their associated tailings are recorded in the latest geo-referenced inventory, but that number will almost certainly grow as many more sites are identified and are verified (SEMARNAT, 2021). Many mines were abandoned when ore bodies were exhausted or when they became unprofitable. Waste materials were left in piles, exposed to rain

2

and wind, and without vegetative cover they spread their toxic loads into their surroundings for years afterwards (IIED 2002). If nothing is done to prevent it, the waste will continue as a source of potentially toxic elements (PTE) spread by the erosion of tailings, in wind-blown dust, and in drainage water. Leaching of the elements can also acidify the soil, ground water and surface water (Dudka and Adriano 1997). The fates of the elements once in the soil depend to some degree on the nature of the soil itself. Most elements are more mobile in acid soil than in alkaline or calcareous soil (Alloway 2012) and more likely to be leached from the soil.

Metals in tailings are among the most damaging legacies of mining in that they can cascade through the environment into plants and animals and eventually into human food (Cross et al. 2017). As above, if no action is taken then the pollution continues to harm the environment and to threaten the health and safety of both humans and their livestock. The most immediate need in most cases is the assessment of the concentrations of the pollutants and their distributions in the affected land.

The distributions of metal pollutants in soil can vary from one spatial scale to another because of the natural variation in the soil itself and differences in land management (Yun et al. 2020). Farmers, their advisors and agencies responsible for restrictions on land use or remediation need to understand where and on what scale pollutants are spatially distributed to decide how to manage the land safely and to develop suitable strategies and methods for soil remediation.

Mexico is rich in mineral ores. From the viceroyalty of the 16th century and into the 20th century extraction was inefficient, and it left large amounts of metal-rich waste. There was little or no concern for the damage it might do to the environment or for regulation (Douglas and Hansen 2008). Detailed records are few, and the extents of lands affected are largely unknown, both in Mexico as a whole and in Sonora in particular. The tailings deposit around San Felipe de Jesús, the study of which we describe below, is one example of the legacy left by mining. The concentrations of lead (Pb), arsenic (As) and zinc (Zn) in the deposit and nearby soil

3

exceed national and international standards (Del Rio et al. 2019). These elements can be taken up by plants (Loredo-Portales et al. 2020), and, given that this land is used for agriculture, they represent a serious threat to the safety of food for human consumption. We surveyed this area to map the distributions of the potentially toxic elements to identify where remediation is urgent or desirable and where the mobility of pollutant metals should be studied.

2. Materials and Methods

2.1 Case study: San Felipe de Jesús

The study was done in San Felipe de Jesús and Aconchi, Sonora, in Northwestern, Mexico (Fig. 1). The two towns lie contiguous to one another along the Sonora River within the Sonora River basin. The regional climate is arid (BSO) with average monthly temperature ranging from 12.3 °C in January to 30.4 °C in July, but maximum temperature can reach 47 °C (Brito-Castillo et al. 2010). The average annual precipitation is approximately 481 mm, with a range from 300 to 600 mm. Most of the rain falls in July and August (summer) in short spells (SMN, 2020). The natural vegetation is thorn-scrub dominated by leguminous trees and cacti (Martínez-Yrízar et al. 2010).

Mining started in the region in about 1900. Sampling from the mine workings in 1932 gave grades up to 16.21 oz/ton (470 g tonne^{-1}) silver, 21.7 % lead, 29.5 % zinc and 27.65 % copper. There are no records of production, but as much as 100 tonnes ore are estimated to have been extracted per day on average. Mining was suspended in 1944 because of low metal prices. Mining resumed briefly from 1957 to 1959 and recommenced again from 1963 to 1968. In 1973 a flotation plant was constructed for processing ore, and that functioned until 1991 (Tietz 2018). The abandoned laboratories still exist, and in them can be seen the remains of the chemicals used to analyse the samples.

Waste from the mine was piled in Aconchi, 0.5 km to the south of San Felipe de

4

Jesús (Fig. 1). The pile is 140 m to 160 m across at its base, covering approximately 16 300 m^2, and with a height varying from 2 to 5 m (Espinoza-Madero 2012). The residues in this pile seem to be the main source of pollution in the neighbouring agricultural land. The pile is still completely free of vegetation, is subject to wind erosion during the dry season, and in the summer heavy bursts of rain erode gullies. During the rainy season, a small stream (named El lavadero) connects the pile to the Sonora River. Additionally, efflorescent minerals consisting of white crusts have precipitated on top of the pile by evaporation. These materials can concentrate toxic elements, and are easily soluble and dispersed by wind, contributing to dispersion of the elements into the surrounding environment (Bea et al. 2010; Del Rio-Salas et al. 2019; Loredo-Portales et al. 2020).

We selected for study an area of 900 ha, most of which is agricultural, along the Sonora River (with its northwest corner at 572305.56 E, 3303770.15 N to its southeast corner at 574670.22 E, 3299861.47 N) and close to the abandoned mine tailing deposit at 572717.27 E, 3302399.27 N (Fig. 1). The soil comprises Regosols, Fluvisols, and Phaeozems (1:250 000; INEGI 2005), with 40 % or more of sand. It has a pH ranging between 6.1 and 8.7 (in water), electrical conductivity 25 to 342 µS m^{-1} (in water), 0.1 to 1.2 % of C, and 0.03 to 0.16 % of N.

Water is extracted from wells and the Sonora River for irrigation. Agriculture and cattle raising are the most economically important activities in the area. Agriculture is practised on the flood plain of the Sonora river. The main crops for human consumption are groundnuts (*Arachis hypogaea*), garlic (*Allium sativum*) and maize (*Zea mays*), whereas alfalfa (*Medicago sativa*) and barley (*Hordeum vulgare*) are the most important forage crops for livestock (SIAP 2019).

2.2. Survey

We mentioned above that the environmental damage and risks to the health of both humans and their livestock caused by toxic elements depends mainly on their concentrations and distributions. Our first task at San Felipe de Jesús was to assess

these for five potentially toxic elements, namely, lead (Pb), arsenic (As), zinc (Zn), copper (Cu) and manganese (Mn), and to map them. All five had been reported to be present in large concentrations in the mine tailings (Del Rio-Salas et al. 2019; Loredo-Portales et al. 2020), no other potentially toxic elements were found to be present in important concentrations. We added calcium (Ca) to our list for analysis, not because it is toxic but because it might help us to understand the mobility of the toxic metals. We also measured the concentration of iron (Fe) since it displays a conservative behaviour in the basin (Calmus et al. 2018). Despite the earlier studies, which focused on the concentrations of the elements in the mine tailings themselves, we knew nothing of the spatial scales of variation of the elements in the agricultural soil and so did not know how densely to sample for mapping, for which we should use kriging, the current best practice. Too sparse sampling could make kriging impracticable for lack of spatial correlation in the data; dense sampling on the other hand might be unnecessarily expensive and exceed the budget. Finding a suitable compromise has been a common problem in environmental science for many years. As Marchant and Lark (2007) pointed out, by sampling in two or more stages one can design efficient surveys for mapping; an initial stage provides rough estimates of the spatial scale(s) of variation, and later stages can fill in the gaps by grid sampling and concentrated where the contamination seems most serious.

2.3. Principles of nested sampling

Pollutants from abandoned mine tailings are spread by wind and water to varying extents and are not all equally mobile. Their distributions on neighbouring land can be further modified by the way the land is managed. So before one can design a sampling scheme suitable for mapping the distributions one needs to know what the spatial scales of variation are, as Lark et al. (2017) pointed out.

Youden and Mehlich (1937) were the first to propose a spatially nested sampling design to discover the spatial scales of variation in soil. They sampled soil at locations arranged hierarchically into clusters separated by fixed distances but with

6

random orientations. Each distance corresponded to one level of the hierarchy, and at each sampling location they selected two substations, and so on. An analysis of variance of their measurements allowed them to partition the variance of the measured properties into components associated with each level of the design. By accumulating the components in sequence from the smallest to the largest distance one can obtain a crude variogram. The technique lay dormant in soil survey until Webster and Butler (1976) resurrected it for a soil survey in the Southern Tablelands of Australia. In both surveys the designs were balanced with four levels. Adding more levels to refine the spatial structure while maintaining balance would soon make the technique unaffordable because the size of the sample would double with each added level. Further, the doubled degrees of freedom at the lower levels would be unnecessarily large for estimation of the components of variance for the smallest separating distances.

Since then the basic design has been elaborated, sacrificing balance for economy. Oliver and Webster (1987), for example, designed a scheme with five levels but without doubling the sampling at the lowest level, and Atteia et al. (1994) extended the principle to six levels without doubling the sampling in the fifth and sixth levels. More recently Lark (2011) devised a strategy for optimizing such nested schemes (see also Webster and Lark 2013), and Lark et al. (2017) applied it in a survey of heavy metals in the soils near a large tailings dam in Zambia. We adapted the strategy for our survey of the polluted soil at San Felipe de Jesús.

2.4. Implementation of nested sampling

Our initial sampling was an unbalanced nested design with six stages with distances increasing in an approximately threefold progression from 3.6 m to 1050 m (3.6, 11, 33, 100, 300 and 1050 m). The first stage comprised eight main centres placed randomly over the region with an average distance between nearest neighbours of approximately 1050 m (Fig. 1). From each main centre, three second sites were chosen 300 m apart on an equilateral triangle (Fig. 2). From each vertex of the

7

triangle, five sites were allocated 100 m away to comprise the third-stage. The next

level contains five sites at 33 m separation, the fifth and sixth levels are composed of

four sites at 11 m separation, and three sites at 3.6 m separation, respectively

(Fig. 2). This gave a total of eight main centres, with 20 points to each, and therefore

160 soil sampling points in all. At each site at any one stage, from the second level

onwards, points were placed on random orientations to comply with the random

effects model. The sampling points are shown by red discs in Fig. 1. Once the site

was located, we used a GPS (Garmin eTrex10) to geo-reference the point. Table 1

sets out the corresponding analysis of variance for this design.

   At each sampling point in the design we took five samples of topsoil (0–30 cm) at

the vertices and centre of a square of 50 m × 50 cm and bulked them. Each sample

was put in paper bag in the field, air-dried in the laboratory and sieved to pass

2 mm. The sieved sample was reduced to 30 mg by coning and quartering, and this

sub-sample was milled in an agate ball mill according to EPA protocol (600 A). The

samples were analysed by a portable X-ray fluorescence spectrometer (XRF, Niton

XL3t Ultra) to measure total concentrations of Pb, As, Zn, Cu, Mn, Fe and Ca. Data

from the manufacturer assured us of its accuracy, and we verified its accuracy against

the reference material NIST-2710a provided by the manufacturer after every 20

samples. There was no significant deviation from known values. The main source of

error in the measurement of elements in soil by the technique is the heterogeneity

within the soil samples themselves, as Ravansari et al. (2020) have pointed out. To

diminish this error measurements were made in triplicate and mean values calculated.

The standard errors are listed in Table 2.

   The structure of the sampling can be represented in a table as for an analysis of

variance (ANOVA). Table 1 lists the degrees of freedom with the corresponding

distances. Our main aim is to estimate the components of variance at these distances,

and so we have used residual maximum likelihood (REML) (Patterson and Thompson

1971) for the purpose because it is more efficient than ANOVA. Lark (2011) sets out

the mathematics of the REML solution, and we do not repeat it here. The estimated components of variance were summed to give rough variograms (Fig. 3). Note that the concentrations for all elements except Fe and Ca were transformed to common logarithms to give distributions that were approximately symmetric. The transformations are listed in Table 2 for the whole data (see below).

The variograms deriving from the nested analysis, and shown in Fig. 3, are too rough for use in kriging. We wanted to improve the estimates between 11 and 33 m, and so we added 50 points 20 m away from 50 of the original 160 sampling points on random orientations. These are shown as green stars in Fig. 1. Finally, as one can see in Fig. 1, there were still large gaps between the nests, and we should want to place further points in these gaps for kriging. Otherwise there would be large errors in the kriged predictions. We therefore added a further 51 points at the nodes of a 220-m regular grid wherever nodes lay more than 200 m from a point in the nests. These points are shown as yellow + symbols in Fig. 1.

Samples of topsoil were taken from these additional locations and analysed by X-ray fluorescence spectroscopy in the same way as for the original 160. We thus had measurements for all elements at a total of 261 of locations from which to map the concentrations.

2.5. Geostatistical analysis: Variograms and their modelling

The complete set of data comprised the measured concentrations on (1) soil sampled at sites of the original nested design, (2) a set of sites chosen close to 20 m from any of the previously sampled sites and (3) sites on a grid at 220-m intervals in those parts of the region with large gaps. In all there were 261 sampling sites providing 261 values, bar a few missing ones, for each metal. Table 2 summarizes the data, both on the original scales and after transformation where desirable. Although we did not analyse Fe geostatistically, we include it in the summary and in the principal components analysis (see below), because it helps to understand the calcium pattern.

We computed the Pearson correlation coefficients among the elements and did a

9

principal components analysis on the correlation matrix for reasons that we explain below. The results are summarized in Table 3, from which one sees that almost 85 % of the variance lies in the leading two components.

The sites are strongly clustered, one consequence of which is that the experimental variograms computed by the usual method of moments have strong peaks and troughs, which make modelling them uncertain. Marchant et al. (2013) found that in such a situation maximum likelihood estimation is better and gives stable results. It also has the advantage of fitting models over the whole range of the region. We used specifically residual maximum likelihood, REML, for the purpose. Having fitted models in this way, we compared the two most plausible models, exponential and spherical, by cross-validation. We did so by omitting each point in turn and predicting the value there by ordinary kriging from the rest of the data. The validated parameters were then ones finally be needed for interpolation and mapping.

Table 4 lists the parameter estimates and cross-validation statistics for the spherical models, which fitted best and for which the equation is

$$
\begin{aligned}
\gamma(h) \quad &= \quad c_0 + c_1 \left\{ \frac{3h}{2r} - \frac{1}{2}\left(\frac{h}{r}\right)^3 \right\} \quad \text{for} \ \ h \leq r \\
&= \quad c_0 + c_1 \quad \text{for} \ \ h > r \\
&= \quad 0 \quad \text{for} \ \ h = 0 \ .
\end{aligned} \tag{1}
$$

The parameters are the variances c_0, the nugget variance, and c_1, and the range r. We have treated the variation as isotropic, so that the lag h is a scalar in distance only.

The cross-validation statistics are the mean error of prediction (ME), the mean squared error of prediction (MSE) and mean square deviation ratio (MSDR), i.e. the ratio of the squared deviation to the kriging variance. They are as follows in which $z(\mathbf{x}_i)$ is the observed value at $\mathbf{x}_i$, $\widehat{Z}(\mathbf{x}_i)$ is the predicted value there and $\sigma_{\mathrm{K}}^2(\mathbf{x}_i)$ is the kriging variance. The averages are over the n data.

$$
\mathrm{ME} \ = \ \frac{1}{n} \sum_{i=1}^{n} z(\mathbf{x}_i) - \widehat{Z}(\mathbf{x}_i) \ .
$$

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} \left\{ z(\mathbf{x}_i) - \widehat{Z}(\mathbf{x}_i) \right\}^2 .$$

$$\text{MSDR} = \frac{1}{n} \sum_{i=1}^{n} \frac{\left\{ z(\mathbf{x}_i) - \widehat{Z}(\mathbf{x}_i) \right\}^2}{\sigma_{\mathrm{K}}^2(\mathbf{x}_i)} .$$

We have added the median of the squared deviation ratio (medSDR):

$$\text{medSDR} = \text{median} \left[\frac{\left\{ z(\mathbf{x}_i) - \widehat{Z}(\mathbf{x}_i) \right\}^2}{\sigma_{\mathrm{K}}^2(\mathbf{x}_i)} \right] .$$

The mean errors are all close to zero, which is to be expected; kriging is an unbiased predictor. The mean squared errors are small. The important diagnostic is the MSDR. Ideally this should be 1; i.e. the squared deviation between the observed and predicted value should equal the kriging prediction error variance. The MSDRs for the first five elements listed in Table 4 are all close to 1; that for calcium is also sufficiently close to justify our accepting the model tabulated. The table includes the variances of the data, s^2, for comparison with the sill variances, $c_0 + c_1$, of the models. The median of the squared deviation ratios should be close to 0.455 for a true model. All are less than this value; only that for copper is close.

3. Results

3.1 REML analysis and variograms

Summary statistics of concentrations for the elements are listed in Table 2. Among the elements, Ca had the largest mean concentration (1.81 %), while the smallest was for As (20.14 mg kg^{-1}). The mean concentrations of other metals in order of magnitude were Mn 786 mg kg^{-1}, Zn 233 mg kg^{-1}, Pb 95.9 mg kg^{-1} and Cu 41.9 mg kg^{-1}. The complete set of data exhibited a wide variation, with total concentrations varying from 0.7 to 4.3 % for Ca, 67.7 to 3128 mg kg^{-1} for Zn, 425 to 1955 mg kg^{-1} for Mn, 18.6 to 896 mg kg^{-1} for Pb, 23.1 to 185 mg kg^{-1} for Cu, and 11.3 to 87.5 for As.

Notice that all except Ca had skewed distributions, and that is why we transformed the concentrations to logarithms to stabilize the variances. At several points, Pb and As exceeded the national guide values (400 and 22 mg kg^{-1}, respectively, DOF 2007).

The REML analysis of the nested sampling revealed that most of the variance occurs at distances between 33 and 100 m. Figure 3 shows that only small proportions of the variances for As, Cu, Mn and Ca occur at less than 33 m. Nevertheless, as Lark and Marchant (2018) pointed out, it is good practice to include sampling points close to one another to ensure that variograms are well estimated at short lag distances because those estimates have a large effect on the uncertainty of kriging predictions. So we refined the nested sampling by choosing new 50 sampling points at 20 m far from any of the previous nested points, and then filled the gaps with 50 more points.

3.2. The maps of concentrations

The distributions of concentrations were spatially dependent, the variograms of the logarithms of the concentrations of Pb, Zn and Mn and of the concentration of Ca (Fig. 4) were in general, well structured with small nugget variances. The variograms of As and Cu had proportionately larger nugget variances; mainly, we think, because the error variances in the measurements are proportionately more. Iron showed no spatial dependence; it seemed to be uniformly distributed in the region.

Figures 5 to 10 show the spatial distributions of the concentration of the elements in the soil of agricultural land. All the elements but Ca (Fig. 10) are strongly concentrated around the tailings deposit, particularly to the west, the principal hot spot. Lead, Mn and Zn have similar spatial patterns with four other areas relatively rich in the three elements (Figures 5, 6 and 7). One such area is in the centre of the region with values 2.85 log$_{10}$(mg kg^{-1}) for Pb, 3.15 log$_{10}$(mg kg^{-1}) for Mn and 2.9 log$_{10}$(mg kg^{-1}) for Zn. Another region relatively rich in these metals is somewhat to the south east of it, though with somewhat smaller concentrations. A fairly narrow belt of land also relatively rich in Pb and Mn extends south from the tailings deposit with concentrations reaching 2.1 log$_{10}$(mg kg^{-1}) for Pb and 2.9 log$_{10}$(mg kg^{-1}) for

12

Mn.

As above, As and Cu are concentrated in the hot spot surroundings of the tailings, In addition, both are relatively rich in the soil on either side of the Sonora River with concentrations of 1.25–1.30 $\log_{10}(\text{mg kg}^{-1})$ for As and 1.45–1.50 $\log_{10}(\text{mg kg}^{-1})$ for Cu. We suggest an explanation below. Elsewhere in the region their concentrations are less.

Calcium is the most abundant metal that we measured. Its spatial distribution is evidently unrelated to the other metals, and it seems unaffected by the tailings (Fig. 10). It is the only element where concentrations are less close to the sources of the tailings than elsewhere, and where Pb, Zn and Mn are richest.

The map of kriging variances for Pb, Fig. 11, shows how the prediction errors depend on the positions of the sampling points. The denser is the sampling, the smaller is the kriging variances. The maps of the kriging variances for the other elements have similar patterns, though the variances themselves are different, of course.

Finally, as noted above, the three elements Pb, Zn and Mn, have similar spatial patterns; their patterns differ from those of As and Cu, and all differ substantially from the distribution of Ca. This distinction is neatly summarized in the correlation circle obtained from the principal components analysis (Fig. 12). The metals Pb, Zn and Mn are strongly correlated with one another and appear as a cluster of points close to the extreme right of the circle. Arsenic and Cu appear away from them, upper right, and Ca, evidently fairly closely related to Fe appears far away in the upper left quadrant.

4. Discussion

Sampling to the nested design and the analysis of the data provided a sound guide for the subsequent grid survey for mapping. It showed at what spacings most of the variance occurs and which turned out to be at less than 100 m. Plots of the data on a map of the region also showed that the largest concentrations were near the pile of

13

tailings. Those plots and the kriged maps show how the pollutant elements are concentrated around the tailings deposit; that deposit is a hot spot and evidently a major source of pollution. The elements Pb, Mn and Zn show strong spatial similarities that suggest a common transport process. In addition, the concentrations of As and Cu have a spatial pattern associated with the Sonora River, indicating an additional source of pollution. In contrast, Ca is less concentrated around the tailings deposit; it seems unrelated to the mining.

The dispersion of the elements around the tailings pile is likely to been caused by the combined effects of water and wind. This combination of processes is widespread in arid and semi-arid regions where erosion by wind and water alternate with the changing seasons and interact with each other; it is a process that differs in its effects from those of wind and water separately (Yang, et al. 2019). Tuo et al. (2014) found that the combined effect of wind and water erosion of the soil surface (0–1 cm) removed fine particles (< 0.01 mm) preferentially, leaving coarser particles (> 0.05 mm) in place. This suggests that the pollutant elements have been carried attached to the finer particles in the tailings and spread by this complex process.

Heavy rain, driven by moderate to strong wind, is especially erosive (Marzen et al. 2017), and it is likely to have re-distributed particles from the tailings in the patterns we observe in Figures 5–10. The rose diagrams in those figures show two predominant directions of the wind, namely towards north north east and south south east. Their velocities, ranging from 12 to 38 km hour^{-1}, combined with heavy rain in short spells during summer are quite sufficient to carry material from the tailings.

The gully erosion of the tailings is likely to have contributed substantially to the enrichment of metals in the surroundings and toward to the Sonora River following the path of the El Lavadero stream. As the soil has a large proportion of sand (40 % or more) and contains little organic matter, it has rather few active sites on to which metals can bind, thereby allowing the metals to be transported by leaching in infiltrating water or in run-off. Efflorescent salts in the San Felipe tailings are rich in

14

Pb and Mn in particular (Del Rio-Salas et al. 2019). The fine fractions of these salts are susceptible to wind erosion because of the weak cohesion between particles (Sanchez-Bisquet et al. 2017). Wind carries significant amounts of dust from tailings deposits following the dominant wind direction (Moreno-Brotons et al. 2010) and creating trends with increasing distance from tailings (Lark et al. 2017; Djebbi et al. 2017). Thus, Pb and Mn could be dispersed several hundred metres from the tailings in the form of efflorescent salts. When the rain is especially heavy flooding spreads the pollutants, both in solution and as particles, over the flood plain to generate the spatial patterns that we observe in our study. The maps show where the pollutants are so concentrated that remediation should be considered. They also show where to prioritize further studies on the mobility of the pollutants in the light of other properties of the soil that are likely to enhance or retard mobility.

It is evident in Figs 8 and 9 that much of the As and Cu derives from the tailings. Arsenic is also spread more widely, with some of its larger concentrations close to the Sonora River. It is likely that some of this As has come from spills from mines in the northern sector of the Sonora basin. Gomez-Alvarez et al. (1990, 1993), SEMARNAT (2014) and Silva-Rodriguez (2019) have documented such spills from mine wastes to the north of our region. Some of those discharges were rich in As and Cu, and after attaching themselves to soil and sediment they remained along the river channel (Rivera-Uria et al. 2018). The metal-enriched material could then be re-mobilized during heavy rain and dispersed downstream on the flood plain (Foulds et al. 2014). It is likely therefore that the current spatial pattern of As and Cu arises from mine discharges at various times in the past. Martín-Peinado et al. (2015) reported similar persistent residual pollution (including As and Cu) 15 years after a mine spill in Aznalcòllar, Spain. Although the remediation measures were implemented immediately, spilled material from the tailings remained mixed with the soil as a major source of pollution (García-Carmona et al. 2019).

Arsenic and Cu are concentrated around the El lavadero stream which connects

the tailings pile with the Sonora River (Fig. 5 and 6). The water-soluble salts of Mn, Zn and Pb in the efflorescent deposits could contaminate water from the tailings (Del Rio-Salas et al. 2019). These salts would then carried via channels into the Sonora River.

Nieva et al. (2021) found that the mineralogical composition of the efflorescent salts depended on climate (specifically climate with alternating dry and wet seasons). They found that during the wet season, copiapite is the dominant mineral in the salts precipitated in the pores of the tailings, where the arsenates substituted the sulfates, converting the copiapite into an As reservoir. This arsenic can be released during the short spells of summer rain. Del Rio-Salas et al. (2019) found that the efflorescent salts of San Felipe de Jesús contain up to 26% of copiapite, so this could be an important process for the spread of As from the tailings.

We still need better understanding of metal pollution in this region. Not only have metals from the mine waste polluted the soil, they are also mobile in the soil and likely to be taken up by plants. Loredo et al. (2020) analysed samples of the soil from close to the front of the mine tailings. They found that Zn and As are highly mobile in the soil there and that Zn and Pb exceeded the threshold limits of phytoaccesibility. Such assessments need to be extended throughout the 900-ha region where concentrations are now seen to be large.

5. Conclusions

Our experience of splitting the survey of pollutant metals in the soil at San Felipe de Jesús in Northwestern Mexico into two stages shows the merit of preceding grid survey for mapping with a nested design and analysis to establish the scale(s) at which most variance occurs. It allowed us to plan an affordable sampling in the second stage that would provide predictions with acceptable error. Despite the several papers setting out the procedure and software now embodying REML for the analysis of spatially nested data the technique seems under-used. We hope our paper will be read and followed by other scientists and technicians who map and try to

16

understand the distribution of pollutants.

The survey revealed widespread large concentrations of Pb, As, Zn, Cu and Mn in the soil of the region. The maps made by kriging from the sample data show clearly that the largest concentrations are associated with the tailings deposit on the western margin of the region. Concentrations of Pb, Zn and Mn decrease with increasing distance from the deposit, and it seems likely that the metals were transported by wind and water from the tailings. Arsenic and Cu are also concentrated close to the Sonora River, almost certainly with material from mine spills north of the basin. Land managers and responsible agencies can now focus on those parts of the region most seriously affected to restrict agriculture and plan feasible remediation.

Authors' Contributions

B. González-Méndez: Conceptualization, investigation, writing–original draft, programming **R. Webster**: Methodology, programming, formal analysis, writing–original draft **R. Loredo-Portales:** Sampling, writing–review and editing **F. Molina-Freaner:** Writing–review and editing, funding acquisition **R. Djellouli:** Visualization, graphics

Conflict of interest

The authors declare that they have no conflict of interest.

References

Alloway BJ (2012) Heavy Metals in Soils: Trace Metals and Metalloids in Soils and their Bioavailability. 3rd edition. Springer, Dordrecht.

Atteia O, Webster R, Dubois JP (1994) Geostatistical analysis of soil contamination in the Swiss Jura. Env Pol 86:315–327

Bea SA, Ayora C, Carrera J, Saaltink MW, Dold B (2010) Geochemical and environmental controls on the genesis of soluble efflorescent salts in coastal

mine tailings deposits: A discussion based on reactive transport modeling. J Cont Hydrol 111(1-4): 65–82

Brito-Castillo L, Crimmins MA, Díaz SC (2010) Clima. In: Molina-Freaner F, Van Devender TR (Eds.), Diversidad Biológica de Sonora. UNAM/CONABIO, México, pp. 73-96

Calmus T, Valencia-Moreno M, Del Río-Salas R, Ochoa-Landín L, Mendivil-Quijada H (2018) A multi-elemental study to establish the natural background and geochemical anomalies in rocks from the Sonora river upper basin, NW Mexico. Rev Mex Ciencias Geol 35(2):158–167. DOI:http://dx.doi.org/10.22201/cgeo.20072902e.2018.2.605

Cross AT, Stevens JC, Dixon KW (2017) One giant leap for mankind: can ecopoiesis avert mine tailings disasters. Plant and Soil 421:1–5

Del Rio-Salas R, Ayala-Ramírez Y, Loredo-Portales R, Romero F, Molina-Freaner F, Minjarez-Osorio C, Pi-Puig T, Ochoa-Landin L, Moreno-Rodríguez, V (2019) Mineralogy and geochemistry of rural road dust and nearby mine tailings: a case of ignored pollution hazard from an abandoned mining site in semi-arid zone. Nat Resour Res 28:1485–1503

Djebbi C, Chaabani F, Font O, Queralt I, Querol X (2017) Atmospheric dust deposition on soils around an abandoned fluorite mine (Hammam Zriba, NE Tunisia). Environ Res 158:153–166

Douglas L, Hansen T (2008) La riqueza escondida en el desierto: la búsqueda de metales preciosos en el noroeste de Sonora durante los siglos XVIII y XIX. Región y Sociedad 20(42):165–190

Dudka S, Adriano DC (1997) Environmental impacts of metal ore mining and processing: a review. J Env Qual 26:590–602

Espinoza-Madero Z (2012) Impacto ambiental producido por los jales de San Felipe de Jesús. Tesis de Licenciatura, Departamento de Geología, Universidad de Sonora, Hermosillo

Foulds SA, Brewer PA, Macklin MG, Haresign W, Betson RE, Rassner SME (2014) Flood-related contamination in catchments affected by historical metal mining: an unexpected and emerging hazard of climate change. Sci Tot Env 476-477:165–180

García-Carmona M, García-Robles H, Turpín Torrano C, Fernández Ondoño E, Lorite Moreno J, Sierra Aragón M, Martín-Peinado FJ (2019) Residual pollution and vegetation distribution in amended soils 20 years after a pyrite mine tailings spill (Aznalcóllar, Spain). Sci Total Environ 650:933–940

Gomez-Alvarez AMT, Yocupicio-Anaya J, Ortega-Romero P (1990) Niveles y distribución de metals pesados en el Rio Sonora y su afluente el Rio Bacanuchi, Sonora, Mexico. Ecologica 1:10–20

Gomez-Alvarez AMT, Yocupicio-Anaya J, Ortega-Romero P (1993) Concentraciones de Cu, Fe, Mn, Pb y Zn en los sedimentos del Rio Sonora y de su afluente el Rio Bacanuchi, Sonora, Mexico. Boletin del Departamento de Geología de la Universidad de Sonora 10:49–62.

IIED (International Institute for Environment and development) (2002) Mining for the future: Appendix C abandoned mines working paper. Mining, Minerals and Sustainable Development 28: 1–20

Lark RM (2011) Spatially nested sampling schemes for spatial variance components: scope for their optimization. Comp & Geosci 37:1633–1641

Lark RM, Hamilton EM, Kaninga B, Maseka KK, Mutondo M, Sakala GM, Watts MJ (2017) Nested sampling and spatial analysis for reconnaisance investigations

of soil: an example from agricultural land near mine tailings in Zambia. Eur J Soil Sci 68:605–620

Lark RM, Marchant BP (2018) How should a spatial-coverage sample design for a geostatistical soil survey be supplemented to support estimation of spatial covariance parameters. Geoderma 319:89–99

Lark RM, Hamilton EM, Kaninga B, Maseka KK, Mutondo M, Sakala GM, Watts MJ (2017) Nested sampling and spatial analysis for reconnaissance investigations of soil: an example from agricultural land near mine tailings in Zambia. Eur J Soil Sci 68:605–620

Loredo-Portales R, Bustamante-Arce J, González-Villa HN, Moreno-Rodríguez V, Del Rio-Salas R, Molina-Freaner F, González-Méndez B, Archundia-Peralta D (2020). Mobility and accessibility of Zn, Pb, and As in abandoned mine tailings of northwestern Mexico. Env Sci Pol Res 27:26606–26620

Marchant BP, Lark RM (2007) Optimized sample schemes for geostatistical surveys. Math Geol 39:113–134

Marchant BP, Viscarra Rossel RA, Webster R (2013) Fluctuations in method-of-moments variograms caused by clustered sampling and their elimination by declustering and residual maximum likelihood estimation. Eur J Soil Sci 64:401–409

Martín-Peinado FJ, Romero-Freire A, García Fernández I, Sierra Aragón M, Ortiz-Bernard I, Simón Torres M (2015) Long-term contamination in a recovered area affected by a mining spill. Sci Total Environ 514:219–223

Martínez-Yrízar A, Felger RS, Búrquez A (2010) Los Ecosistemas de Sonora: un diverso capital natural. In: Molina Freaner F, Van Devender TR (Eds) Diversidad biológica de Sonora. UNAM/CONABIO, México, pp. 129-156.

Marzen M, Iserloh T, de Lima JLMP, Fister W, Ries JB (2017) Impact of severe
    rain storms on soil erosion: Experimental evaluation of wind-driven rain and its
    implications for natural hazard management. Sci Total Environ
    590-591:502–513

Moreno-Brotons J, Romero-Diaz A, Alonso-Sarria F, Belmonte-Serra F (2010) Wind
    erosion on mining waste in southeast Spain. Land Deg Dev 21:196–209

Nieva NE, Garcia MG, Borgnino L, Borda LG (2021) The role of efflorescent salts
    associated with sulfide-rich mine wastes in the short-term cycling of arsenic:
    Insights from XRD, XAS, and XRF studies. J Haz Mat
    https://doi.org/10.1016/j.jhazmat.2020.124158

DOF (2007). NOM-147-SEMARNAT/SSA1-2004, que establece criterios para
    determinar las concentraciones de remediación de suelos contaminados por As,
    Ba, Be, Cd, Cr, Hg, Ni, Ag, Pb, Se, Tl y V. Secretaría de Medio Ambiente y
    Recursos Naturales Mexico 35–94

Oliver MA, Webster R (1987) The elucidation of soil pattern in the Wyre Forest of
    the West Midlands. II. Spatial distribution. J Soil Sci 38:293–307

Patterson HD, Thompson R (1971) Recovery of inter-block information when block
    sizes are unequal. Biometrika, 58, 545–554

Ravansari R, Wilson SC, Tighe M (2020) Portable X-ray fluorescence for
    environmental assessment of soils: not just a point and shoot method.
    Environment International, 134, article 105250.

Rivera-Uria MY, Ziegler-Rivera RA, Diaz-Ortega J, Prado-Pano B, Romero FM
    (2018) Effect of an acid mine spill on soils in Sonora river basin:
    micromorphological indicators. Span J Soil Sci 8:258–274

Sánchez-Bisquert D, Castejón JMP, García-Fernández G (2017) The impact of
    atmospheric dust deposition and trace elements levels on the villages

surrounding the former mining areas in a semi-arid environment (SE Spain). Atm Env 152:256–269

SEMARNAT (2014) Derrame de sulfato de cobre en el Río Bacanuchi, afluente del Río Sonora. Consulted on November 2020. `https://www.gob.mx/cms/uploads/attachment/file/338899/` `21DPpresentacion_conferencia_derrame.pdf`

SEMARNAT (2021) Inventario homologado preliminar de presas de jales. Consulted on October 2021. `https://geomaticaportal.semarnat.gob.mx/arcgisp/` `apps/webappviewer/index.html?id=95841aa3b6534cdfbe3f53b3b5d6edfa`

SIAP (2019) Estadística de Producción Agrícola. Consulted on November 2020. `http://infosiap.siap.gob.mx/gobmx/datosAbiertos_a.php`

Silva-Rodriguez JM (2019) 100 años de contaminación de la minería en ríos de Sonora 1908-2014. In: Vega Deloya, H. (Coord.), Los derechos ambientales como paradigma social y de gobierno en Sonora: el caso del Rio Sonora y otros estudios. Universidad de Sonora, Departamento de Historia y Antropología, Hermosillo, Sonora

Tietz P (2018) Technical report and estimated resources for the San Felipe Project, Sonora, Mexico. Americas Silver Corporation, Toronto, Canada

Tuo DF, Xu MX, Ma XX, Zheng SQ (2014) Impact of wind-water alternate erosion on the characteristics of sediment particles. Chin. J. Appl. Ecol. 25:381–386.

UNEP (2001) Abandoned Mines: Problems, Issues and Policy Challenges for Decision Makers. Summary report. Chilean Copper Commission, Santiago, Chile

Webster R, Butler BE (1976) Soil classification and survey studies at Ginninderra. Austr J Soil Res 14:1–24

Webster R, Lark RM (2013) Field Sampling for Environmental Science and
    Management. Routledge, London.

Yang H, Zou X, Wang J, Shi P (2019) An experimental study on the influences of
    water erosion on wind erosion in arid and semi-arid regions. J Arid Land
    11(2):208–216

Youden WJ, Mehlich A (1937) Selection of efficient methods for soil sampling.
    Contributions of the Boyce Thompson Institute for Plant Research 9:59–70

Yun SW, Choi DK, Yu C (2020) Spatial distributions of metal(loid)s and their
    transport in agricultural soils around abandoned metal mine sites in South
    Korea. J Soils and Sediments, https://doi.org/10.1007/s11368-020-02663-7.

587 Table 1: Degrees of freedom for the nested sampling and the corresponding distances.

Source	Degrees of freedom	Distance/m
Stage 1	7	1050
Stage 2	16	300
Stage 3	29	100
Stage 4	30	33
Stage 5	39	11
Stage 6 (residual)	38	3.6

590 Table 2: Summary statistics of concentrations of the elements and of their transformations. All the
591 concentrations are in mg kg^{-1}, except for Ca and Fe which are % by mass

592

| Element | N^1 | SE^2 | Original scale | | | | | | Transformed scale | | | |
			Mean	Min.	Max.	Median	Std dev.	Skew	Mean	Median	Std dev.	Skew
Pb	257	4.3	95.9	18.6	896.4	48.2	120.7	3.91	1.57	1.48	0.52	0.24
As	248	1.5	20.1	11.3	87.5	19.30	9.84	4.18	1.27	1.28	0.16	0.26
Zn	258	9.3	233.0	67.7	3128.0	139.7	310.3	5.56	2.23	2.15	0.29	1.47
Cu	227	2.6	41.9	23.1	185.1	38.6	15.96	4.31	1.60	1.59	0.12	1.22
Mn	258	26.9	786.9	425.2	1955.0	721.9	255.1	1.61	2.88	2.86	0.12	0.76
Ca	258	0.04	1.81	0.70	4.30	1.84	0.62	0.33	-	-	-	-
Fe	258	0.07	2.35	1.05	4.56	2.34	0.49	0.69	-	-	-	-

594 [1]N is the number of sampling locations for which valid data are available.

595 [2]All the statistics are based on means of three replicate measurements for which the standard errors

596 are listed in the column headed SE.

598   Table: 3 Eigenvalues of correlation matrix

Order	Eigenvalue	Percentage of variance	Cumulative percentage
1	4.107	58.67	58.67
2	1.801	25.74	84.41
3	0.480	6.86	91.27
4	0.335	4.79	96.06
5	0.193	2.75	98.81

599

Tabel 4: Parameters of spherical models and cross-validation statistics.

Element	s^2	c_0	c_1	r/m	ME	MSE	MSDR	medSDR
Pb	0.27030	0.00776	0.24930	1051.4	-0.01070	0.0376	1.043	0.2627
As	0.02724	0.00518	0.01596	1041.3	-0.00033	0.00778	1.121	0.3078
Zn	0.08303	0.00482	0.06204	1022.1	-0.00391	0.01100	1.153	0.2894
Cu	0.01548	0.00303	0.01480	818.3	0.00227	0.00503	0.993	0.4297
Mn	0.01547	0.00183	0.00957	993.2	-0.00081	0.0032	1.140	0.3250
Ca	0.38505	0.04254	0.48697	934.3	0.0244	0.1340	1.328	0.2441

605 **Fig. 1** The region surveyed and the sampling points in Northwestern, Mexico.

606 **Fig. 2** The unbalanced nested design used: (a) the topological tree of the design; (b)

607 the design as it might appear on the ground, the red point is the main station; blue

608 lines represent nodes spaced 300 m apart, green lines indicate 100 m, purple lines link

609 points 33 m apart, orange lines indicate 11 m, and black lines are nodes separated by

610 3.6 m.

611 **Fig. 3** Variograms from the nested sampling in phase 1.

612 **Fig. 4** Variogram models fitted by REML from the whole set of data.

613 **Fig. 5** Map of Pb. The red polygon is the mine tailing, the blue dashed line

614 corresponds to the El lavadero stream, the continuous blue line is the Sonora River,

615 and the yellow lines are the levels of Pb in (log mg/kg). The rose wind was taken

616 from Del Rio-Salas et al. (2019)

617 **Fig. 6** Map of Mn. The red polygon is the mine tailing, the blue dashed line

618 corresponds to the El lavadero stream, the continuous blue line is the Sonora River,

619 and the yellow lines are the levels of Mn in (log mg/kg). The rose wind was taken

620 from Del Rio-Salas et al. (2019)

621 **Fig. 7** Map of Zn. The red polygon is the mine tailing, the blue dashed line

622 corresponds to the El lavadero stream, the continuous blue line is the Sonora River,

623 and the yellow lines are the levels of Zn in (log mg/kg). The rose wind was taken

624 from Del Rio-Salas et al. (2019)

625 **Fig. 8** Map of As. The red polygon is the mine tailing, the blue dashed line

626 corresponds to the El lavadero stream, the continuous blue line is the Sonora River,

627 and the yellow lines are the levels of As in (log mg/kg). The rose wind was taken

628 from Del Rio-Salas et al. (2019)

629 **Fig. 9** Map of Cu. The red polygon is the mine tailing, the blue dashed line

630 corresponds to the El lavadero stream, the continuous blue line is the Sonora River,

631 and the yellow lines are the levels of Cu in (log mg/kg). The rose wind was taken

from Del Rio-Salas et. al. (2019)

Fig. 10 Map of Ca. The red polygon is the mine tailing, the blue dashed line corresponds to the El lavadero stream, the continuous blue line is the Sonora River, and the yellow lines are the levels of Ca in (%)The rose wind was taken from Del Rio-Salas et. al. (2019)

Fig. 11 Error map of Pb on the logarithmic scale. The red polygon is the mine tailing, the blue dashed line corresponds to the El lavadero stream, the continuous blue line is the Sonora River, and the red lines are the estimated variance of Pb in (log mg/kg). The red to pink discs are the eight nested sampling nodes, the additional points are shown as green stars and yellow crosses

Fig. 12 Correlations between the elements and the first two principal components plotted in the unit circle.