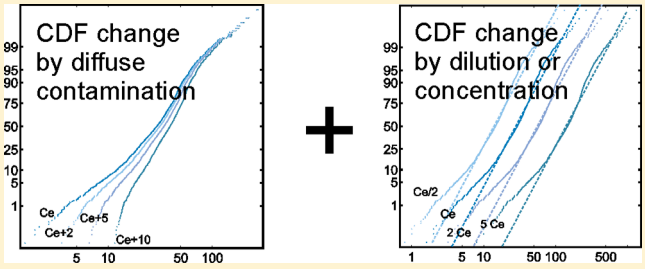
Environmental Science & Technology
ACS Editors' Choice
Article
pubs.acs.org/est

Quantifying Diffuse Contamination: Method and Application to Pb in Soil

Karl Fabian,*[†] Clemens Reimann,[†] and Patrice de Caritat[‡]

[†]Geological Survey of Norway (NGU), P.O. Box 6315 Sluppen, N-7491 Trondheim, Norway
[‡]Research School of Earth Sciences, Australian National University, Canberra, Australian Capital Territory 2601, Australia

ABSTRACT: A new method for detecting and quantifying diffuse contamination at the continental to regional scale is based on the analysis of cumulative distribution functions (CDFs). It uses cumulative probability (CP) plots for spatially representative data sets, preferably containing >1000 determinations. Simulations demonstrate how different types of contamination influence elemental CDFs of different sample media. It is found that diffuse contamination is characterized by a distinctive shift of the low-concentration end of the distribution of the studied element in its CP plot. Diffuse contamination can be detected and quantified via either (1) comparing the distribution of the contaminating element to that of an element with a geochemically comparable behavior but no contamination source (e.g., Pb vs Rb), or (2) comparing the top soil distribution of an element to the distribution of the same element in subsoil samples from the same area, taking soil forming processes into consideration. Both procedures are demonstrated for geochemical soil data sets from Europe, Australia, and the U.S.A. Several different data sets from Europe deliver comparable results at different scales. Diffuse Pb contamination in surface soil is estimated to be <0.5 mg/kg for Australia, 1−3 mg/kg for Europe, and 1−2 mg/kg, or at least <5 mg/kg, for the U.S.A. The analysis presented here also allows recognition of local contamination sources and can be used to efficiently monitor diffuse contamination at the continental to regional scale.



INTRODUCTION

Because soil formation is an extremely slow process, quantification of threats to soil quality, such as changes in element concentrations due to diffuse contamination, is of eminent importance for society. Local point-source contamination by potentially toxic elements (PTEs) near a metal smelter, power plant, industry, city, or highway is easily recognized.[1,2] Quantification of anthropogenic diffuse contamination at the country, continental or global scale is still a challenge although the European Commission identified diffuse contamination as one of the eight major threats to soil quality in Europe.[3]

On the basis of experience from the local scale, for example from the proximity of a metal smelter,[1,2] it is widely assumed that contamination always results in unusually high concentrations of the emitted element in soil, and monitoring activities are focusing on extreme values. However, dependent on signal intensity local contamination becomes indistinguishable from the natural background variation at a characteristic distance from the source, usually measured in meters to some kilometers, and in extreme cases 100−200 km.[1,4,5] It is therefore a major challenge to identify and quantify diffuse contamination in soil as distinguished from local or regional contamination, because it cannot be clearly separated from a large natural background variation. Here a new method for recognizing and quantifying diffuse contamination is presented that avoids known problems.[6−8] It compares the cumulative distribution function (CDF) of a contaminant in a near surface sampling medium with that of either (1) another element with a similar geochemical behavior that is not normally emitted by human activities, or (2) the same element in a deeper soil horizon while accounting for the effects of soil forming processes.

Cumulative distribution functions (CDFs) and cumulative probability diagrams (CP plots) have been introduced to exploration geochemistry as early as the 1950s,[9] and their use in identifying anomalies related to mineral deposits in geochemical data sets is well established[10−13] with a focus on the detection of various geochemical processes diagnosed by one or several breaks at the upper end, usually in the high 90th percentile range, of the CDF. These high values have been used to identify exploration targets for mineral deposits. CDFs also have been used in environmental geochemistry to identify unusually high element concentrations due to contamination.[14] Different types of CDF diagrams and their construction have been comprehensively discussed.[15] The prevailing view that solely the high end of the CDFs are relevant in environmental geochemistry has been challenged, and it has been pointed out that in many cases element deficiency at the continental scale

Received: February 9, 2017
Revised: April 26, 2017
Accepted: April 28, 2017
Published: April 28, 2017

DOI: 10.1021/acs.est.7b00741
Environ. Sci. Technol. 2017, 51, 6719−6726

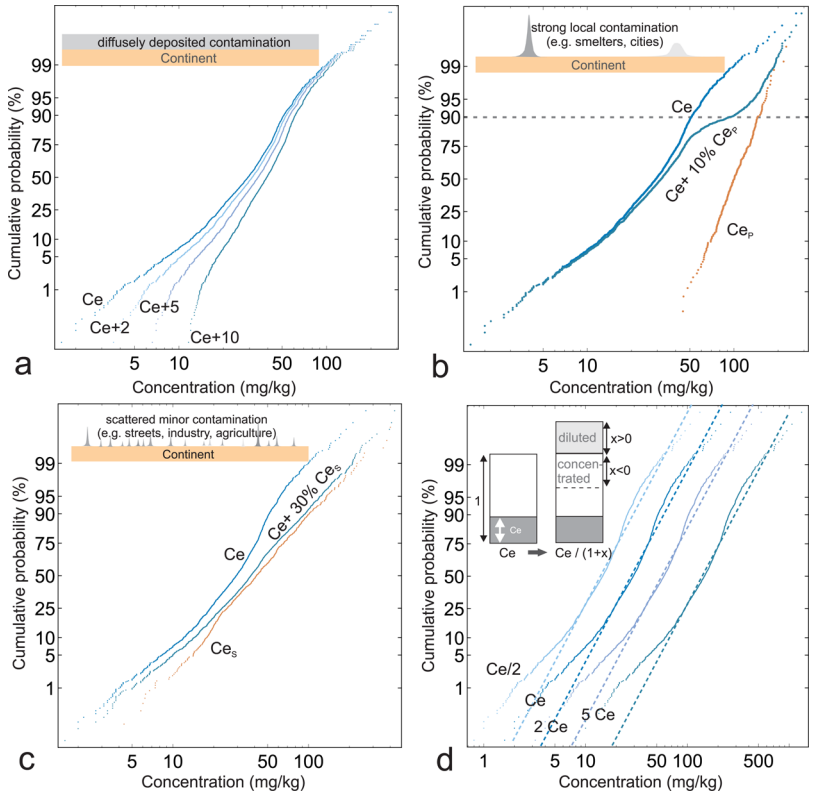


Figure 1. (a) Cumulative frequency distribution of original cerium (Ce) concentrations in European agricultural soil (GEMAS Project[18]) and after modeled addition of 2 (Ce+2), 5 (Ce+5) and 10 (Ce+10) mg/kg Ce to simulate diffuse contamination at the continental scale. (b) Effect of a major local contamination source (Ce_p) characterized by very high element concentrations on the original distribution of an otherwise pristine element (Ce). The Ce + 10%Ce_p curve shows the impact of local contamination on the distribution function. Because the contamination source is local, only a small proportion of all samples (here 10%) will be contaminated with unusually high element concentrations. The effect is a break or bulge in the curve, usually well above the 90th percentile. (c) Effect of a variety of scattered, relatively minor contamination sources (Ce_S) on the overall Ce distribution (here affecting 30% of sites, Ce+30% Ce_S). (d) Dilution or up-concentration, for example due to soil forming processes, shifts the CDF in the CP plot without changing its slope or shape.

poses a more important problem than element toxicity.[16−18] Thus, the low end of the CDF requires more attention and it is shown here that the signal of diffuse contamination characteristically modifies the lower end of the CDF of an element. A prerequisite for the CDF method are a spatially representative, large data sets with preferably >1000 determinations. These are provided by geochemical mapping projects from various geological surveys.

DATA SETS AND METHOD

Data Sets. Five published data sets from different soil geochemical mapping projects at the continental (10 000 000 km^2) to regional (25 000 km^2) scale are used:

1. The continental scale GEochemical Mapping of Agricultural Soil project (GEMAS),[18]
2. The subcontinental scale Baltic Soil Survey project (BSS),[19]
3. The continental scale North American Soil Geochemical Landscapes project (NASGL),[20]
4. The continental scale National Geochemical Survey of Australia project (NGSA),[21,22]
5. The regional scale Nord Trøndelag project (NTR).[23]

Sampling and analytical procedures are described in the cited sources. All data sets contain between 750 and almost 5000 samples. Further requirements for the investigated elements are low detection limits and high analytical precision at the lower end of the data distribution.

CDF Signatures of Contamination and Sample Media. Three different types of contamination will be distinguished: (1) diffuse, (2) strong local to regional, and (3) from scattered local sources. All three processes lead to overabundant concentrations of the studied element which characteristically distort its pristine CDF (Figure 1a−c).

Diffuse contamination results from large scale atmospheric transport and mixing of material from many different sources of local or widespread anthropogenic activities at present and in the past, e.g., industrial and traffic emissions, or the use of fertilizers, including manures, sewage sludges, and other wastes. In contrast, local contamination is deposited in the vicinity of the source and causes a clear, most often exponential (e^{-d}) or power-law (e.g., $1/d^2$) concentration gradient in the vicinity of a point source.[2,4,24,25] Diffuse contamination at the continental scale will not necessarily create such a concentration gradient and should not be expected to result in exceedingly high element concentrations.

How diffuse contamination affects the pristine CDF of an element can be modeled in a simple way. To do this, Figure 1a first shows the distribution of Ce, an element that is to date not significantly influenced by anthropogenic activities. Its CP plot is based on the GEMAS data[18] for European agricultural soil (Ap horizon) with a median of about 30 mg/kg. The other curves in Figure 1a show the effects of adding various amounts

DOI: 10.1021/acs.est.7b00741
Environ. Sci. Technol. 2017, 51, 6719−6726

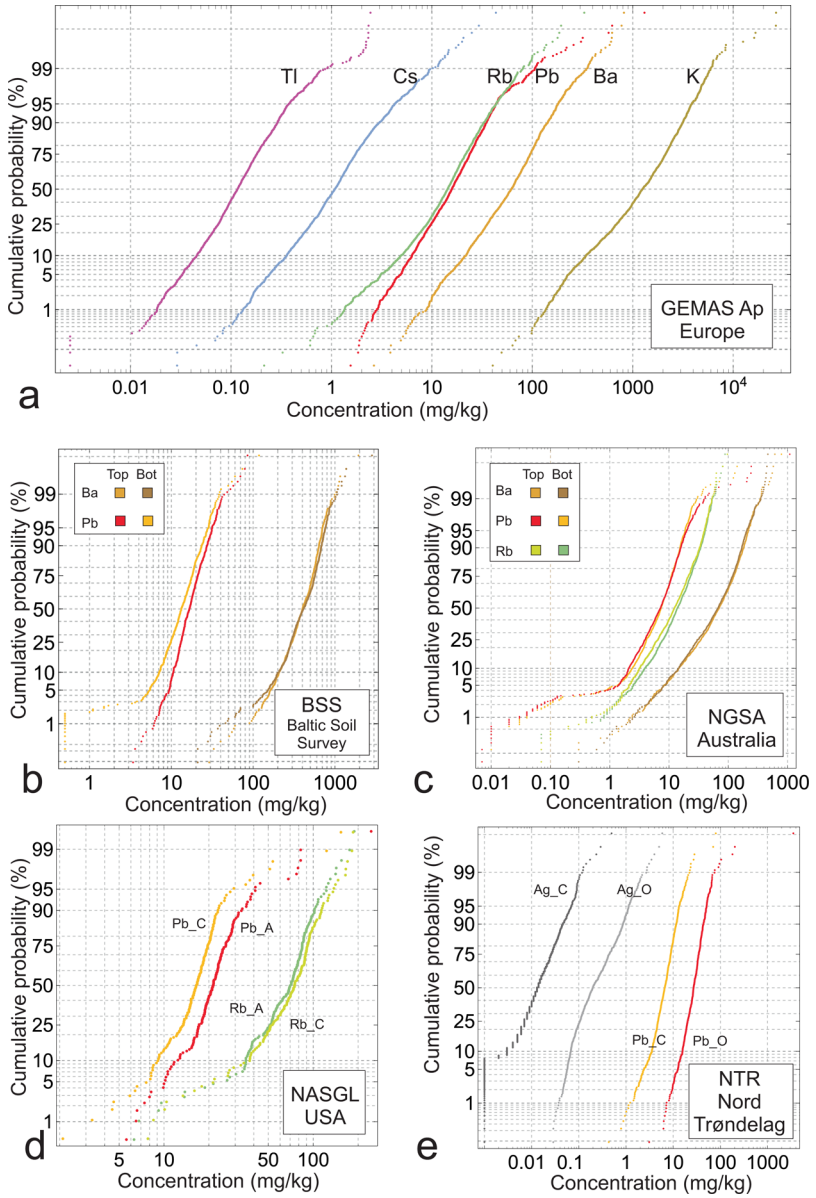


Figure 2. CP diagrams for five large scale geochemical soil projects. (a) Tl, Cs, Rb, Pb, Ba and K for GEMAS Ap.[18] (b) Pb and Ba in an hydrofluoric acid extraction of Top (0−25 cm) and subsoil (Bot, 50−75 cm) from the BSS project.[19] (c) Analytical results of the top (Top; 0−10 cm) and subsoil (Bot; 60−80 cm) samples for Pb, Ba and Rb from the NGSA project.[21,22] (d) Pb and Rb in soil A and C horizon samples collected for the NASGL project.[20] (e) Ag and Pb from the Nord Trøndelag project.[23]

of constant "diffuse contamination" to all data points (2, 5, and 10 mg/kg). They demonstrate that even minor amounts of diffuse contamination (<10% of the median) distort the CP plot noticeably in the region of low natural concentrations, and that the added amount is estimated by the visual intercept of the curve on the abscissa, approximating the zero-CP limit. Such a diffuse addition does not result in unusually high element concentrations but rather in a steepening of the curve in the low-concentration region of the CP plot. Thus, diffuse contamination will have its strongest impact at the lower end of the concentration distribution and not at the high values.

A strong local contamination source will in contrast primarily affect the high concentrations at a local scale. Its effect will become invisible at a distance of some kilometres from the source, and will not affect the whole distribution. The result is a "bulge" in the CDF in the area of the high concentrations (Figure 1b). In extreme cases major contamination sources can of course also dominate a whole region. In such cases the curve in the CP plot shows a distinct break at a much lower percentile. For example the emissions of the Ni and Cu smelters on the Kola Peninsula affect more than 60% of the survey area by substantial atmospheric Ni and Cu deposition.[1] Accordingly, the break in the CDF of the Kola moss data occurs at the 40th percentile in the CP plot.[1]

More difficult to recognize in the CP plot is the impact of many small and local contamination sources that are spatially scattered throughout the survey area, like industry, roads, agriculture, or settlements. Their signal unevenly shifts the pristine curve in the CP plot toward slightly higher values over the whole concentration range (Figure 1c), but also changes the pristine CDF shape by smoothing out breaks due to natural variability.

Linear Concentration Shifts. To estimate contamination based on comparison between different sample media, or

DOI: 10.1021/acs.est.7b00741
Environ. Sci. Technol. 2017, 51, 6719−6726

different elements within the same medium, it is important to consider the effects of element dilution and up-concentration. These occur through soil development processes that mainly add or remove *other* elements, for example decomposition of organic phases, weathering, aeolian winnowing, fluvial, and colluvial deposition and removal, or precipitation of secondary phases. A direct example of element dilution is liming, where through the addition of $CaCO_3$, the Si content is reduced. Figure 1d shows that on a logarithmic scale, they shift the whole element CDF linearly in one direction. In a simple model, the mass y of the element itself is not changed, while other components are added or removed (Figure 1d inset), such that the total amount of material is changed from 1 to $(1+x)$ (for $-1 < x$). The elements concentration thereby changes from y to $y/(1+x)$ leading to a *linear concentration shift* (LCS) of the CDF in the CP diagram. Processes generating LCS are most relevant when studying different media at the same site, for example when comparing different soil horizons. Comparison of Figure 1, parts d to a–c shows that an LCS cannot reasonably be explained by contamination, neither diffuse, nor local or regional, because this would imply that by pure chance the contamination CDF has the same slope and shape as the natural reference CDF.

Thus, the proposed new investigation method can potentially discriminate between several types of contamination. Diffuse contamination is diagnosed on a CP plot by a shift of only the lower concentration range toward higher values resulting in a steeper gradient (Figure 1a). A major point-source contamination will result in a bulge in the CDF at the higher end of the distribution with no effect on the lower end (Figure 1b). A large number of scattered but modest contamination sources will affect the whole distribution curve, with possibly a diagnostic shallowing of the slope (Figure 1c). In the following the focus is on diffuse contamination signals, because they are most controversial and most clearly distinguishable.

Detecting Diffuse Contamination. If the pristine precontamination CDF of an individual element concentration in soil samples were known, the contamination could be determined by deconvolving the contamination distribution by the pristine distribution. Unfortunately the pristine CDF is unknown for elements such as Pb which in Europe, for instance, is influenced by thousands of years of anthropogenic impact. To overcome this problem, two approaches are proposed here.

1. Because a rather large group of elements, e.g., Ba, Cs, K, Rb, and, with limitations, Tl, have a comparable geochemical behavior as Pb[26] and can be assumed to be almost undisturbed by anthropogenic impact, one of their CDFs can serve as a reference shape for the pristine Pb CDF. A prerequisite is that, after some linear shift that accounts for LCS processes, the shapes of the CDFs of reference and contaminated elements are not significantly separated by weathering and soil formation processes. Fortunately, the assumption of similar CDF can be tested by comparison of the CDFs with other elements in the group. Figure 2a shows the CP plots for the above suite of elements in the GEMAS Ap samples.
2. In many geochemical surveys samples from different soil horizons have been collected. In these cases it can be tested on geochemically similar elements whether the deeper horizon, usually less influenced by anthropogenic contamination (if at all), is similar to the top horizon in terms of the shape of the CDF. The curve for Pb in the lower horizon can then be used to estimate the shape of the pristine background CDF of Pb in the top horizon. For example in the Baltic Soil Survey,[19] the NGSA project,[21,22] or the Nord Trøndelag project[23] the CDFs of the subsoil samples, after LCS correction, can be used to represent the uncontaminated CDF of top-soil Pb.

In the following, one or both approaches will be applied to published geochemical data sets to estimate the amount of diffuse contamination with Pb, one of the most critical PTEs. Diffuse contamination by any other element can be estimated analogously.

Quantitative Estimates. To obtain objective and quantitative estimates of diffuse Pb contamination the two CDFs, that of Pb and that of a reference element R, are compared. In the simplest case the natural background of Pb is obtained by shifting the reference distribution of R by a constant factor a (LCS), and diffuse contamination adds a constant concentration b, such that the CDF of $aR + b$ becomes statistically indistinguishable from the CDF of Pb. This statistical similarity can be tested by either a two sample Kolmogorov–Smirnov test,[27,28] or better by a two sample Cramér-von-Mises (C-vM) test.[29,30] One approach to estimate a and b is to maximize the p-value of the C-vM test as a function of a and b. It turns out that the diffuse Pb concentrations b thus derived are well below the median of the original elemental concentration as in Figure 1a, while local and regional effects corresponding to the scenarios in Figure 1b,c still affect the CDF at higher concentrations. To better target the lower concentration percentiles, a second method is devised to overlay the CDFs of Pb and $aR + b$. This is done by listing the same percentiles for both data sets Pb and R, whereby the percentiles are chosen to put more weight on lower concentrations and to disregard high concentrations. Besides this, the details of the choice of the percentiles are not very influential. In the following the values $\mathbf{Q} = (0.01, 0.02, 0.03, 0.05, 0.1, 0.2, 0.3, 0.5, 0.75)$ are chosen, which result in the percentile lists $\mathbf{p} = (p_{0.01}, ..., p_{0.75})$ for Pb, and $\mathbf{r} = (r_{0.01}, ..., r_{0.75})$ for R. Because for $a, b \geq 0$ the corresponding percentile list for $a\ R + b$ is

$$a\mathbf{r} + \mathbf{b} = a(r_{0.01}, ... r_{0.75}) + (b, b, ..., b) = (ar_{0.01} + b, ..., ar_{0.75} + b)$$

the optimal fit values for a, b can be found by a non-negative least-square fit minimizing $\|\mathbf{p} - a\ \mathbf{r} + \mathbf{b}\|^2$, which is performed using a standard algorithm.[31]

Three different methods will thus be used to estimate the amount of "diffuse contamination" for each of the case studies presented below: (1) The amount of diffuse contamination can be calculated by *percentile fitting* which minimizes $\|\mathbf{p} - a\ \mathbf{r} + \mathbf{b}\|^2$ for the percentile list $\mathbf{Q}$. This results in two fit parameters a_Q, b_Q, and the corresponding p-value p_Q of the C-vM test for these parameters. (2) Maximizing the p-value of the C-vM test provides the C-vM fit a_C, b_C between the two full empirical distributions with optimal p-value p_C. (3) A *visual fit* provides the amount b_V of diffuse contamination by a visual estimate based on the CP diagram when the two CDFs are overlain. Although this method has limited precision and is more subjective it may result in more accurate values because it is possible to abstract from individual irregularities of the CDFs due to measurement uncertainties at low concentrations or disturbances in the CDFs from mineralization or local contamination.

DOI: 10.1021/acs.est.7b00741
Environ. Sci. Technol. 2017, 51, 6719–6726

Table 1. Results for Three Different Approaches to Estimate the Input of Diffuse Contamination Based on the CDF of Pb in Top and Sub-Soil Samples or the CDF of Pb and an Element with Geochemical Similar Behaviour[a]

project	Pb	R	Pb_M	a_Q	b_Q	p_Q	a_C	b_C	p_C	b_V	b/Pb_M
GEMAS	Pb	Rb	**16**	1.01	**1.4**	0.37	1.04	**0.9**	0.72	**1−2**	6−12%
BSS	Pb_{top}	Pb_{bot}	**16.6**	0.84	**5.3**	0.13	0.99	**3.3**	0.84	**3−3.5**	20−32%
		Ba_{top}		0.033	**3.5**	0.21	0.033	**3.4**	0.23	**3**	18−21%
NGSA	Pb_{TOS}	Pb_{BOS}	**7.2**	0.99	**0.0**	0.18	0.98	**0.0**	0.23	**<0. 5**	<5%
NASGL	Pb_A	Pb_C	**21**	1.22	**1.5**	0.77	1.2	**1.5**	0.88	**1. 4**	7%
		Ba_A		0.032	**4.9**	0.35	0.033	**4.9**	0.38	**1.7**	8−23%
		Rb_A		0.25	**4.5**	0.55	0.24	**4.5**	0.62	**1.7**	8−21%
NTR	Pb_O	Pb_C	**27**	3.55	**3.5**	0.78	3.57	**3.5**	0.81	**2**	8−13%

[a]See text for detailed explanation of the parameters a_Q, a_C, a_V, b_Q, b_C, b_V, b_Q, b_C, and b_V are the diffuse contamination estimates using the three suggested methods. For comparison the median Pb concentration Pb_M in the sample material at top for the project is also provided. All bold numbers are in mg/kg.

An important indicator of the usefulness of the first two quantitative fits are their p-values. A necessary condition for a reference parameter to be usable is that its CDF can be fit to the original parameter in a reasonable way. This condition can be quantitatively tested by requesting that the CDF after fitting the reference distribution through the parameters a,b is statistically indistinguishable from the original CDF at a certain significance level of the C-vM test. Typically a p-value of 0.05 is chosen, indicating that the null-hypothesis that both data sets are drawn from the same probability distribution cannot be rejected at a 5% probability level.

RESULTS AND DISCUSSION

The estimates of diffuse contamination from the three methods obtained for the different case studies are summarized in Table 1 and will be discussed in the following subsections. In Table 1 the target is always the Pb concentration of the topmost soil layer, and diffuse contamination is estimated with respect to the reference element R from either the same or a deeper soil horizon. Only reference elements that are chemically comparable and yield a final significance of at least $p = 0.05$ are chosen.

Case Study 1: The Continental Scale GEMAS Project. Figure 2a shows the CDFs of Ba, Cs, K, Pb, and Tl in the GEMAS Ap samples.[18] Disregarding the individual LCS due to different median concentrations, the five elements display remarkably similar shapes of their CDFs. When comparing the CDFs of Pb and Rb, which have very similar concentration, one observes clear deviations in the Pb distribution at high (>50 ppm) and at low (<10 ppm) concentrations. The upper deviation indicates unusually high Pb in about 3% of the samples. It resembles Figure 1b and probably reflects the combined effects of mineralization (Pb ore deposits) and local anthropogenic contamination.[32] The deviation at low concentrations consists of a slight shift of the Pb curve toward higher concentrations and closely resembles the predicted signal from diffuse contamination in Figure 1a. Visually, this shift suggests a diffuse deposition of $b_V = 1-2$ mg/kg Pb at the European scale. When quantitatively fitting the two curves in the lower quantiles, a diffuse deposition of $b_Q = 1.4$ mg/kg Pb is recovered. A full fit of the two CDFs using the C-vM test yields only $b_C = 0.9$ mg/kg Pb, but inspection reveals that in order to improve the CDF correspondence at high concentrations the lower percentiles fit worse than in the lower-quantile fit. At first glance, even $b_Q = 1.4$ mg/kg may appear to be a low figure, however, compared to the median Pb concentration $Pb_M = 16$ mg/kg in the GEMAS Ap samples this is a 9% addition. In a suggested "90% contamination" scenario[33] the Pb CDF in the CP-diagram would become extremely steep. A spatial distribution map of Pb would then be completely unrelated to geology. Observations from the GEMAS Pb and Pb-isotope maps, however, show a very strong relation to geology[32] besides some few local hotspots of contamination, e.g., Paris and London.

Case Study 2: The Subcontinental Scale BSS Project. Direct comparison of Pb in top and subsoil samples from BSS, and comparison with the chemically comparable element Ba in the same medium in Figure 2b demonstrate that two different reference CDFs yield similar estimates of diffuse contamination. As discussed in the literature,[6,7] when comparing top and subsoil samples, the intrinsic LCS due to relative dilution or enhancement by the variation of all other element concentrations has often been ignored, although it is intrinsic to the rationale of using different sample media. Neglecting material differences between more organic top soil and minerogenic subsoil, and the affinity of many metals to organic matter,[34] inevitably leads to inaccurate quantitative estimates of the contamination signal. Because also element ratios are affected, this mistake is implicit—though unfortunately well hidden—in the definition of enrichment factors as originally introduced,[35,36] which are still widely used. For two reasons, this problem is not substantial in the BSS data set. The first reason is accidental in that the two soil horizons studied are both from agricultural and predominantly minerogenic soil. They practically have identical loss on ignition values[19] and can thus be directly compared. The visible differences in the CP plots of Pb in top and subsoil (Figure 2b) mainly resemble Figure 1a with a diffuse deposition contribution of $b_V = 3-3.5$ mg/kg Pb at the northern European scale. The second reason is that all proposed estimates of diffuse contamination b take into account a dilution or enhancement through the factor a. This reason is more profound and applies to other sample sets, where the media are more distinct.

The estimate $b_Q = 5.3$ mg/kg in Table 1 is higher than the visual estimate and leads to a relatively bad overall fit as indicated by the low $p_Q = 0.13$. However, the very good overall fit $b_C = 3.3$ with $p_C = 0.84$ agrees better with the visual impression. The origin of this difference is that the subsoil Pb_{bot} contains >1% of determinations with Pb concentration below the detection limit, while the CDF for Pb_{top} shows few samples with concentration of <4 mg/kg. In spite of this problem the visual and the C-vM estimates coincide well with the continental scale estimate based on the GEMAS data, while

DOI: 10.1021/acs.est.7b00741
Environ. Sci. Technol. 2017, 51, 6719−6726

the apparently too high quantitative fits still yield the right order of magnitude.

From the group of geochemically similar reference elements in Figure 2a, for the BSS samples only Ba in the strong extraction (HF) has sufficient quality and a sufficiently low detection limit. The CP plots for Pb and Ba show similar CDF shapes in the subsoil, while the topsoil shapes are steeper, reflecting a narrower concentration range. That the Ba curve also varies slightly in shape between subsoil and topsoil samples may be due to liberation of Ba through weathering, regional input of wind-blown dust, or possibly even anthropogenic Ba deposition in connection with K-fertilizers. Comparison of the shifted topsoil Ba curve with the topsoil Pb curve indicates increased abundance of Pb in about 2% of the samples and again a diffuse deposition signal of 3−3.5 mg/kg Pb. This amount is slightly larger than the independently determined value for the GEMAS Ap soil. Both methods show that the impact of diffuse contamination on the overall Pb distribution is in the 20% range of the median Pb concentration reported for the BSS project.

Case Study 3: The Continental Scale NGSA Project. The NGSA project[21,22] besides topsoil, also sampled deep subsoil, where no contamination effect is expected. Due to the mostly warm and dry climate, the Australian topsoil generally is not substantially more organic than the subsoil, and these two media should be more directly comparable than in Europe. Because analyses of Pb in both top and subsoil samples as well as analyses of Rb and Ba in an aqua regia extraction of the samples are available, both reference methods to estimate diffuse contamination can be applied in parallel (see Figure 2c and Table 1 for results). The results confirm that the CDFs of top soil and subsoil in the CP plots of the Australian samples are extremely similar. Unusually high Pb concentrations are observed for the highest 5% of the topsoil samples, probably related to regional mineralization, mining, or local contamination.[37] When shifting the topsoil CDF to match the subsoil CDF within the low 10−40 percentile range, only a very small diffuse deposition of less than 0.5 mg/kg Pb can be inferred. This is in agreement with the fact that 4% of the topsoil samples contain ≪1 mg/kg Pb in total and therefore cannot have been exposed to diffuse deposition of more than 1 mg/kg Pb. More interesting may be the slight shift displayed at the upper end of the CDF, pointing at a certain amount of local point-source contamination. All other methods deliver very comparable estimates of essentially no diffuse contamination. The comparison with the CDFs for Ba and Rb (not shown) do also not deliver any indication of measurable diffuse contamination at the continental scale in Australia.[37]

Case Study 4: The Continental Scale NASGL Project. The samples for the North American Soil Geochemical Landscapes (NASGL) project[20] were taken in relation to pedogenically determined soil horizons instead of at fixed depth intervals. Its topsoil samples from the A horizon are substantially more organic than the minerogenic C horizon soil which induces a pronounced LCS (Figure 2d), while the characteristic CDF shapes are remarkably similar. As pointed out above, such a shift cannot be explained by anthropogenic input because the amount of added mass of contaminant would have to be proportional to the naturally available element concentration at each site, clearly an absurd assumption, especially in light of the geologically related spatial patterns of the Pb distribution,[20] and the fact that also elements with negligible contamination, like Rb in Figure 2d, display a LCS.

With Pb_C as a reference, both quantile fit and C-vM fit yield b = 1.5 mg/kg which agrees with the visual estimate (Table 1). The large LCS implies that statistically the A horizon has 20% higher Pb concentrations than the C horizon. Deviations in shape between the shifted CDFs of Pb_A and Pb_C occur at high concentrations for about 15% of the samples, indicating substantial Pb contamination by local anthropogenic sources and the existence of mineralization. A bulge in the CDF is also observed for about 5% of the samples at lower concentrations which show an increased concentration of about 5 mg/kg Pb, probably effected by smaller local Pb sources.

Repeating the quantile and C-vM fits for the reference elements Rb and Ba in the A horizon results in higher estimates of b = 4.5−4.9 mg/kg which have lower p-values and deviate from the lower visual estimates of b_V = 1.7 (Table 1). This indicates that the CDF comparison of different elements in the same horizon is less reliable than the comparison of curves displayed for the same element in different soil horizons, as long as the LCS assumption holds. Altogether, diffusely deposited Pb contamination is between 10 and 20% of the median Pb concentration for the North American soil, probably with a tendency toward the lower estimate.

Case Study 5: The Regional Scale NTR Project. To test the CDF methods at a regional scale, the geochemical mapping project of Nord Trøndelag is chosen. It provides forest soil O- and C-horizon samples collected over an area of about 25 000 km^2 in central Norway,[23] where no single major Pb emission source is known and anthropogenic Pb should predominantly originate from diffuse contamination. As in the NASGL project, LCS between the soil horizons is substantial, due to a highly organic O horizon. Because the C horizon is the geological substrate for the O horizon, it has been argued that element concentrations in top and subsoil at the same site can be compared either directly, or via ratios or double ratios as they occur in enrichment factors,[38] that are based on the popular but apparently false[6−8] assumption of invariant element ratios across the C and O horizons. One of the strengths of CP plots is that they allow the differentiation between CDF changes resulting from the true addition of material and proportional LCS. Figure 2e demonstrates that in Nord Trøndelag the difference between the CDFs of Pb in C and O horizons is dominantly due to LCS. This is reinforced by the similar or even larger LCSs observed in other elements like Ag shown in Figure 2e. As in the previous NASGL project, this interpretation is also strongly supported by the very similar CDF shapes after LCS correction confirmed by the high p-values for NTR in Table 1.

The optimal quantile and C-vM fits between the soil horizons both yield $a \approx 3.6$ and b = 3.5 mg/kg, whereas visual inspection suggests a somewhat lower $b_v \approx 2$ mg/kg. Overall, the contribution of diffuse Pb contamination to the O horizon is 8−13% of the median Pb concentration.

Statistical Soil Monitoring. Comparing CDFs from soil surveys carried out at different times will be advantageous in soil monitoring programs, provided that laboratory detection limits are sufficiently low, and analytical quality sufficiently high to detect subtle changes at the low-concentration end of the CDF. If this is the case, then the statistical monitoring has the huge advantage that comparison of two surveys does not require the same individual sample locations, as long as sample medium, density, and distribution are sufficiently similar.

In terms of environmental impact of contamination a focus on the local rather than the regional to continental scale is

DOI: 10.1021/acs.est.7b00741
Environ. Sci. Technol. 2017, 51, 6719−6726

required. Yet, even low or median levels of anthropogenic contributions, depending on grain size and chemical mobility, or organic availability of the compounds, may have a greater influence on organisms than a high natural background. The present focus of environmental science on high concentrations is thus largely misguided. The proposed CDF-based method requires high data quality and low detection limits. The latter are still problematic because, due to the intuitive focus on high concentrations, very often the precision of measurements and reports is insufficient for low concentrations.

Overabundant Concentrations Mark Contamination. The term *overabundant concentration* denotes the concept that values in some concentration range due to contamination occur more often than in the pristine CDF. For example if the Pb concentration 20 mg/kg occurs too often, such that one finds a bulge in the CDF that is not present in the LCS-corrected reference CDF this indicates a frequent occurrence of local contamination adding <20 mg/kg Pb. To characterize contamination, overabundant element concentrations are at least equally important as high concentrations which may be caused by both, mineralization or contamination. Overabundant concentrations are not easily detected by classical enrichment factors due to the influence of LCS, and because local comparison of the contamination element and the reference can be randomly distorted for example by the different mixing length-scales of different soil horizons.

Diffuse soil contamination at the continental scale, although clearly visible in the CDFs, presently does not generate exceedingly high PTE concentrations. It remains invisible in continental scale geochemical maps because the natural variation is still considerably larger and dominates spatial patterns of high concentrations. This solves the puzzle that in spite of substantial anthropogenic emissions, geochemical maps of Pb at the continental scale consistently reflect natural conditions, usually both geology and climate. Pb contamination therefore is rarely marked by high total Pb concentration, but rather by overabundant lower and medium concentrations of Pb. It is of utmost importance to develop and apply geochemical tools that reliably detect and evaluate these overabundant concentrations.

AUTHOR INFORMATION

Corresponding Author

*Phone: +47 7390 4203; e-mail: karl.fabian@ngu.no (K.F.).

ORCID

Karl Fabian: 0000-0002-3504-3292

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The helpful comments of four reviewers are thankfully acknowledged.

REFERENCES

(1) Reimann, C. et al. *Environmental Geochemical Atlas of the Central Barents Region*; NGU GTK CKE special publication; Geological Survey of Norway: Trondheim, Norway, 1998; p 745.

(2) Bonham-Carter, G.; Henderson, P.; Kliza, D.; Kettles, I. Comparison of metal distributions in snow, peat, lakes and humus around a copper smelter in western Quebec, Canada. *Geochem.: Explor., Environ., Anal.* **2006**, *6*, 215−228.

(3) Commission of the European Communities. *Thematic Strategy for Soil Protection*; 2006; Vol. COM(2006)231; p 12.

(4) Caritat, P. de; Reimann, C.; Chekushin, V.; Bogatyrev, I.; Niskavaara, H.; Braun, J. Mass balance between emission and deposition of airborne contaminants. *Environ. Sci. Technol.* **1997**, *31*, 2966−2972.

(5) Henderson, P.; McMartin, I.; Hall, G.; Percival, J.; Walker, D. The chemical and physical characteristics of heavy metals in humus and till in the vicinity of the base metal smelter at Flin Flon, Manitoba. *Environ. Geol.* **1998**, *34*, 39−58.

(6) Reimann, C.; Caritat, P. de Intrinsic flaws of element enrichment factors (EFs) in environmental geochemistry. *Environ. Sci. Technol.* **2000**, *34*, 5084−5091.

(7) Reimann, C.; Caritat, P. de Distinguishing between natural and anthropogenic sources for elements in the environment: regional geochemical surveys versus enrichment factors. *Sci. Total Environ.* **2005**, *337*, 91−107.

(8) Sucharova, J.; Suchara, I.; Hola, M.; Marikova, S.; Reimann, C.; Boyd, R.; Filzmoser, P.; Englmaier, P. Top-/Bottom-soil ratios and enrichment factors: What do they really show? *Appl. Geochem.* **2012**, *27*, 138−145.

(9) Tennant, C.; White, M. Study of the distribution of some geochemical data. *Econ. Geol. Bull. Soc. Econ. Geol.* **1959**, *54*, 1281−1290.

(10) Sinclair, A. Selection of threshold values in geochemical data using probability graphs. *J. Geochem. Explor.* **1974**, *3*, 129−149.

(11) Sinclair, A. *Application of Probability Graphs in Mineral Exploration*; Association of Exploration Geochemists: Toronto, 1976; Vol. *4*; p 95.

(12) Sinclair, A. In *Exploration Geochemistry: Design and Interpretation of Soil Surveys*; Fletcher, W., Hoffman, S., Mehrtens, A., Sinclair, V., Thompson, I., Eds.; Reviews in Economic Geology; Society of Economic Geologists: Chelsea, MI, USA, 1986; Vol. *3*; Chapter Statistical interpretation of soil geochemical data, pp 97−115.

(13) Sinclair, A. A fundamental approach to threshold estimation in exploration geochemistry: probability plots revisited. *J. Geochem. Explor.* **1991**, *41*, 1−22.

(14) Davies, B. A graphical estimation of the normal lead content of some British soils. *Geoderma* **1983**, *29*, 67−75.

(15) Reimann, C.; Filzmoser, P.; Garrett, R.; Dutter, R. Statistical Data Analysis Explained. *Applied Environmental Statistics with R. Wiley, Chichester* **2008**, 343.

(16) *Communication from the European Commission to the European Council and the European Parliament: 20 20 by 2020: Europe's Climate Change Opportunity*; European Commission: Brussels, Belgium, 2008.

(17) Plant, J.; Smith, D.; Smith, B.; Williams, L. Environmental geochemistry at the global scale. *Appl. Geochem.* **2001**, *16*, 1291−1308.

(18) Reimann, C.; Birke, M.; Demetriades, A.; Filzmoser, P.; O'Connor, P. *Chemistry of Europe's Agricultural Soils, Part A: Methodology and Interpretation of the GEMAS Data Set*; Geologisches Jahrbuch Reihe B; Schweizerbart Science Publishers: Stuttgart, 2014; Vol. *B 102*; p 523.

(19) Reimann, C.; Siewers, U.; Tarvainen, T.; Bityukova, L.; Eriksson, J.; Gilucis, A.; Gregorauskiene, V.; Lukashev, V.; Matinian, N.; Pasieczna, A. *Agricultural Soils in Northern Europe: A Geochemical Atlas*; Geologisches Jahrbuch, Sonderhefte, Reihe D; 2003; Vol. *5*p 279.

(20) Smith, D.; Cannon, W.; Woodruff, L.; Solano, F.; Ellefsen, K. *Geochemical and Mineralogical Maps for Soils of the Conterminous United States*; Open-File Report; U.S. Geological Survey: U.S., 2014; Vol. *2014−1082*; p 386.

(21) Caritat, P. de; Cooper, M. *National Geochemical Survey of Australia: the Geochemical Atlas of Australia (Volume 1)*; Geoscience Australia: Canberra, Australia, 2011; pp 1−268; Available at: http://www.ga.gov.au/corporate_data/71971/Rec2011_021_Vol1.pdf.

(22) Caritat, P. de; Cooper, M. *National Geochemical Survey of Australia: The Geochemical Atlas of Australia (Volume 2)*; Geoscience Australia: Canberra, Australia, 2011; p 269−478, Available at: http://www.ga.gov.au/corporate_data/71971/Rec2011_021_Vol2.pdf.

(23) Reimann, C.; Fabian, K.; Schilling, J.; Roberts, D.; Englmaier, P. A strong enrichment of potentially toxic elements (PTEs) in Nord

DOI: 10.1021/acs.est.7b00741
Environ. Sci. Technol. 2017, 51, 6719−6726

Trøndelag (central Norway) forest soil. *Sci. Total Environ.* **2015**, *536*, 130−141.
(24) McMartin, I.; Henderson, P.; Nielsen, E. Impact of a base metal smelter on the geochemistry of soils of the Flin Flon region, Manitoba and Saskatchewan. *Can. J. Earth Sci.* **1999**, *36*, 141−160.
(25) McMartin, I.; Henderson, P.; Plouffe, A.; Knight, R. Comparison of Cu-Hg-Ni-Pb concentrations in soils adjacent to anthropogenic point sources: examples from four Canadian sites. *Geochem.: Explor., Environ., Anal.* **2002**, *2*, 57−73.
(26) Goldschmidt, V. *Geochemistry*; Clarendon Press: Oxford, United Kingdom, 1999; p 730.
(27) Kolmogorov, A. N. Sulla determinizione empirica di una legge di distribuzione. *Giorn. Ist. Ital. Attuari* **1933**, *4*, 83−91.
(28) Smirnov, N. V. On estimating the discrepancy between empirical distribiution curves for two independent samples. *Byull. Moskov. Gos. Univ. Ser. A* **1933**, 3−14. (In Russian).
(29) Cramér, H. On the Composition of Elementary Errors. *Scandinavian Actuarial Journal* **1928**, *1*, 13−74.
(30) von Mises, R. *Wahrscheinlichkeit, Statistik und Wahrheit*; Julius Springer Verlag: Wien, 1928; p 192.
(31) Lawson, C. L.; Hanson, R. J. *Solving Least Squares Problems*; Prentice Hall: New York, 1974; p 337.
(32) Reimann, C.; Flem, B.; Fabian, K.; Birke, M.; Ladenberger, A.; Négrel, P.; Demetriades, A.; Hoogewerff, J.; Team, T. G. P. Lead and lead isotopes in agricultural soils of Europe - The continental perspective. *Appl. Geochem.* **2012**, *27*, 532−542.
(33) Nriagu, J. Global inventory of natural and anthropogenic emissions of trace metals to the atmosphere. *Nature* **1979**, *279*, 409−411.
(34) Goldschmidt, V. M. The principles of distribution of chemical elements in minerals and rocks. *J. Chem. Soc.* **1937**, 655−673.
(35) Zoller, W.; Gladney, E.; Duce, R. Atmospheric concentrations and sources of trace metals at the South Pole. *Science* **1974**, *183*, 198−200.
(36) Duce, R.; Hoffman, G.; Zoller, W. Atmospheric trace metals at remote northern and southern hemisphere sites: pollution or natural? *Science* **1975**, *187*, 59−61.
(37) Reimann, C.; Caritat, P. Establishing geochemical background variation and threshold values for 59 elements in Australian surface soil. *Sci. Total Environ.* **2017**, *578*, 633−648.
(38) Blaser, P.; Zimmermann, S.; Luster, J.; Shotyk, W. Critical examination of trace element enrichments and depletions in soils: As, Cr, Cu, Ni, Pb, and Zn in Swiss forest soils. *Sci. Total Environ.* **2000**, *249*, 257−280.

DOI: 10.1021/acs.est.7b00741
Environ. Sci. Technol. 2017, 51, 6719−6726