

GEMAS: Establishing geochemical background and threshold for 53 chemical elements in European agricultural soil



Clemens Reimann ^{a, *}, Karl Fabian ^a, Manfred Birke ^b, Peter Filzmoser ^c, Alecos Demetriades ^d, Philippe Négrel ^e, Koen Oorts ^f, Jörg Matschullat ^g, Patrice de Caritat ^h, The GEMAS Project Team

^a Geological Survey of Norway (NGU), PO Box 6315 Sluppen, 7491 Trondheim, Norway

^b Federal Institute for Geosciences and Natural Resources (BGR), Stilleweg 2, D-30655 Hannover, Germany

^c Institute of Statistics & Mathematical Methods in Economics, Vienna University of Technology, Wiedner Hauptstr. 8-10, A-1040 Vienna, Austria

^d Institute of Geology and Mineral Exploration (IGME), 1 Spirou Louis St., Olympic Village, Acharnae, 13677 Athens, Hellas

^e Geological Survey of France (BRGM), Laboratories Division, Orléans, France

^f Arche Consulting, Liefkensstraat 35D, 9032 Wondelgem, Belgium

^g Interdisciplinary Environmental Research Centre (IÖZ), TU Bergakademie Freiberg, Brennhausgasse 14, 09599 Freiberg, Germany

^h Research School of Earth Sciences, The Australian National University, Canberra ACT 2601, Australia

ARTICLE INFO

Article history:

Received 7 November 2016

Received in revised form

24 January 2017

Accepted 30 January 2017

Available online 16 February 2017

Editorial handling by Dr. R. Fuge.

Keywords:

Background variation
Geochemical threshold
Agricultural soil
Continental scale
Europe

ABSTRACT

The GEMAS (geochemical mapping of agricultural soil) project collected 2108 Ap horizon soil samples from regularly ploughed fields in 33 European countries, covering 5.6 million km². The <2 mm fraction of these samples was analysed for 53 elements by ICP-MS and ICP-AES, following a HNO₃/HCl/H₂O (modified aqua regia) digestion. Results are used here to establish the geochemical background variation and threshold values, derived statistically from the data set, in order to identify unusually high element concentrations for these elements in the Ap samples. Potentially toxic elements (PTEs), namely Ag, B, As, Ba, Bi, Cd, Co, Cr, Cu, Hg, Mn, Mo, Ni, Pb, Sb, Se, Sn, U, V and Zn, and emerging 'high-tech' critical elements (HTCEs), i.e., lanthanides (e.g., Ce, La), Be, Ga, Ge, In, Li and Tl, are of particular interest. For the latter, neither geochemical background nor threshold at the European scale has been established before. Large differences in the spatial distribution of many elements are observed between northern and southern Europe. It was thus necessary to establish three different sets of geochemical threshold values, one for the whole of Europe, a second for northern and a third for southern Europe. These values were then compared to existing soil guideline values for (eco)toxicological effects of these elements, as defined by various European authorities. The regional sample distribution with concentrations above the threshold values is studied, based on the GEMAS data set, following different methods of determination. Occasionally local contamination sources (e.g., cities, metal smelters, power plants, agriculture) can be identified. No indications could be detected at the continental scale for a significant impact of diffuse contamination on the regional distribution of element concentrations in the European agricultural soil samples. At this European scale, the variation in the natural background concentration of all investigated elements in the agricultural soil samples is much larger than any anthropogenic impact.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

To identify areas with unusually high (or low) concentrations of potentially toxic elements (PTEs) or 'high-tech' critical elements (HTCEs) in Europe, geochemical background and threshold values

need to be established at the continental scale. Note that many different definitions of the terms 'background' and 'threshold' are in use (see discussion below). Geochemical threshold values are most often used to identify locations with unusually high element concentrations. A lower threshold could, of course, be defined for the lower end of the data distribution to identify areas with unusually low element concentrations. In fact, Reimann et al. (2014a) argue that element deficiency in European agricultural soil is a spatially much more widespread problem than toxicity.

* Corresponding author.

E-mail address: clemens.reimann@ngu.no (C. Reimann).

High element concentrations in soil may be due to the occurrence of mineralisation, unusual rock types, like serpentinite, black shale or alkaline intrusions, or may be caused by human activities. Depending on bioavailability, these concentrations may present environmental risks due to element toxicity. Proper risk assessment of soil entails a comparison of exposure concentrations with effect thresholds for environmental and human health derived from (eco)toxicological data. This approach preferentially takes into account the effect of abiotic soil properties on bioavailability and toxicity of the element – see Smolders et al. (2009), Oorts and Schoeters (2014), Oorts et al. (2016) or Birke et al. (2016) for methods and examples.

Sufficient (eco)toxicological knowledge required for derivation of such effect-based thresholds is, however, presently only available for a limited number of elements (Ag, As, Ba, Cd, Co, Cr, Cu, Hg, Mo, Ni, Pb, Sb, V, and Zn) and ecotoxicological information allowing a site-specific risk assessment is even restricted to fewer elements (Ag, Co, Cu, Mo, Ni, Pb and Zn). Furthermore, this site-specific approach requires a substantial amount of additional data (e.g., pH, grain size, cation exchange capacity, total organic carbon) being available for each site. For most HTCEs basic (eco)toxicological knowledge to derive generic threshold values for (eco)toxicological risk assessment is not available. A straightforward approach to identify a geochemical (non-toxicological) threshold, simply defined as the value above which the concentration of an element in a given data set is ‘unusually high’, may thus have merits in a first screening for sites that may require attention.

Unusually high element concentrations can be due to anthropogenic activities, such as urbanisation, industrial activities, mining, and agricultural practices. They may also be of natural origin and indicate areas with geochemically unusual rock types (e.g., serpentinite, alkaline intrusions or black shale) or areas having a high potential for the occurrence of mineral deposits of these or related elements (e.g., Reimann et al., 2016). The separation of these three different causes of high element levels in soil requires substantial expert knowledge about, e.g., the location of possible contamination sources (cities, metal smelters, power plants, industry), climate, vegetation zones, geology, element dispersion processes and mineral deposits of Europe.

Reimann and Garrett (2005) showed that the terms *background* and *threshold* are rather ill-defined and their meaning can differ widely between different fields of science. The use of these terms in applied geochemistry is originally derived from exploration geochemistry (e.g., Hawkes, 1957; Hawkes and Webb, 1962), with the intention of obtaining a reasonable number of cases for follow-up exploration surveys in order to locate potential mineralisation targets. At that time, the background was defined to cover the variation of the Mean \pm 2 Standard Deviations and is thus a range. The upper threshold was defined as the exact number of the Mean + 2 Standard Deviations. This approach resulted, if the element concentrations were normally distributed (which is rarely the case when dealing with geochemical data – Reimann and Filzmoser, 2000), in a manageable 2.3% of all values that had to be followed up in the hope of locating mineralisation.

In environmental investigations, a threshold is often defined by its protection goal (e.g., accepted effect level for specific organisms or human health) and different thresholds can be derived for the same element in soil. Examples of a set of environmental threshold values can be found in Provoost et al. (2006), Carlon et al. (2007), Ministry of the Environment of Finland (2007), Department of Environment and Conservation of Western Australia (DEC, 2010) or on the EPA website (<https://www.epa.gov/chemical-research/interim-ecological-soil-screening-level-documents>). An exhaustive list of the different ‘action levels’ or ‘Soil Guideline Values’ (SGVs), as defined for a variety of elements in soil by different

European countries, can be found in Carlon et al. (2007) or, more recently, in Reimann et al. (2014b – Appendix A).

The best method of defining the geochemical background and threshold is a matter of much discussion in the literature (e.g., Chaffee, 1977; Matschullat et al., 2000; Reimann and Garrett, 2005; Reimann et al., 2005; Reimann, 2007; Jarva et al., 2010; Cave et al., 2012; Johnson et al., 2012; Ander et al., 2013; Mali et al., 2015; Rothwell and Cooke, 2015; Reimann and Caritat, 2017). Here it will be discussed whether single geochemical threshold values, based on the distribution of background concentrations, can be established and used for identifying sites at risk, areas containing geochemically unusual rock types, or areas of mineral potential for the investigated elements at the continental scale.

Geochemical threshold values for 53 elements will be estimated using a variety of methods suggested in the literature. Results will be displayed on maps of the exceedances of threshold values for some of the most important PTEs in Europe. The method best suited to calculating general threshold values for a continent will be discussed. Reasons for the exceedance of a threshold value will also be discussed. It will be shown that while geochemical threshold values can be used to identify sites with ‘unusually high element concentrations’ at a variety of scales, these cannot be used to separate anthropogenic from geogenic element sources. Rather, areas in Europe will be indicated that may require follow-up surveys or denote ‘more work to be done’. Geochemical threshold values are also compared with (eco)toxicologically based soil guideline values to evaluate if they can be used as a protective value for (eco)toxicological effects.

2. Methods

2.1. The GEMAS project

GEMAS (geochemical mapping of agricultural soil) is a cooperation project between the Geochemistry Expert Group of Euro-GeoSurveys (EGS) and Eurometaux (Reimann et al., 2014a,b). The GEMAS project was aimed at producing consistent soil geochemical data at the continental scale in accordance with the REACH regulation requirements (EC, 2006). REACH specifies that industry must prove that it can produce and handle its substances safely. Risks from exposure to a substance during its production and use at local, regional and European scales all need to be assessed. Industries handling metals need harmonised data on the natural distribution of chemical elements, and of soil properties governing element availability at the continental scale.

The REACH regulation requires that exposure and risk assessment be performed according to land use. The GEMAS project focused on agricultural and grazing land soil, both linked to the human food chain. According to REACH, the sample depth should be 0–20 cm for agricultural soil (arable land, A_p-horizon) and 0–10 cm for grazing land soil (land under permanent grass cover), and the <2 mm grain size is the fraction to be analysed. The samples were not pulverised prior to analysis, and a sample weight of 15 g was used for the digestion to overcome possible sample inhomogeneity issues. Risk assessment for most elements is typically based on their hot aqua regia extractable concentration (e.g., Oorts and Schoeters, 2014). Except for sample density, sampling specifications were rigidly fixed by external requirements.

As a result of historical and current anthropogenic input from diffuse sources on arable and grazing land due to, e.g., the use of fertilisers including manure, the direct measurement of the natural geochemical background concentrations in Europe is challenging. Therefore, measured soil element concentrations in these land use types can be considered as ‘ambient background’ concentration (i.e., the sum of the natural background of an element and diffuse

anthropogenic input in the past or present, without the massive influence of point sources) as defined in the REACH guidance document on Environmental risk assessment for metals and metal compounds (ECHA, 2008).

With regard to sample density, it was decided to follow the example of an earlier project covering north-eastern Europe (Baltic Soil Survey: Reimann et al., 2003) and to sample one site per 2500 km² (50 × 50 km grid). The coordinates of the grid cells were centrally provided, but the sampling teams were free to decide where within a grid cell the two samples of agricultural and grazing land soil were taken. Sample materials, and especially the certified trace element free bags used for storing the samples, were centrally provided to all field teams.

2.2. Sampling

Samples of agricultural and grazing land soil were collected in 2008 and early 2009 according to the Field Manual provided to all participants of the GEMAS project (EuroGeoSurveys Geochemistry Working Group, 2008; Reimann et al., 2014a). In total, 2108 samples of agricultural soil (Ap) and 2024 samples of grazing land soil (Gr) were collected, with 33 countries participating in the survey. The Ap samples were collected as composite samples (c. 3.5 kg) from the uppermost 20 cm of regularly ploughed fields. Grazing land soil samples were collected as composite samples from the uppermost 10 cm of land under permanent grass cover. Each sampling site was carefully documented and field duplicates were taken at every 20th site for quality control purposes. The focus of this article will be on the Ap samples; results of Gr samples are comparable.

2.3. Sample preparation and analysis

All samples were air-dried and sieved using a 2 mm nylon screen in the same laboratory. Subsequently, they were homogenised and split into ten aliquots for analysis and future reference. The samples were then randomised, and analytical replicates and aliquots of the two GEMAS project standards Ap and Gr were inserted every 20th sample (Mackovych and Lucivjanský, 2014).

One such split of all samples was sent to ACME Labs (now Bureau Veritas Minerals) in Vancouver, Canada. Prior to analysis, the soil samples underwent a modified hot aqua regia digestion, consisting of equal parts of concentrated ACS grade HCl and HNO₃ and de-mineralised H₂O. Fifteen grams of each sieved and unmilled mineral soil sample were digested in 90 ml of the HNO₃/HCl/H₂O mixture for one hour in a hot (95 °C) water bath. After cooling, the solution was made up to a final volume of 300 ml with 5% HCl. The final sample-weight-to-solution-volume ratio was 1 g per 20 ml.

The filtered solutions were analysed using a Spectro Ciros Vision emission spectrometer (ICP-AES) and a Perkin Elmer Elan 6000/9000 inductively coupled plasma quadrupole mass spectrometer (ICP-QMS) for 53 elements (Ag, Al, As, Au, B, Ba, Be, Bi, Ca, Cd, Ce, Co, Cr, Cs, Cu, Fe, Ga, Ge, Hf, Hg, In, K, La, Li, Mg, Mn, Mo, Na, Nb, Ni, P, Pb, Pd, Pt, Rb, Re, S, Sb, Sc, Se, Sn, Sr, Ta, Te, Th, Ti, Tl, U, V, W, Y, Zn and Zr). The focus of this manuscript is on the following 28 elements being either (a) potentially toxic elements (PTEs): Ag, As, Ba, Cd, Co, Cr, Cu, Hg, Mo, Ni, Pb, Sn, V and Zn; (b) 'high-tech' critical elements (HTCEs): Be, Ce, Ga, Ge, In, La, Li and Tl; or (c) additional elements that sometimes are discussed in environmental legislation as further PTEs: B, Bi, Mn, Sb, Se and U as shown in Table 1. For all other elements, results are included in the Supplementary Material (Table 1SM). For the elements Te, Pd, Pt, Re and Ta, between 50% (Te) and 94% (Ta) of all results were below the detection limit (see Appendix B in Reimann et al., 2014b). They are included here in the tables and figures in the Supplementary Material, because some of

the discussed methods still allow the establishment of a threshold even for such strongly censored data (e.g., the 98th percentile can be provided for all elements).

Results of an international round robin test carried out using the two GEMAS project standards Ap and Gr have demonstrated that ACME's HNO₃/HCl/H₂O digestion, described above, delivered analytical results for most elements that are comparable to a classical aqua regia digestion (Kriete, 2011; Reimann et al., 2012a); a summary of the report can also be found in Reimann et al. (2014a). For elements such as Cs, Ga, K, Na and Ti, which are predominantly bound in the lattice of silicates, this digestion delivers lower values than obtained from a 'classical' aqua regia digestion. For results of the round robin test, see Table 7.2 in Reimann et al. (2014a, p. 63–65). All quality control results of the aqua regia digestion are presented in Reimann et al. (2009a) and summarised in Reimann et al. (2014a).

2.4. Determination of geochemical threshold values

Methods for deriving geochemical threshold values for the identification of samples with unusually high element concentrations vary substantially. The original 'simplistic' approach, to calculate Mean + 2 Standard Deviations (SD) of a given data set, was developed in exploration geochemistry to detect data outliers (Hawkes and Webb, 1962; Matschullat et al., 2000; Reimann and Garrett, 2005). This approach is outdated because, among other shortcomings, it did not consider the multimodal nature of geochemical data sets (Tennant and White, 1959; Lepeltier, 1969; Sinclair, 1974, 1976, 1983, 1986, 1991; Reimann and Filzmoser, 2000).

Reimann et al. (2005) suggested that it may be possible to replace the 'old' Mean + 2 SD approach by using Median + 2 Median Absolute Deviations (MAD), where the Median is defined for a sample x_1, \dots, x_n as $\text{median}_i(x_i)$, and the MAD as:

$$\text{MAD}_i(x_i) = 1.48 \cdot |\text{median}_i(x_i) - \text{median}_j(x_j)|$$

If the sample follows a normal distribution, this definition of the MAD with the constant 1.48 leads to a consistent estimator of the underlying standard deviation.

This definition is much more robust against the effect of data outliers, common in geochemical data sets. The disadvantage of this method if applied to raw, untransformed data, is that it delivers very conservative (low) threshold values (quite often around the 90th percentile), i.e., produces a lot of sites that need to be checked. The reason is that geochemical data distributions are most often strongly right-skewed, while, when using the above formula, the underlying assumption is a symmetrical (not necessarily normal) data distribution. The correct approach to using this formula would thus be to calculate Median + 2 MAD on the log-transformed data (e.g., using log base 10) and to then back-transform the result and use these values as threshold according to the formula:

$$\text{Threshold} = 10^b$$

where

$$b = (\text{median}_i(\log_{10}(x_i)) + 2 \cdot \text{MAD}_j(\log_{10}(x_j)))$$

These values are provided in Table 2. Such an approach will result in values that are often quite comparable to the TIF (Tukey inner fence, or upper whisker of the Tukey boxplot), which is discussed below.

Another approach, which again stems from exploration geochemistry, is to study data distributions in a cumulative probability (CP) diagram (Tennant and White, 1959; Lepeltier,

Table 1

Statistical summary of HNO₃/HCl/H₂O (modified aqua regia)-soluble concentrations of selected 28 PTEs and HTCEs in European agricultural soil (GEMAS Ap samples, 0–20 cm, <2 mm - data source: Reimann et al., 2014a,b). n: number of samples; DL: detection limit; Min: minimum concentration; Q25: 25th percentile of the data set; MEDIAN: 50th percentile of the data set; Q75: 75th percentile of the data set and Max: maximum concentration. Median absolute deviation (MAD) and interquartile range (IQR) are provided as measures of spread. For each element, these values are given separately for the whole European data set and for northern and southern Europe. For a table of all 53 elements see Table 1SM in the Supplementary Material. All element concentration values in mg/kg.

Element	Material	n	DL	%<DL	Min	Q25	Median	Q75	Max	MAD	IQR
Ag	Ap	2108	0.002	0.4	<0.002	0.023	0.038	0.060	2.95	0.025	0.037
	Ap-north	814	0.002	0.74	<0.002	0.020	0.031	0.051	0.26	0.021	0.031
	Ap-south	1294	0.002	0.15	<0.002	0.027	0.042	0.063	2.95	0.026	0.036
As	Ap	2108	0.05	0.1	<0.05	2.5	5.5	9.5	666	4.87	7.02
	Ap-north	814	0.05	0.25	<0.05	1.2	2.3	3.7	37	1.78	2.5
	Ap-south	1294	0.05	0	0.191	5.4	8	12	666	4.58	6.53
B	Ap	2108	0.5	7.4	<0.5	1.23	2.42	4.68	49	2.18	3.46
	Ap-north	814	0.5	11	<0.5	0.89	1.7	3.3	16	1.48	2.39
	Ap-south	1294	0.5	4.9	<0.5	1.6	3.0	5.6	49	2.54	3.99
Ba	Ap	2108	0.5	0	2.6	36	62	95	818	42.4	59.7
	Ap-north	814	0.5	0	3.9	26	43	71	426	29.8	45.6
	Ap-south	1294	0.5	0	2.6	48	74	108	818	44.8	60.2
Be	Ap	2108	0.05	1.8	<0.05	0.28	0.52	0.81	11	0.385	0.53
	Ap-north	814	0.05	3.9	<0.05	0.16	0.28	0.46	2.0	0.212	0.301
	Ap-south	1294	0.05	0.46	<0.05	0.45	0.68	0.98	11	0.379	0.525
Bi	Ap	2108	0.005	0.05	<0.005	0.094	0.16	0.26	3.2	0.118	0.164
	Ap-north	814	0.005	0.1	<0.005	0.060	0.093	0.14	2.3	0.0548	0.0774
	Ap-south	1294	0.005	0	0.009	0.15	0.22	0.31	3.2	0.111	0.157
Cd	Ap	2108	0.01	0.2	<0.01	0.11	0.18	0.28	7.5	0.12	0.17
	Ap-north	814	0.01	0.2	<0.01	0.079	0.12	0.18	1.1	0.0744	0.104
	Ap-south	1294	0.01	0.2	<0.01	0.148	0.22	0.34	7.5	0.134	0.189
Ce	Ap	2108	0.1	0	1.6	18.5	28.4	39.5	265	15.3	21
	Ap-north	814	0.1	0	1.6	17	26	38.2	158	15.9	21.7
	Ap-south	1294	0.1	0	2	20	30	40	265	15.2	20.5
Co	Ap	2108	0.1	0.05	<0.1	3.7	7.5	12	126	6.06	8.22
	Ap-north	814	0.1	0	0.12	2.1	4.1	7.1	66	3.4	4.98
	Ap-south	1294	0.1	0.1	<0.1	6.1	10	14	126	5.74	7.72
Cr	Ap	2108	0.2	0	0.4	11	20	33	696	15.3	21.2
	Ap-north	814	0.2	0	0.6	7.2	13	23	221	10.7	15.8
	Ap-south	1294	0.2	0	0.4	16	25	36	696	14.6	20.2
Cu	Ap	2108	0.01	0	0.3	8.3	15	24	395	11.1	15.7
	Ap-north	814	0.01	0	0.65	5.4	9.4	16	78	6.84	10.1
	Ap-south	1294	0.01	0	0.3	12	19	29	395	12.1	17
Ga	Ap	2108	0.05	0	0.14	2.2	3.4	4.9	20	1.96	2.66
	Ap-north	814	0.05	0	0.14	1.6	2.7	4.1	17	1.83	2.5
	Ap-south	1294	0.05	0	0.19	2.7	3.9	5.2	20	1.87	2.54
Ge	Ap	2108	0.02	26	<0.02	<0.02	0.037	0.058	0.21	0.0311	0.0475
	Ap-north	814	0.02	2.2	<0.02	<0.02	0.035	0.059	0.21	0.0363	0.0487
	Ap-south	1294	0.02	30	<0.02	0.022	0.039	0.057	0.21	0.0262	0.0354
Hg	Ap	2108	0.003	0.8	<0.003	0.018	0.030	0.048	1.6	0.0207	0.03
	Ap-north	814	0.003	0.74	<0.003	0.016	0.024	0.035	0.914	0.0135	0.0192
	Ap-south	1294	0.003	0.77	<0.003	0.021	0.036	0.058	1.56	0.0251	0.0364
In	Ap	2108	0.003	3.8	<0.003	0.0102	0.0176	0.026	0.348	0.0114	0.0158
	Ap-north	814	0.003	6	<0.003	0.01	0.012	0.018	0.073	0.00812	0.0113
	Ap-south	1294	0.003	3	<0.003	0.014	0.021	0.029	0.348	0.0108	0.0149
La	Ap	2108	0.1	0	1	9.5	14	20	109	7.66	10.5
	Ap-north	814	0.1	0	1	8.9	14	20	93	7.9	11.1
	Ap-south	1294	0.1	0	1.1	9.9	15	20	109	7.43	10
Li	Ap	2108	0.1	0	0.16	5.8	11	19	136	9.55	13.7
	Ap-north	814	0.1	0	0.16	3.4	6.4	12	58	5.32	8.13
	Ap-south	1294	0.1	0	0.17	9.1	15	22	136	9.66	13.2
Mn	Ap	2108	1	0	2	241	445	701	14969	334	460
	Ap-north	814	1	0	12	154	280	437	14969	199	283
	Ap-south	1294	1	0	2	354	584	846	14605	361	492
Mo	Ap	2108	0.01	0	0.026	0.25	0.42	0.72	14	0.294	0.473
	Ap-north	814	0.01	0	0.031	0.19	0.32	0.66	13	0.255	0.471
	Ap-south	1294	0.01	0	0.026	0.31	0.46	0.74	14	0.282	0.427
Ni	Ap	2108	0.1	0.05	<0.1	6.8	15	27	2475	13.7	20.5
	Ap-north	814	0.1	0	0.32	3.9	8.2	15	99	7.24	10.6
	Ap-south	1294	0.1	0.1	<0.1	12	21	34	2475	15.8	22.2
Pb	Ap	2108	0.01	0	1.6	9.9	16	24	1309	10.1	13.9
	Ap-north	814	0.01	0	1.6	6.7	9.7	14	52	4.97	6.86
	Ap-south	1294	0.01	0	2.1	15	21	29	1309	9.48	13.6
Sb	Ap	2108	0.02	0.7	<0.02	0.12	0.23	0.42	17	0.192	0.293
	Ap-north	814	0.02	1.8	<0.02	0.073	0.11	0.16	1.0	0.0638	0.0875
	Ap-south	1294	0.02	0	0.022	0.24	0.35	0.55	17	0.208	0.309
Se	Ap	2108	0.05	2.2	<0.05	0.24	0.35	0.49	3.8	0.18	0.248
	Ap-north	814	0.05	3.3	<0.05	0.20	0.31	0.43	3.5	0.175	0.233
	Ap-south	1294	0.05	1.6	<0.05	0.27	0.39	0.52	3.8	0.18	0.245

(continued on next page)

Table 1 (continued)

Element	Material	n	DL	%<DL	Min	Q25	Median	Q75	Max	MAD	IQR
Sn	Ap	2108	0.4	0.38	<0.1	0.46	0.72	1.1	30	0.462	0.672
	Ap-north	814	0.1	0.4	<0.1	0.32	0.48	0.76	8	0.279	0.437
	Ap-south	1294	0.1	0.4	<0.1	0.63	0.87	1	30	0.448	0.685
Tl	Ap	2108	0.0005	0.2	<0.0005	0.075	0.12	0.18	2.5	0.0732	0.106
	Ap-north	814	0.0005	0	0.0025	0.054	0.086	0.14	0.58	0.0554	0.0845
	Ap-south	1294	0.0005	0	0.0025	0.092	0.14	0.20	2.5	0.0736	0.107
U	Ap	2108	0.04	0.05	<0.04	0.52	0.77	1.2	24	0.43	0.678
	Ap-north	814	0.04	0.1	<0.04	0.46	0.68	1.3	24	0.436	0.852
	Ap-south	1294	0.04	0	0.071	0.56	0.81	1.2	21	0.402	0.59
V	Ap	2108	0.2	0	1	16	25	37	188	15	21
	Ap-north	814	0.2	0	1	11	19	31	104	13	19
	Ap-south	1294	0.2	0	1.7	20	29	40	188	14.2	20
Zn	Ap	2108	0.1	0	2.8	27	45	65	1396	28	38
	Ap-north	814	0.1	0	2.9	20	30	48	139	19	28
	Ap-south	1294	0.1	0	2.8	38	53	72	1396	25	34

1969; Sinclair, 1974, 1976, 1983, 1986, 1991; Chaffee, 1977). This is a very powerful plot displaying the statistical data distribution from which one can detect processes that cause deviation from the overall data distribution (Reimann et al., 2008, 2011; Johnson, 2011). Upper outliers can most often be detected as a break in the data distribution in these plots. In Great Britain, CP diagrams and percentiles (most frequently the 95th percentile) have been used to detect samples deviating from the 'normal background variation' and to identify a case-specific threshold (e.g., Cave et al., 2012; Johnson et al., 2012; Ander et al., 2013). Using CP diagrams is scientifically the most rigorous procedure, but requires an expert to carefully assess the element concentration distributions within any given data set, and it involves a certain amount of subjectivity in population splitting. Values obtained for the GEMAS Ap data set are provided in Table 2 (see also Table 2SM for all elements).

Other researchers have suggested using a percentile-based approach, and proposed a simplistic method with the 90th, 95th or 98th percentile of a given data set defining the threshold (see Ander et al., 2013). The 98th percentile, which identifies 2% of all samples as upper 'outliers', comes closest to the original method of calculating the Mean + 2 SD in the case of a normal distribution, which would result in 2.3% of upper outliers. A common feature of all these statistical methods is that it will not necessarily be possible to establish a meaningful single threshold valid for a whole country, because the background varies spatially (see below). Furthermore, there exists no valid scientific reason why 2, 5 or 10% of the samples should be considered as 'anomalous' regardless of the statistical data distribution. Values for the 90th, 95th and 98th percentile are provided in Table 2 (Table 2SM for all elements).

Kürzl (1988) originally suggested using Tukey boxplot boundaries (Tukey, 1977) for class selection in geochemical mapping. The International Organization for Standardisation (ISO, 2005; Annex B) has proposed to detect outliers in a soil geochemical composition based on the Tukey inner (upper) fence (TIF) or upper whisker in a boxplot. The method has many advantages: it is based on the boxplot, an exploratory data analysis tool originally introduced by Tukey (1977), and solely depends on the data distribution; it allows the definition of a threshold for outliers even if none are present in the data set (i.e., Max < TIF), as it extrapolates from the robust inner core (25th to 75th percentiles) of the data's structure. The TIF is calculated as follows:

$$\text{TIF} = \text{Q3} + 1.5 \cdot \text{IQR}$$

where Q3 stands for the 3rd quartile (equivalent to the 75th percentile), and IQR is the interquartile range (75th–25th percentile). The multiplying factor of 1.5 in the formula is based on the

assumption of a symmetrical data distribution. Note that the TIF, when dealing with geochemical data, which are most often right-skewed, must be calculated on the log- (or otherwise) transformed data to achieve 'symmetry' (Reimann et al., 2008). Because the calculation of the TIF is based on a multiple of IQR, it can happen that in narrow data distributions no values are detected above the TIF in a given data set for certain elements. TIF values for the GEMAS Ap data set are again provided in Table 2 (Table 2SM for all elements). The TIF can be considered as one of the most reliable and powerful tools to calculate meaningful threshold values for any given data set (Reimann and Caritat, 2017). Jarva et al. (2010) apply this approach for identifying areas in Finland needing closer attention. The TIF is also used in the REACH guidance to detect outliers in monitoring data sets used for exposure assessment (ECHA, 2016).

Soil Guideline Values (SGVs) have been defined for at least some chemical elements in many European countries (see e.g., Carlon et al., 2007 or Appendix A in Reimann et al., 2014b). For a proper comparison of these values, one needs to know their purpose and derivation method (data, protection goal and level, etc.). For example, there is an important difference between SGVs used as screening values for risk assessment (e.g., Predicted No Effect Concentrations, or PNEC, in REACH, or threshold values as defined in Finland) and soil clean-up values (e.g., Finnish guideline values). The former are conservative concentrations of chemical substances found in soil below which no adverse effects are expected to occur; they are indicators that soil concentrations above this level may pose a risk for the environment or human health. In contrast, soil clean-up values are thresholds for unacceptable effects for a certain land use and were set to define the need for intervention (e.g., remediation). Furthermore, some thresholds depend on ecotoxicological effects, while others are driven by effects on human health or potential for groundwater contamination. In the early days of soil protection, these guideline values, action levels or maximum admissible concentrations were quite often based on geochemical thresholds derived from often rather local soil mapping projects combined with some toxicological considerations (e.g., Kloke, 1980). The existing values for the 28 elements discussed here and their range in European legislation are summarised in Table 3.

Environmental risk assessment according to the methods outlined in Tristán et al. (2000), Smolders et al. (2009), Demetriades (2011), Oorts and Schoeters (2014), Oorts et al. (2016) and Birke et al. (2016) are scientifically the most rigorous procedures. These approaches start with the existence of good ecotoxicological data of the elements for a variety of soil organisms and acknowledge that a site-specific risk assessment must be carried out, because the

Table 2

Geochemical threshold values for the selected 28 elements (same selection as used in Table 1), estimated by a variety of methods: Median + 2 Median Absolute Deviation (MED+2MAD), the 90th, 95th and 98th percentiles, the break in the CP diagram and the Tukey Inner Fence (upper whisker of the boxplot – TIF); n: number of samples. Values are provided for the complete GEMAS Ap data set (Ap) and for northern (Ap-north) and southern (Ap-south) Europe separately. The threshold values for all 53 elements analysed are provided in Table 3SM in the Supplementary Material.

Element	Material	n	Threshold values all HNO ₃ /HCl/H ₂ O extraction, all in mg/kg						Number of samples above threshold n>					
			MED+2MAD	Q90	Q95	Q98	CP plot	TIF	MED+2MAD	Q90	Q95	Q98	CP plot	TIF
Ag	Ap	2108	0.15	0.096	0.13	0.19	0.25	0.23	67	211	106	43	27	27
Ag	Ap-north	814	0.13	0.088	0.12	0.14	0.14	0.21	26	82	41	17	16	2
Ag	Ap-south	1294	0.16	0.098	0.14	0.23	0.23	0.23	52	130	65	26	26	26
As	Ap	2108	36	15	21	34	60	71	37	211	106	43	17	9
As	Ap-north	814	12.0	5.6	7.2	10	12	17	12	82	41	17	10	2
As	Ap-south	1294	26	19	27	44	55	38	70	130	65	26	19	30
B	Ap	2108	18	7.64	10	13.8	20	30	19	210	106	43	16	3
B	Ap-north	814	12	5.6	7.0	8.1	10	16	5	82	40	17	10	0
B	Ap-south	1294	19	8.9	11	16	18	30	17	130	65	26	18	3
Ba	Ap	2108	264	142	180	252	400	400	40	211	106	43	11	11
Ba	Ap-north	814	196	100	130	177	180	330	8	82	41	17	17	2
Ba	Ap-south	1294	249	162	209	303	310	367	41	130	65	26	25	15
Be	Ap	2108	2.5	1.2	1.4	2.0	3	4	25	211	106	43	15	12
Be	Ap-north	814	1.3	0.70	0.88	1.12	1.1	2	6	82	41	17	19	1
Be	Ap-south	1294	2.1	1.31	1.69	2.39	3	3	41	130	65	26	15	15
Bi	Ap	2108	0.71	0.38	0.49	0.76	0.8	1.2	51	211	106	43	38	17
Bi	Ap-north	814	0.31	0.22	0.28	0.36	0.5	0.45	27	82	41	17	3	6
Bi	Ap-south	1294	0.63	0.45	0.61	0.94	1	0.88	60	130	65	26	23	32
Cd	Ap	2108	0.73	0.45	0.61	0.98	2	1.1	71	211	106	43	12	36
Cd	Ap-north	814	0.43	0.26	0.35	0.48	0.7	0.7	23	82	41	17	4	7
Cd	Ap-south	1294	0.75	0.53	0.74	1.2	2	1.2	64	130	65	26	12	28
Ce	Ap	2108	86	51	62	81	150	118	34	210	106	43	6	12
Ce	Ap-north	814	90	52	64	85	100	126	13	82	41	17	9	3
Ce	Ap-south	1294	85	51	62	78	100	114	22	130	65	26	13	9
Co	Ap	2108	39	17	22	29	35	66	17	211	106	43	21	2
Co	Ap-north	814	24	11	13	17	17	31	4	81	41	17	18	3
Co	Ap-south	1294	33	20	26	32	35	45	24	130	65	26	20	9
Cr	Ap	2108	92	48	70	107	250	157	66	211	106	43	10	22
Cr	Ap-north	814	72	37	48	63	70	118	10	82	41	17	10	1
Cr	Ap-south	1294	84	57	89	142	200	121	72	130	65	26	18	33
Cu	Ap	2108	69	38	50	78	150	118	52	211	106	43	10	21
Cu	Ap-north	814	44	23	30	39	75	73	10	82	41	17	1	1
Cu	Ap-south	1294	73	44	61	100	100	104	45	130	65	26	26	24
Ga	Ap	2108	11	6.4	7.6	9.1	12	14	19	211	106	43	13	5
Ga	Ap-north	814	11	5.6	6.7	7.9	10	14	5	82	41	17	5	1
Ga	Ap-south	1294	10	6.7	8.1	9.7	12	14	20	130	65	26	10	4
Ge	Ap	2108	0.17	0.078	0.094	0.13	0.2	0.21	8	210	106	43	3	1
Ge	Ap-north	814	0.22	0.082	0.10	0.14	0.12	0.21	0	82	41	17	29	1
Ge	Ap-south	1294	0.15	0.076	0.090	0.11	0.14	0.21	7	130	65	26	8	0
Hg	Ap	2108	0.13	0.076	0.10	0.19	0.5	0.21	72	211	106	43	9	36
Hg	Ap-north	814	0.076	0.051	0.066	0.10	0.1	0.11	26	82	41	17	19	13
Hg	Ap-south	1294	0.160	0.088	0.126	0.25	0.25	0.26	48	130	65	26	27	24
In	Ap	2108	0.065	0.036	0.042	0.055	0.075	0.1	21	211	106	43	17	9
In	Ap-north	814	0.049	0.026	0.033	0.041	0.05	0.07	6	82	41	17	6	1
In	Ap-south	1294	0.059	0.039	0.047	0.063	0.08	0.08	28	130	65	26	13	13
La	Ap	2108	42	26	32	42	60	57	44	211	106	43	18	18
La	Ap-north	814	45	27	34	44	60	66	16	82	41	17	8	5
La	Ap-south	1294	41	25	30	40	60	54	26	130	65	26	10	11
Li	Ap	2108	63	28	33	45	75	95	15	211	106	43	9	3
Li	Ap-north	814	38	20	24	28	27	58	5	82	41	17	26	0
Li	Ap-south	1294	55	31	38	52	50	84	24	130	65	26	33	4
Mn	Ap	2108	2141	1080	1415	1831	4000	3345	26	211	105	43	5	7
Mn	Ap-north	814	1288	641	781	1100	1100	2018	14	82	41	17	17	5
Mn	Ap-south	1294	2114	1296	1611	2022	2500	3122	23	130	65	26	12	5
Mo	Ap	2108	2.0	1.2	1.7	3.1	7.5	3.5	81	211	106	43	6	35
Mo	Ap-north	814	1.9	1.3	2.0	3.6	3	4.3	42	82	41	17	25	12
Mo	Ap-south	1294	1.7	1.2	1.6	2.7	2	2.6	54	130	65	26	43	29
Ni	Ap	2108	105	43	66	139	500	216	59	211	106	43	7	21
Ni	Ap-north	814	55	24	30	37	40	99	5	82	41	17	14	0
Ni	Ap-south	1294	100	56	101	193	300	167	68	130	65	26	14	30
Pb	Ap	2108	58	35	44	73	75	87	59	211	106	43	42	32
Pb	Ap-north	814	27	18	22	28	75	39	18	82	41	17	0	3
Pb	Ap-south	1294	54	41	54	97	75	73	65	130	65	26	42	43
Sb	Ap	2108	1.40	0.70	1.0	1.8	3	2.6	60	211	106	43	11	21
Sb	Ap-north	814	0.36	0.22	0.28	0.34	0.3	0.5	14	82	41	17	30	9
Sb	Ap-south	1294	1.20	0.91	1.4	2.5	2	1.9	79	130	65	26	33	35
Se	Ap	2108	1.00	0.68	0.88	1.2	2.5	1.4	71	211	106	43	4	29
Se	Ap-north	814	0.98	0.63	0.77	1.1	1	1.3	21	82	41	17	20	11
Se	Ap-south	1294	0.98	0.73	0.93	1.2	1	1.3	53	130	65	26	51	22

(continued on next page)

Table 2 (continued)

Element	Material	n	Threshold values all HNO ₃ /HCl/H ₂ O extraction, all in mg/kg						Number of samples above threshold n>					
			MED+2MAD	Q90	Q95	Q98	CP plot	TIF	MED+2MAD	Q90	Q95	Q98	CP plot	TIF
Sn	Ap	2108	2.8	1.7	2.4	4.3	10	4.4	85	211	106	43	11	40
Sn	Ap-north	814	1.7	1	2	2	2	2.7	29	82	41	17	10	4
Sn	Ap-south	1294	2.5	2	3	5	10	3.9	92	130	65	26	11	50
Tl	Ap	2108	0.43	0.29	0.38	0.62	1	0.67	62	211	106	43	18	34
Tl	Ap-north	814	0.34	0.24	0.30	0.38	0.4	0.56	0	82	40	17	13	2
Tl	Ap-south	1294	0.42	0.31	0.46	0.74	0.75	0.63	105	130	65	26	26	42
U	Ap	2108	2.6	2.2	3.4	5.2	10	4.2	148	211	106	43	6	63
U	Ap-north	814	2.8	2.7	4.0	5.9	6	6.1	78	82	41	17	16	15
U	Ap-south	1294	2.4	1.8	2.6	4.2	5	3.4	86	130	65	26	17	46
V	Ap	2108	86	51	63	85	120	123	42	211	104	43	14	13
V	Ap-north	814	83	43	51	63	60	105	8	82	40	17	23	0
V	Ap-south	1294	79	57	72	95	150	112	52	130	65	26	10	17
Zn	Ap	2108	161	87	104	129	150	239	19	211	105	42	22	8
Zn	Ap-north	814	110	73	87	110	110	139	16	82	41	17	16	1
Zn	Ap-south	1294	139	94	112	139	180	184	26	130	64	26	18	17

spatial variation of soil parameters, such as cation exchange capacity (CEC), total organic carbon (TOC), grain size (especially clay content) and pH value, may have an impact on the bioavailability and toxicity of elements in soil. It thus requires that all these parameters be determined, in addition to the concentration of the chemical element in question. Moreover, the calculation of (no) effect thresholds and corresponding risks must be carried out for each site and a new set of risk characterisation maps will result (see Cu and Mo in Oorts and Schoeters, 2014; Oorts et al., 2016). Good

and robust ecotoxicology data and bioavailability correction models, considering the effect of varying soil properties on the toxicity of these elements to soil organisms, are presently only available for a limited number of elements (Smolders et al., 2009; Oorts et al., 2016). Furthermore, even ecotoxicology will, in the end, depend on a spatial component because organisms are adapted to natural differences in their habitats (Reimann and Garrett, 2005), resulting in a spatial variation in occurrence of specific organisms. Thus, the performance of a number of simple

Table 3
Comparison of soil guideline values as defined by different authorities for PTEs in Europe and Australia. Data sources for: EU REACH (ECHA dissemination website, <https://echa.europa.eu/information-on-chemicals>), Finland (MEF, 2007), Australia (DEC, 2010), Lithuania, Denmark, Hungary and RANGE EU data provided in Appendix A of Reimann et al. (2014b) have been used. PNEC: Predicted No Effect Concentration; THRESH: threshold value; SGV: Soil Guideline Value; EIL: Ecological Investigation Level (comparable to the European SGVs); HIL 'E': Health Investigation Level class 'E' (parks, recreational open space and playing fields). Note that no guideline values exist so far for several of the elements discussed in this paper. Geochemical threshold values for these elements are provided in Table 2. All element concentration values in mg/kg.

Element	EU REACH	Finland		Lithuania	Denmark	Hungary	RANGE EU	Australia	
	PNEC	THRESH	GUIDEL	SGV	SGV	SGV	SGV	EIL	HIL 'E'
Ag	1.4			2	50	2	2–500		
As	0.7	5	50	10	20	15	10–200	20	200
B	5.7 ^a			50		1000	5–1000		
Ba	208			600	100	250	100–600	300	
Be				10			7–10		
Bi	67.6						0.5–1		
Cd	0.9 (secondary poisoning to mammals and birds) 1.1 (direct toxicity to soil organisms) 0.9 (secondary poisoning to mammals and birds) 1.1 (direct toxicity to soil organisms)	1	10	3	0.5	1	0.5–20	3	40
Ce	0.45 (Ce III)								
Co	10.9 (9.3–83.4) ^b	20	100	30		30	20–100	50	200
Cr	3.2 (Cr III) ^a	100	200	100	500	75	30–1000	400	200
Cu	65 (34.6–135)	100	150	100	500	75	40–1000	100	2000
Ga									
Ge									
Hg	0.02 ^a	0.5	2	1.5	1	0.5	0.5–80	1	30
In									
La	10.7								
Li	0.26								
Mn	15.7 ^a			1500			1500	500	3000
Mo	9.9 (8.1–213)			5	5	7	2.5–60	40	nd
Ni	29.9 (13.3–93.4)	50	100	75	30	40	30–300	60	600
Pb	109 (41–240) (secondary poisoning to mammals and birds) 212 (132–468) (direct toxicity to soil organisms)	60	200	100	40	100	40–750	600	600
Sb	37			10			10–50		
Se	0.1 (Se IV)			5	20	1	1–20		
Sn								50	
Tl					1		0.5–1		
U				20			20		
V	7.2 ^a	100	150	150			100–220	50	
Zn	106.8 (55.4–313) ^a	200	250	300	500	200	60–2500	200	14000

^a added approach (geochemical background not taken into account).

^b range is 5th - 95th percentile of EU soils (based on GEMAS data) for arable soil).

statistical procedures to identify sites with an unusually high element concentration is still of interest as a first screening for locations that may require attention for elements lacking proper ecotoxicological data.

2.5. Statistical data analysis

The statistical analysis of geochemical data must take into account the 'closure' effect due to their compositional nature (Aitchison, 1986). Significant problems with closure occur in bivariate plots, based on element concentrations within one sample material, because each data point – representing one sample – could be shifted along a straight line ($y = a \cdot x$) through this point by changing the abundance of other elements within this particular sample. The effect can be substantial (Filzmoser et al., 2010; Reimann et al., 2012b). Calculated bivariate relations on raw data (e.g., in mg/kg units), therefore, must be interpreted as being exploratory and qualitative at best, or an appropriate transformation must be performed prior to statistical treatment (e.g., Filzmoser et al., 2010; Sealey et al., 2015). Geochemical data cannot be quantitatively compared if they relate element concentrations within the same sample. To compare two sample materials (e.g., Ap vs. Gr), bivariate plots can be used, because the values plotted do not come from a single, closed composition, and are, therefore, independent. In this paper, either univariate statistics are used or two sample media are compared. Thus no log-ratio-transformation of the raw geochemical data was used, but rather the simple log-transformation (base 10) was applied where appropriate. For data analysis the DAS package (Dutter et al., 1992), a combination of R scripts (e.g., StatDA 1.6.9 by Filzmoser, 2015), Wolfram Mathematica (Wolfram Research Inc, 2016) and ArcGIS 10 (Esri, 2016) software were used.

3. Results

Results for the GEMAS agricultural (Ap) and grazing land (Gr) soil samples turned out to be very comparable (Reimann et al., 2014a,b). This paper thus focuses on the 2108 Ap samples, collected from regularly ploughed fields.

Fig. 1 shows six examples of regional distribution maps of element concentrations in European agricultural soil. A major break in element concentrations between southern and northern Europe becomes visible for most elements (typical examples are As and Pb), which follows exactly the southern limit of the most recent glaciation (Reimann et al., 2014a; Scheib et al., 2014). It divides the young postglacial soils of northern Europe from the much older and more weathered soils of southern Europe. The difference in concentration is so large that the definition of just one general threshold value for the European continent appears non-sensical. At least two different background regimes need to be considered at the European scale, and many more occur at the regional and local scales. A general location map, identifying the European countries and their capitals in addition to four sites discussed in more detail in the text, is provided as Fig. 1SM in the Supplementary Material.

Fig. 2 shows CP diagrams for selected elements. Diagrams for all 53 elements are provided in the Supplementary Material (Fig. 2SM). To plot these diagrams, the data set was split into samples from 'northern Europe' and 'southern Europe' along the southern limit of glaciation. The 98th percentile for the two sub-data sets is indicated on the plot. Furthermore, where applicable, the range of Soil Guideline Values (SGVs), as defined by various authorities in the EU member countries (for values see Table 3), is shown by two arrows connected by a stippled line. If only one arrow is shown (for example Cd), Table 3 needs to be checked for an

SGV that is higher than the scale of the diagram permits to plot (for Cd: 20 mg/kg). A stippled green line for the eight example plots (but not in Fig. 2SM) indicates the TIF for all samples in the European data set.

Table 1 presents a statistical overview of the analytical results for the GEMAS Ap samples in the ACME-modified aqua regia digestion for 28 of the analysed 53 elements, with a focus on PTEs and HTCEs. A complete table for all elements can be found in the Supplementary Material (Table 1SM). The lower Detection Limit (DL), number and proportion of samples < DL, minimum, 25th percentile, 50th percentile (median), 75th percentile, maximum concentration and MAD and IQR, as measures of variation, are presented in Table 1.

Table 2 provides the geochemical threshold values for agricultural soil at the European scale and for northern and southern Europe separately, using a variety of established methods for their estimation (TIF, CP diagram, 98th percentile, 95th percentile, 90th percentile). The number of samples exceeding the different threshold values for the 28 elements are shown in Tables 1 and 3. In the Supplementary Material, these different threshold values are provided for all 53 elements analysed (Table 2SM).

Table 3 supplies (eco)toxicological based soil guideline values, such as PNEC values derived by the European REACH Regulation, the so-called soil thresholds and soil guidelines, as provided by the Ministry of the Environment of Finland (MEF, 2007), a selection of soil guideline values, as determined in different European countries (Lithuania, Denmark and Hungary chosen here because they provide SGVs for many elements), and their range in European legislation (taken from Appendix A in Reimann et al., 2014b), together with Environmental Impact Levels (EILs) and Health Impact Levels (HILs), as defined by the Department of Health and Conservation of Western Australia (DEC, 2010). Note that no such values exist for a number of elements discussed here.

Fig. 3a and b shows a selection of maps of the sites where the different thresholds are exceeded for one or several of the selected elements (As, Cd, Hg, Mo, Ni and Pb). Maps for all elements are provided in the Supplementary Material (Fig. 3SM).

4. Discussion

Based on statistics alone, a general geochemical background and threshold to identify unusually high element concentrations in agricultural soil can, of course, be defined for all of Europe, based on the GEMAS data set (see Tables 1 and 3), or any other European scale data set (e.g., the LUCAS data - Tóth et al., 2013, 2016a,b). These values can then be used to estimate what are low, normal, or high element concentrations in European agricultural soil at the continental scale. It thus allows European countries to link their own data to a fully harmonised European reference data set – one of the core aims of the IUGS Global Geochemical Baselines mapping programme (Darnley et al., 1995; Smith et al., 2012).

The spatial distribution maps (Fig. 1), as well as the CP diagrams (Fig. 2), demonstrate that it may not make much sense in Europe, because of variable geology and glacial history, to define a single geochemical background range or a threshold representative for the entire continent. The same conclusion was already reached during the FOREGS (Forum of European Geological Surveys) Geochemical Atlas of Europe project (Salminen et al., 2005; De Vos et al., 2006). At least two different geochemical background regimes immediately emerge on most maps: northern Europe with clearly lower concentrations for many elements and southern Europe with clearly higher concentrations. This break in element concentration is directly related to the geological history of Europe. Thus, statistics and different threshold values are presented for three cases in Table 2: (1) all of Europe, (2) northern Europe and (3)

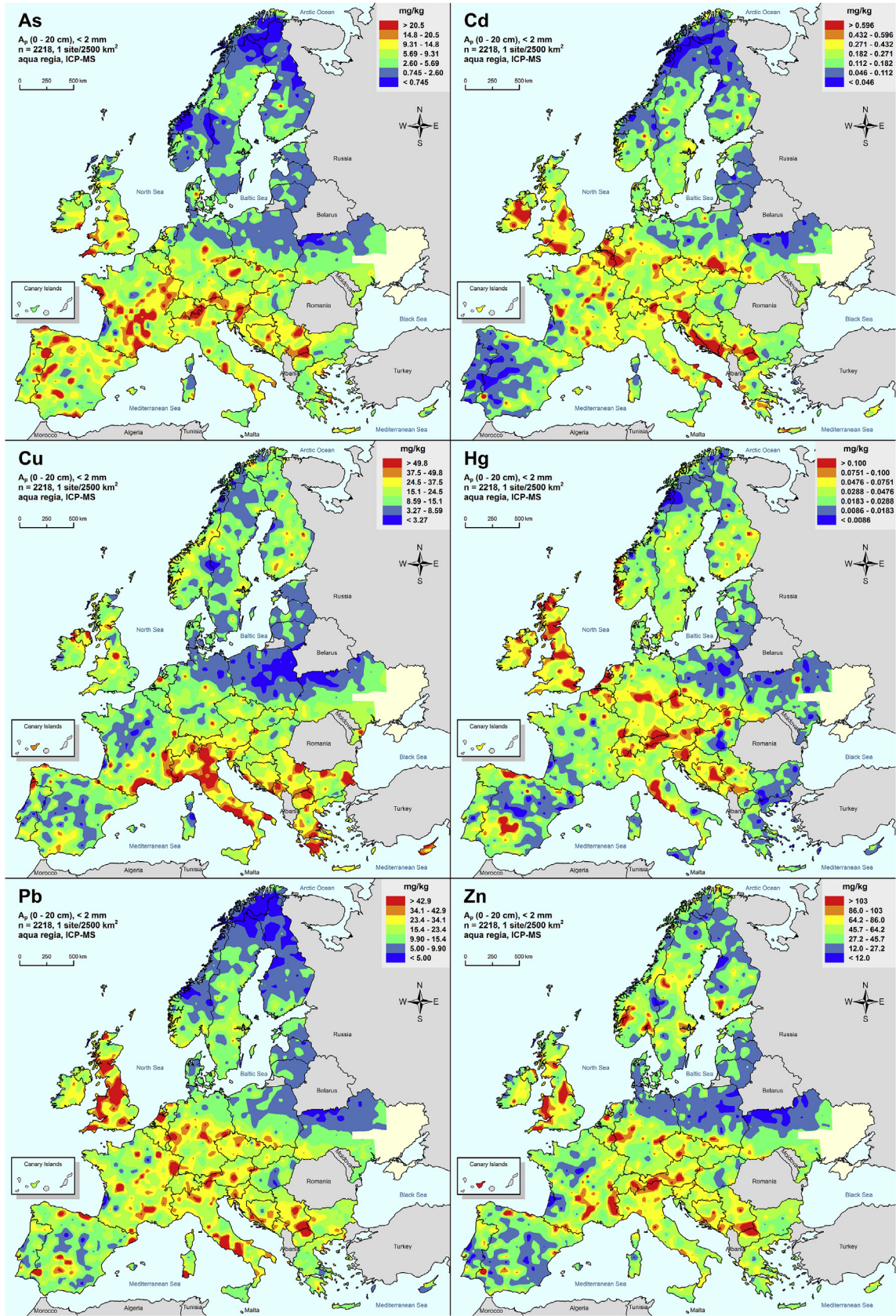


Fig. 1. Regional distribution of the PTEs As, Cd, Cu, Hg, Pb and Zn in the GEMAS agricultural soil samples (Ap, 0–20 cm, <2 mm, modified aqua regia digestion) from Europe. Note the large difference in concentration levels between northern and southern Europe for most elements. Maps for all elements can be found in the GEMAS atlas (Reimann et al., 2014a).

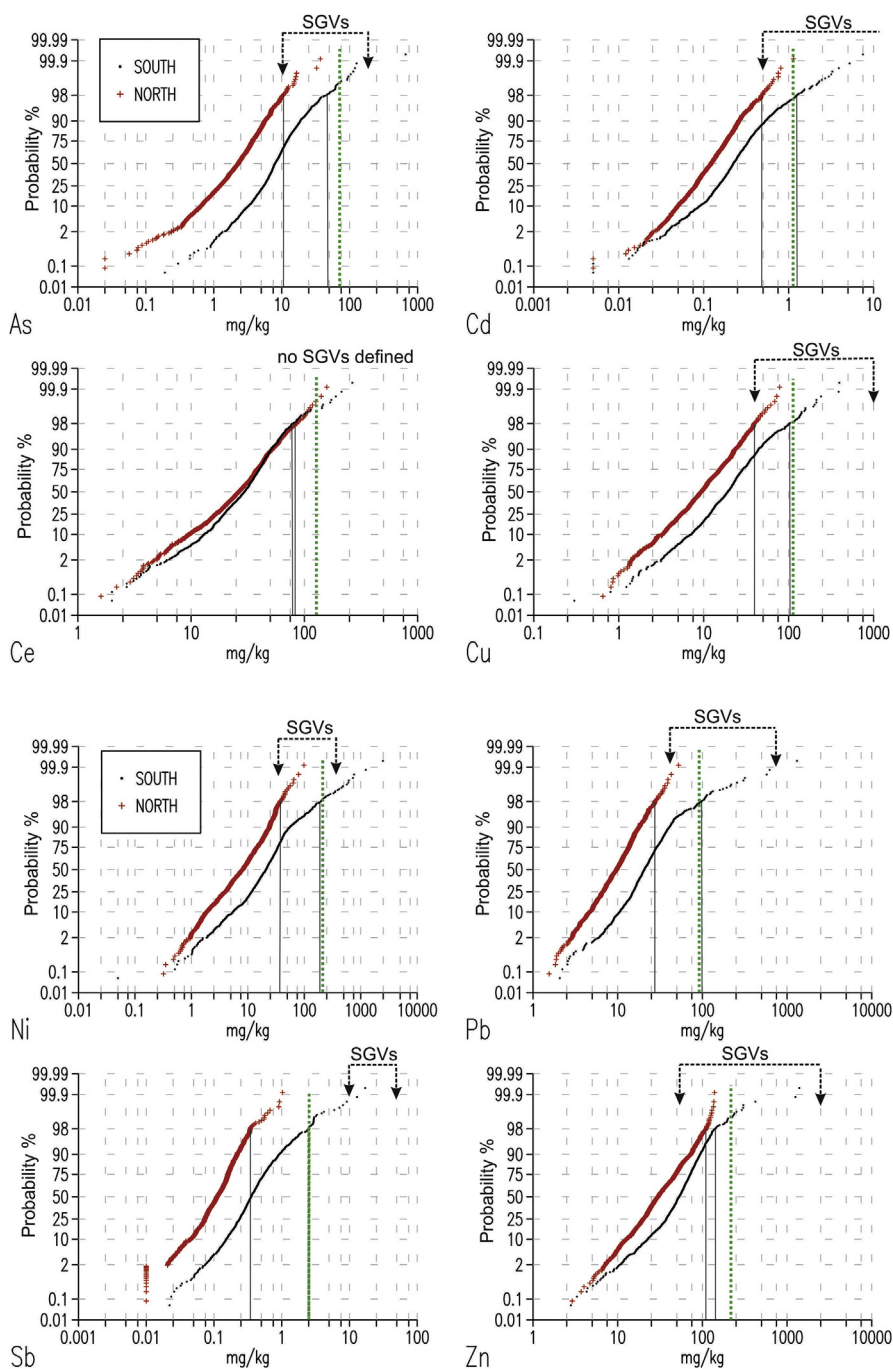


Fig. 2. Cumulative probability (CP) distribution diagrams for eight selected elements (As, Cd, Ce, Cu, Ni, Pb, Sb and Zn), as analysed in a modified aqua regia digestion on the coarse (<2 mm) fraction of the GEMAS Ap samples (red crosses: samples from northern Europe; black dots: samples from southern Europe). The range of Soil Guideline Values (SGVs), as defined by different European authorities, is shown by black arrows connected by stippled lines (Data taken from Appendix A in Reimann et al., 2014b). Note that for Cd the upper limit of guideline values falls outside the diagram (20 mg/kg) and that no SGVs are defined for Ce. The 98th percentiles for the two sub-data sets are indicated by vertical black lines. The TIF, as calculated for the European data set (not differentiated by north and south), is shown as a stippled green vertical line. CP plots for all elements are provided in Fig. 2SM in the Supplementary Material. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

southern Europe. This is a simplistic approach, because an even finer separation of different background areas depending on local geological conditions will become necessary at a more local scale. Quite different geochemical threshold values would need to be estimated for each such scale. This is, however, exactly what the Ministry of the Environment of Finland suggested in its guideline (MEF, 2007). That alone prohibits the use of, for example, the Finnish (or other national) ecotoxicological threshold values for all

of Europe and partly explains the great variation of soil guideline values in Europe, as displayed in Table 3. To further complicate the matter for the guideline values, the variation is also due to differences in the protection goal (e.g., human health vs. environment) or protection level (screening vs. remediation).

Fig. 2 shows CP diagrams for eight selected elements (diagrams for all 53 analysed elements can be found in the Supplementary Material as Fig. 2SM). The diagrams show that very different

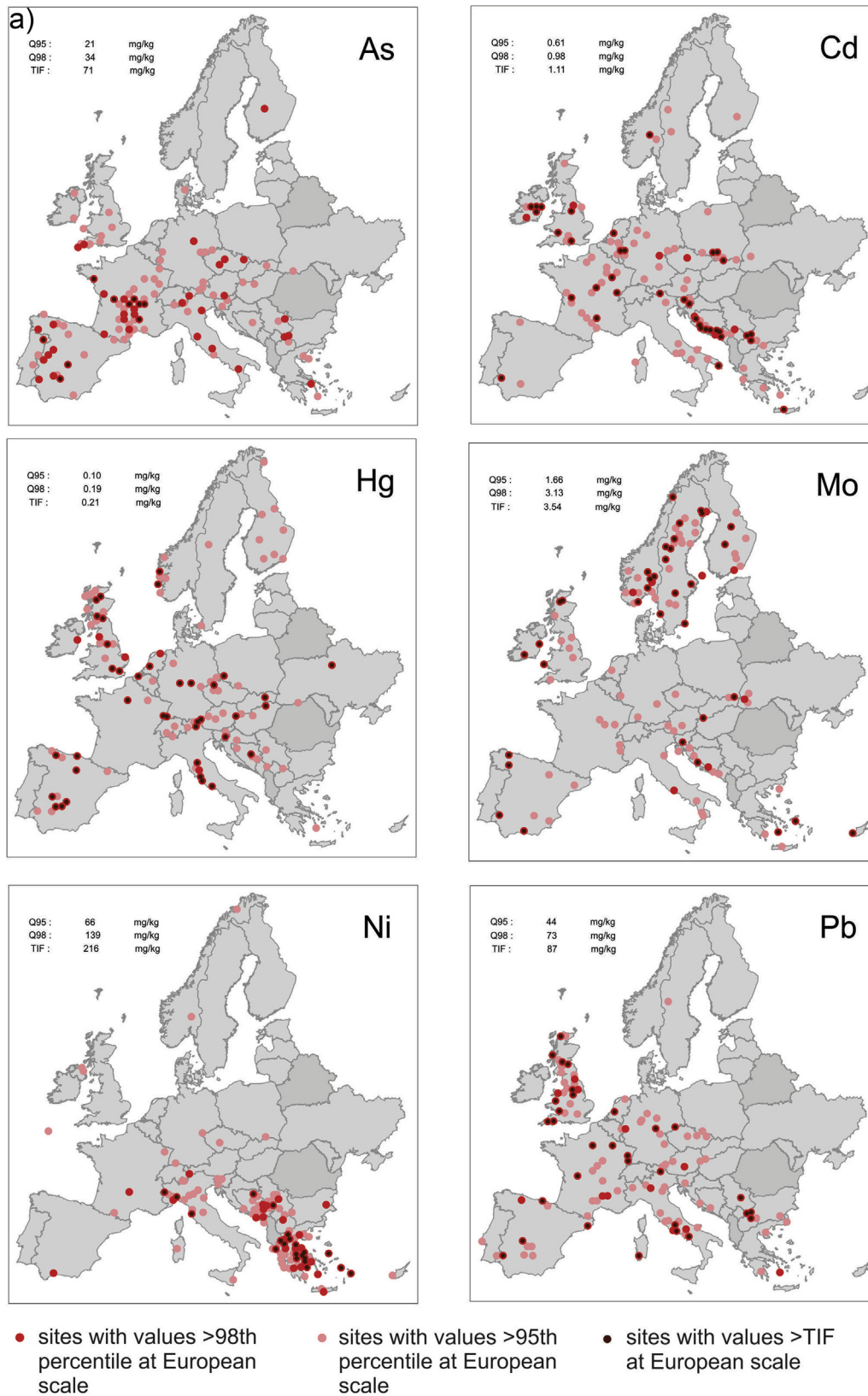


Fig. 3. Maps showing sites with threshold exceedances: (a) values above the 95th, 98th percentile and the TIF calculated for all of Europe; (b) values above the 95th, 98th percentile and the TIF calculated for northern and southern Europe separately. Where a symbol occurs in the Atlantic Ocean south-west of the U.K. (e.g., Ni), this stands for an anomaly of this element on Tenerife (Canary Islands). Maps for all 53 elements can be found in the [Supplementary Material \(Fig. 3SMa,b\)](#). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

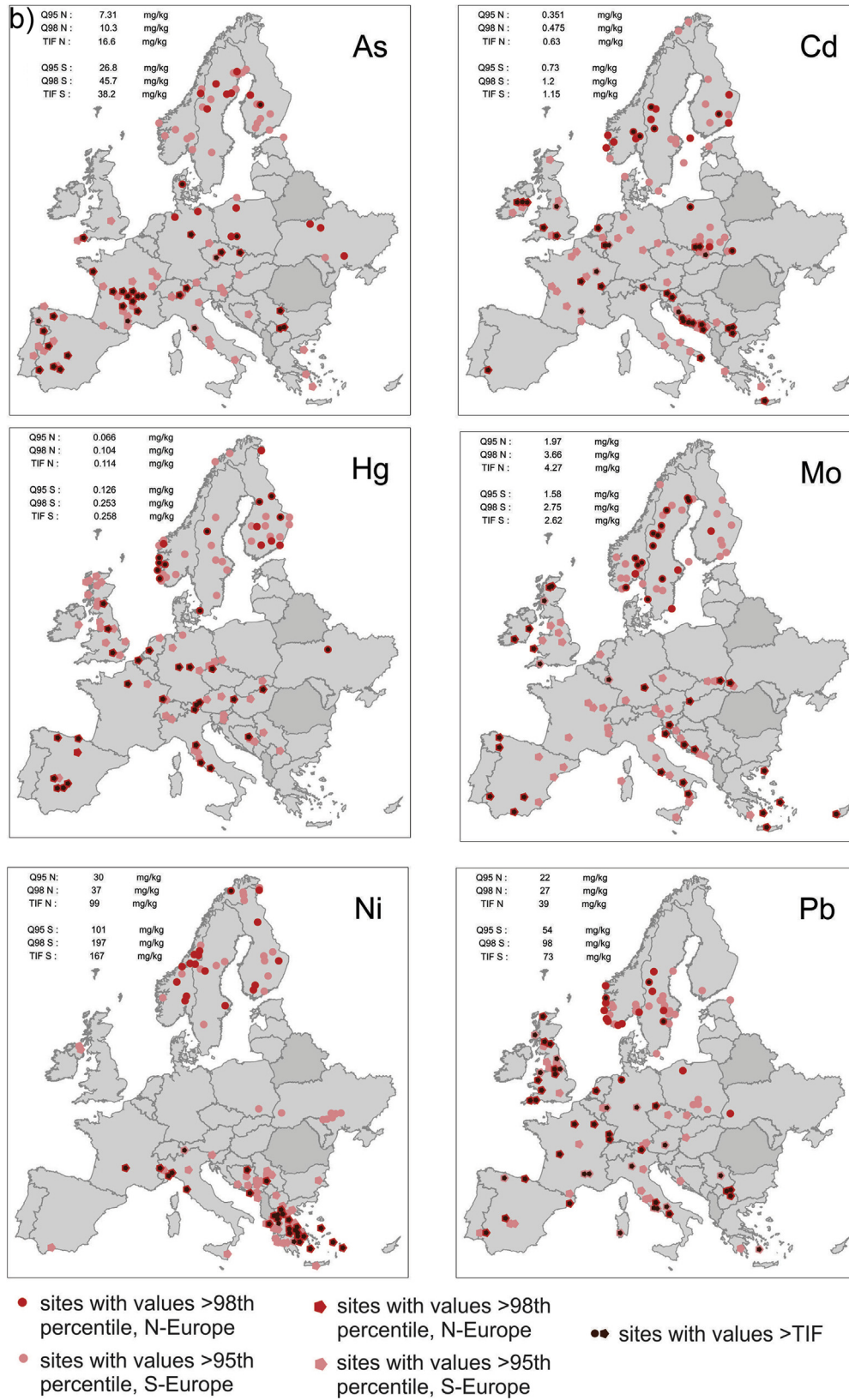


Fig. 3. (continued).

geochemical threshold values emerge for northern and southern Europe for most elements. By using one of the simplest approaches, the 98th percentile, a threshold of 10 mg/kg for As emerges for

northern Europe, while it is almost 50 mg/kg for southern Europe. For Pb, it is about 30 mg/kg in northern Europe and 100 mg/kg in southern Europe. If these values are calculated for single European

countries, the differences are even larger: for As, the 98th percentile is 3.7 mg/kg in Lithuania and 220 mg/kg in Portugal; for Pb, the 98th percentile is 12.5 mg/kg in Lithuania and 439 mg/kg in The Netherlands. *What is, thus, an 'optimum' screening value? Is it meant to detect mineralisation or to protect the general population or the environment or to identify anthropogenic contamination or areas that require follow-up, more detailed monitoring or even remediation?*

There are only very few elements that show almost exactly the same value for the 98th percentile in northern and southern Europe, e.g., Ce, La, Nb, P, Re and Ti (see Fig. 2 for examples or Fig. 2SM for all elements). Even fewer elements exist where the 98th percentile is higher for the samples from northern Europe than for those from southern Europe: Ge, Mo, S and U. All statistical (and spatial) distributions at the continental scale are governed by geology.

When comparing the different statistical methods to derive a geochemical threshold, it is evident that the 'Median + 2 MAD' approach results in a very conservative estimate (low values for the threshold), somewhere between the 90th and 95th percentile (Table 2). This 'conservative' performance may be preferred by Environmental Agencies from the point of view of protection of the environment or human health, as it delivers precautionary levels and, as such, limits the 'false negatives', i.e., sites posing a risk for human health or environment that pass the screening criteria (Rothwell and Cooke, 2015). It will provide, however, a very high number of sites needing attention, and may thus be counter-productive and costly by taking attention and resources away from the few sites that may really present a problem. In the end, any risk analysis should be accompanied by a cost-benefit analysis to ensure that the decisions taken are economically, socially and environmentally viable and fit-for-purpose (Demetriades, 2011).

The 98th percentile delivers a fixed number of cases for each element that needs attention – in case of the GEMAS project agricultural soil data set, with 2108 samples these are 43 sites per element. They are shown in Fig. 3 (a) for the threshold determined for the complete European data set and (b) for the threshold determined for northern and southern Europe separately. That is a large number of sites if 14 PTEs (or even more elements) need to be considered. However, many sites are marked by an exceedance in several elements, which reduces the number of sites needing attention. According to the CP plots, the 98th percentile is, in many cases, also very close to the break in the cumulative distribution curve, which would be used by an expert to identify 'unusual' values in a given data set. The 98th percentile is closest to what the old 'Mean + 2 SD' method identified as 'outliers' in case of a normal distribution (2.3% at each end of the distribution).

The TIF delivers estimates that are most often quite close to the 98th percentile. It has the great advantage that the definition of the threshold depends on the statistical distribution of the robust inner core of the data. In the case of a smooth and symmetrical distribution, no outliers will occur (e.g., many elements, including Ni and Zn from the PTE list, in the northern European subset). As such, it is probably the best method to identify the sites that will need closer attention because they are 'unusual' (Reimann and Caritat, 2017).

A comparison of the calculated geochemical threshold values using a variety of statistical methods to risk-based SGVs, as defined in some European countries, demonstrates the problem of using national guideline values, which are often based on the experience of single member countries for the entire continent. Geochemical conditions and protection goals, based on land use, may indeed differ strongly among member states and, therefore, one cannot use national guideline values in other regions without a good understanding of their basis. Application of improper SGVs may be either over- or under-conservative, resulting in an unnecessary economical or ecological cost,

respectively. Region-specific geochemical threshold values may help to evaluate the appropriateness of SGVs for other regions than originally developed for. By comparing the soil guideline values for Finland and Lithuania, only the guideline values for Cr and Cu are identical. In all other cases, the guideline values are often substantially higher in Lithuania (by a factor of 1.5–3). It must be noted that the Finnish guidelines (MEF, 2007) state that their SGVs are flexible and may need adjustment depending on local geochemical background. This was, for example, overlooked in Tóth et al. (2016a,b) and, consequently, led to their over prediction of risks in soil from southern Europe.

Matters become even more confusing when SGVs from other European countries like Denmark and Hungary are considered in this comparison. Differences between European countries can be larger than a factor of 100 for some elements (see column 'RANGE EU' in Table 3). The Finnish values, for example, are often quite close to the 98th percentile for northern Europe, but deliver a high number of exceedances in southern Europe. Consequently, at the continental scale, using single SGVs per element does not help in the decision-making process as to whether an area is contaminated, poses a risk to human health and environment, or whether monitoring or further action is needed or not.

Site-specific risk assessment based on ecotoxicological data will lead to quite different results, because it not only highlights areas with 'high values' of an element, but also considers soil properties that determine the bioavailability of the elements at each site (Smolders et al., 2009; Oorts and Schoeters, 2014; Oorts et al., 2016). As such, this approach may also identify soil with lower concentrations to have a potential risk because of their poor buffer capacity and, consequently, a high bioavailability of the element for organisms (see range in site-specific PNEC values for Cu, Ni, Mo, Zn, Co and Pb in Table 3). This further reinforces that one single threshold value is not fit for the full range of soil conditions found in Europe. This approach is, however, considerably more data demanding than just determining the 98th percentile or calculating the TIF and, in addition, the necessary information is not currently available for all elements or sites.

The exceedance maps for the 95th and 98th percentiles and TIF based on the GEMAS data set at the European scale show the expected result – almost all locations identified as anomalies occur in southern Europe (exception Ag and Mo – see Fig. 3a for Mo or Fig. 3aSM for all elements). A different picture emerges when calculating the 95th, 98th percentile and the TIF separately for northern and southern Europe (Fig. 3b), based upon the different threshold values for the northern and southern European soil samples. As a first result, the 'outliers' are now more evenly spread across Europe. Whatever choice is made, most anomalies can be linked directly to geology and, thus, do not really need any attention from an anthropogenic impact related risk perspective, follow-up, monitoring or remediation (Fig. 3a and b). Few areas emerge where a sizeable region at the continental scale is affected by high PTE concentrations.

One such example is the Massif Central in France (see Fig. 1SM for location) with many sites showing high As concentrations. Though this is clearly a geological signal, it is one of the areas in France that is under political attention at the national scale (Blum et al., 2002). Drinking water supply and agricultural products may be at risk – though fortunately there exist only few plants, e.g., rice, that take up As – and may thus pose a risk for As poisoning via the food chain (Marin et al., 1993; Abedin et al., 2002). Another such large area that is invisible at the European scale when using the 98th percentile of the entire data set as the threshold occurs for As in central Sweden (Fig. 3a). It becomes, however, prominent when instead the northern European threshold is used (Fig. 3b). Arsenic concentrations in soil are thus high at the northern European scale,

but not at the European scale. Still, Swedish authorities have identified numerous serious problems with excessive As concentration in drinking water wells in this area (e.g., Piteå Kommun, 2005). Again, the source of elevated As concentrations is geogenic (Reimann et al., 2009c) and not related to contamination.

The cluster of high Co, Cr, Cu and Ni values in Hellas (see Fig. 3a and b Ni or Fig. 3SM for maps of all elements) relates to the occurrence of mafic-ultramafic rocks and is thus geogenic (Economou-Eliopoulos and Vacondios, 1995; Economou-Eliopoulos et al., 1999; Rassios and Smith, 2000; Saccani et al., 2011).

The Mo anomaly marking the undeveloped Nordli Mo deposit (see Fig. 1SM for location) in Norway (Pedersen, 1986) and a belt of black shale in Scandinavia has a geogenic origin too (Fig. 3a and b). High Mo, however, should catch the attention of the local authorities because sheep and other ruminants can develop molybdenosis in such areas (e.g., Frank, 1998). Molybdenum taken up by ruminants forms an insoluble complex with Cu and S in the rumen leading to Cu deficiency. This mechanism means that molybdenosis is more likely to occur if dietary Cu is low. Liming should, for example, be avoided on soil high in Mo, because it will release Mo and bind Cu. The Mo anomalies along the Croatian coast (together with Cd, Co, V) are related to special conditions of soil formation and the occurrence of Terra Rossa (Fig. 3a and b).

The Hg, Pb and Cd anomaly at the south-western coast of Norway is a special case (Fig. 3b). It almost disappears when the European scale threshold values are used (compare Fig. 3a with Fig. 3b). The Pb anomaly at the southern tip of Norway, visible in Fig. 3b, has for a long time been attributed to the impact of long range atmospheric transport (LRT) from central and north-western European sources on the clean Norwegian environment (e.g., Steinnes et al., 1994), and was as such interpreted to show the strongest impact of diffuse contamination in all of Europe. Reimann et al. (2009b) have demonstrated, however, that the concentration gradients of many elements, which are observed at the southern tip of Norway (with Au showing the strongest gradient), have a natural explanation. The occurrence of these high element concentrations is related to vegetation and climate and, thus, to natural processes. The Hg anomaly on the west coast of Norway occurs in an area that receives some of the highest amounts of precipitation in Europe. Soil very rich in organic material develops under such conditions and Hg, regardless of source, binds strongly to organic material and will be enriched where humus is abundant. The Pb anomaly, however, continues into the rain shadow at the south coast of Norway and is directly related to a major change in vegetation zones over the first 200 km from the coast (Reimann et al., 2009b). The original source of Pb is the local C horizon.

A few definitely anthropogenic sources of high element concentrations in agricultural soil can also be detected. Examples are cities (London: Ag, Hg; Paris: Ag, Hg, Pb; Kiev: Hg), agriculture (Cu in vineyards in southern Europe), and some usually single-point anomalies, which are most likely related to the vicinity of metal industry or coal-fired power plants (e.g., Cd in Upper Silesia, Hg in the Czech Republic/German/Polish border area). Even a historical impact of anthropogenic activities can be identified on the GEMAS Pb map in eastern France where the battle fields of World War I near Verdun (see Fig. 1SM for location) are marked by a Pb anomaly (Reimann et al., 2012c).

When studying the difference between TIF, the 98th and 95th percentiles it is clear that the lower the geochemical threshold, the more sites needing attention emerge (Fig. 3a and b). Quite often, the areal extent of an anomaly increases too. In other cases, new areas are marked by single-point anomalies and require additional interpretation or investigation. When combining threshold exceedances of all the PTEs in a single map of Europe based, e.g., on

the 90th percentile or the Median + 2 MAD approach (not shown), Europe's soil quality appears as a 'disaster area'. However, the vast majority of these anomalies are natural in origin and, thus, do not justify alarm.

When using the TIF, as the most data-driven approach to identify threshold values, the lowest number of exceedances of the two thresholds for northern and southern Europe is observed for most elements, and follow-up is easily accomplished (see Fig. 3b and Fig. 3SM for all maps). As for exceedances above the 95th and 98th percentiles, most of the anomalies are related to geology. No large-scale anthropogenic process requiring continental-scale attention emerges. Impact from diffuse contamination on the distribution of elements at this scale, though it certainly exists, remains invisible against the large geochemical background variation.

Each of the 53 maps in Fig. 3SM shows the sites with an 'unusually high' concentration of the mapped element either at the European or at the northern and southern European scales. Independent of whether the high element concentration is caused by natural conditions (geology, mineralisation, climate) or an anthropogenic impact (cities, industry, power plants, WW I), all the sites needing additional attention become visible on these simple maps. Some authors, e.g., Baritz et al. (2014) or Tóth et al. (2016b) try to derive higher resolution maps via correlation/regression based models, introducing further high resolution data sets such as topography and climate. Such modelled maps look very impressive at first glance due to the high resolution inherited from the covariables used. These modelled maps may make sense for parameters like carbon, where the aim is to obtain a general overview. However, the aim of regional geochemistry in relation to trace elements has always been to find the sites with an unusual element concentration (whether high as discussed here or low). These most interesting sites likely will get lost in a correlation/regression-based approach to the production of high resolution maps. Consequently, the ability to detect the really important sites in a geochemical trace element data set is not improved.

The general message from the maps in Fig. 3 is clear: the use of a generic threshold is not appropriate for the identification of contamination in agricultural land at the continental scale. Contamination is not identified by exceedingly high PTE concentrations in agricultural soil. The highest element concentrations at the continental scale, observed in the GEMAS data set, are mainly due to geogenic causes. Contamination due to diffuse anthropogenic emissions has a relatively small contribution to the overall natural geochemical background, and its impact on the large natural variation of an element remains usually invisible. The largest effect of diffuse contamination on the overall element concentration in a soil sample should be expected to occur in the low concentration range (or at the local scale where the input may be massive and no longer 'diffuse'). One example of an unusual local source is the impact of shelling in the World War I on the Cu concentrations in the soil around Ypres (Van Meirvenne et al., 2008). Diffuse contamination, however, leads to an overall shift to slightly higher concentrations in the CP diagram. Thus, the affected locations cannot be identified by a simple statistical procedure or on maps.

Decades of geochemical mapping by the Geological Surveys of Europe at a variety of scales and resolutions have led to the realisation that serious soil contamination, resulting in concentrations exceeding the natural range in background concentrations, leads only locally to very high element concentrations. This is very similar to the effect of mineral occurrences, which cannot usually be identified at the continental scale. Many examples of local impact can be found in the results of urban geochemical mapping projects (Johnson et al., 2011). In general, the GEMAS project data provide the opportunity to link the countless local- to regional-

scale data sets that exist in various countries or jurisdictions to the European scale. The data set presented by Bednářová et al. (2016) is also of interest in this connection. These authors show results for more than 50,000 topsoil samples collected at the scale of the Czech Republic and conclude that the reliable detection and monitoring of anthropogenic contamination needs to be carried out at a very local scale. This would be prohibitively expensive (more than 5 million samples needed) at the scale of the European Union. The GEMAS results, in contrast, demonstrate that providing a reference network of samples at the European scale and mapping of the European background variation can be carried out with as little as 2000 samples.

5. Conclusions

The CP diagram, the 98th percentile, and the TIF emerge as the three methods that will deliver useful geochemical threshold values to identify locations with an unusually high element concentration, whatever the scale of a survey. It must be recognised, however, that completely different geochemical threshold values will be obtained for different areas. This reflects the wide diversity of geological conditions at the European scale. Such geochemical threshold values, based on the distribution of background concentrations, will not be able to separate anthropogenic contamination from naturally elevated element concentrations, due to the occurrence of mineral deposits or lithologies with high concentrations of certain elements (e.g., the high Co, Cu, Cr, Ni concentrations of ultramafic rocks). Furthermore, they will not provide any indication of potential toxicity risks to human health or the environment. In such cases, local site-specific risk assessment is needed, preferentially with thresholds derived from ecotoxicological data.

Maps of the threshold exceedances reveal that, for the vast majority of cases, the detected 'high values' have a natural (geology or climate) cause. There are only a few areas at the continental scale where a sizeable region is affected by high concentrations of PTEs. One such example is the Massif Central in France with high As concentrations. Other examples are the comparatively low, but high at the northern European scale, As values in soil of central Sweden; high values of elements like Cr and Ni related to the occurrence of mafic-ultramafic rocks in Hellas; the Mo anomaly marking the undeveloped Nordli Mo deposit in Norway and a belt of black shale in Scandinavia; the anomalies related to the occurrence of Terra Rossa soil along the Croatian coast, and the Hg, Pb and Se anomaly at the west coast of Norway. The latter is related to local climatic conditions and vegetation zones and not only to geology. Few anomalies due to anthropogenic sources can still be detected even at this scale and sample density. Examples are large cities (London: Ag, Hg; Paris: Ag, Hg, Pb; Kiev: Hg), agriculture (Cu in vineyards in southern Europe, intensive husbandry in NW Germany), and some, most often single point, anomalies, which are most likely related to the vicinity of power plants or industry.

In terms of follow-up at the European scale, the following measures can be recommended, based on the results of the GEMAS project:

- Fill in the GEMAS data set with a higher sample density around major cities and/or other known emission sources (power plants, industry, smelters). A few sample sites per location would probably be sufficient to make them stand out even on European-scale maps and to see their relative spatial impact on the background variation provided by the original GEMAS data set.
- Link and level the many high-density local soil geochemical surveys, which the Geological Surveys of the different member

states have carried out, to the GEMAS data set as the European scale reference.

- A significant contribution from diffuse pollution at the continental scale will become visible at sites with a low natural concentration in the element in question. Monitoring of diffuse pollution at the European scale should thus be focussed on a few such sites.
- To study element uptake and cycling between the different spheres of the environment, a number of detailed high sample density studies in selected small catchments across Europe, representative for the variable natural conditions (or including a contamination source), based on multi-media sampling (rock, soil, water, plants), and multi-element analyses, could be used to substantially improve our understanding of natural conditions versus anthropogenic effects.

Acknowledgements

The GEMAS project is a cooperation project of the EuroGeoSurveys Geochemistry Expert Group with a number of outside organisations (e.g., Alterra in The Netherlands, the Norwegian Forest and Landscape Institute, several Ministries of the Environment and University Departments of Geosciences in a number of European countries, CSIRO Land and Water in Adelaide, Australia) and Eurometaux. Sampling was covered by the participating organisations in their home country, the analytical work was co-financed by the following organisations: Eurometaux, Cobalt Development Institute (CDI), European Copper Institute (ECI), Nickel Institute, Europe, European Precious Metals Federation (EPMF), International Antimony Association (i2a), International Manganese Institute (IMnI), International Molybdenum Association (IMoA), ITRI Ltd. (on behalf of the REACH Tin Metal Consortium), International Zinc Association (IZA), International Lead Association-Europe (ILA-Europe), European Borates Association (EBA), the (REACH) Vanadium Consortium (VC) and the (REACH) Selenium and Tellurium Consortium. The Geological Survey of Norway managed the GEMAS project with financial support from EuroGeoSurveys. The Directors of the European Geological Surveys and the additional participating organisations are thanked for enabling sampling of almost all of Europe in a tight time schedule. The authors thank David Smith and an anonymous reviewer for a thorough and insightful review of the manuscript, resulting in many improvements.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.apgeochem.2017.01.021>.

References

- Abedin, M.J., Feldmann, J., Meharg, A.A., 2002. Uptake kinetics of arsenic species in rice plants. *Plant Physiol.* 128, 1120–1128.
- Aitchison, J., 1986. *The Statistical Analysis of Compositional Data*. Chapman & Hall, London, p. 416.
- Ander, E.L., Johnson, C.C., Cave, M.R., Palumbo-Roe, B., Nathanail, C.P., Lark, R.M., 2013. Methodology for the determination of normal background concentrations of contaminants in English soil. *Sci. Total Environ.* 454–455, 604–618.
- Baritz, R., Ernstsen, V., Zirlewagen, D., 2014. Carbon concentrations in European agricultural and grazing land soil. In: Reimann, C., Birke, M., Demetriades, A., Filzmoser, P., O'Connor, P. (Eds.), *Chemistry of Europe's Agricultural Soils – Part B: General Background Information and Further Analysis of the GEMAS Data Set*. Geologisches Jahrbuch (Reihe B103). Schweizerbarth, Hannover, pp. 117–129.
- Bednářová, Z., Kalina, J., Hájek, O., Sánka, M., Komprdová, K., 2016. Spatial distribution and risk assessment of metals in agricultural soils. *Geoderma* 284, 113–121.
- Birke, M., Reimann, C., Oorts, K., Rauch, U., Demetriades, A., Dinelli, E., Ladenberger, A., Halamic, J., Gosar, M., Jähne-Klingberg, F., GEMAS Project Team,

2016. Use of GEMAS data for risk assessment of cadmium in European agricultural and grazing land soil under the REACH Regulation. *Appl. Geochem.* 74, 109–121.
- Blum, A., Chery, L., Barbier, J., Baudry, D., Petelet-Giraud, E., 2002. Rapport final, Rapport BRGM RP-51549-FR, 5 volumes (in French). Contribution à la caractérisation des états de référence géochimique des eaux souterraines. Outils et méthodologie.
- Carlou, C., D'Alessandro, M., Swartjes, F., 2007. Derivation Methods of Soil Screening Values in Europe. A Review and Evaluation of National Procedures towards Harmonization. EUR 22805 EN. European Communities, Luxembourg.
- Cave, M.R., Johnson, C.C., Ander, E.L., Palumbo-Roe, B., 2012. Methodology for the determination of normal background contaminant concentrations in english soils. *Br. Geol. Surv. Comm. Rep. CR/12/003*. Available at: <http://nora.nerc.ac.uk/19959/>.
- Chaffee, M.A., 1977. Some thoughts on the selection of threshold values as practiced in the branch of exploration research of the U.S. Geological Survey. *Explore* No. 21, 14–16. Available at: <https://www.appliedgeochemists.org/images/Explore/Explore.%20Number%2021%20January%201977.pdf>.
- Darnley, A.G., Björklund, A., Bølviken, B., Gustavsson, N., Koval, P.V., Plant, J.A., Steinfeld, A., Tauchid, M., Xuejing, X., Garrett, R.G., Hall, G.E.M., 1995. A Global Geochemical Database for Environmental and Resource Management. Final Report of IGBP Project 259. Earth Sciences. 19. UNESCO Publishing, Paris, p. 122. Available at: http://www.globalgeochemicalbaselines.eu/wp-content/uploads/2012/07/Blue_Book_GGD_IGCP259.pdf.
- De Vos, W., Tarvainen, T., Salminen, R., Reeder, S., De Vivo, B., Demetriades, A., Pirc, S., Batista, M.J., Marsina, K., Ottesen, R.T., O'Connor, P.J., Bidovec, M., Lima, A., Siewers, U., Smith, B., Taylor, H., Shaw, R., Salpeteur, I., Gregorauskiene, V., Halamic, J., Slaninka, I., Lax, K., Gravesen, P., Birke, M., Breward, N., Ander, E.L., Jordan, G., Duris, M., Klein, P., Locutura, J., Bel-Lan, A., Pasieczna, A., Lis, J., Mazreku, A., Gilucis, A., Heitzmann, P., Klaver, G., Petersell, V. (Eds.), 2006. Geochemical Atlas of Europe. Part 2—Interpretation of Geochemical Maps, Additional Tables, Figures, Maps and Related Publications. Geological Survey of Finland, Espoo, Finland, p. 692. Available at: <http://weppi.gtk.fi/publ/foregsatlas/part2.php>.
- DEC (Department of Environment and Conservation, Western Australia), 2010. Contaminated Sites Management Series – Assessment Levels for Soil, Sediment and Water. Version 4, Revision, vol. 1, p. 53. Available at: www.dec.wa.gov.au/contaminatedsites.
- Demetriades, A., 2011. Hazard and exposure assessment in contaminated land investigations and environmental management. Chapter 11. In: Johnson, C.C., Demetriades, A., Locutura, J., Ottesen, R.T. (Eds.), *Mapping the Chemical Environment of Urban Areas*. Wiley-Blackwell, John Wiley & Sons Ltd, Chichester, U.K., pp. 135–172.
- Dutter, R., Leitner, T., Reimann, C., Wurzer, F., 1992. Grafische und geostatistische Analyse am PC. Beiträge zur Umweltstatistik. Schriftenr. Tech. Univ. Wien Bd 29, 78–88. Esri, 2016. ArcGIS for Desktop. Available at: <http://www.esri.com/software/arcgis/arcgis-for-desktop>.
- EC (European Commission), 2006. Regulation (EC) No 1907/2006 of the European parliament and of the council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. *Official J. Eur. Communities*, 30.12.2006, L396, 1–849.
- ECHA (European Chemicals Agency), 2008. Guidance on Information Requirements and Chemical Safety Assessment, Appendix R.7.13–2: Environmental Risk Assessment for Metals and Metal Compounds. European Chemicals Agency, Helsinki, Finland. Available at: https://echa.europa.eu/documents/10162/13632/information_requirements_r7_13_2_en.pdf/0497e68d-4bb5-4b12-a4db-52ce0c1bc237.
- ECHA (European Chemicals Agency), 2016. Guidance on Information Requirements and Chemical Safety Assessment. Chapter R.16: Environmental Exposure Assessment. Available at: https://echa.europa.eu/documents/10162/13632/information_requirements_r16_en.pdf/b9f0f406-ff5f-4315-908e-e5f83115d6af.
- Economou-Eliopoulos, M., Vacondos, I., 1995. Geochemistry of chromitites and host rocks from the pindos ophiolite complex, Greece. *Chem. Geol.* 122, 99–108.
- Economou-Eliopoulos, M., Tarkian, M., Sambanis, G., 1999. On the geochemistry of chromitites from the pindos ophiolite complex, Greece. *Chem. Erde* 59, 19–31.
- EGS (EuroGeoSurveys Geochemistry Working Group), 2008. Geochemical Mapping of Agricultural and Grazing Land in Europe (GEMAS) – Field Manual. Norges Geologiske Undersøkelse Report, 2008.038, 46 pp. Available at: http://www.ngu.no/upload/Publikasjoner/Rapporter/2008/2008_038.pdf.
- ESRI (Environmental Systems Research Institute), 2016. ArcGIS for Desktop. Available at: <http://www.esri.com/software/arcgis/arcgis-for-desktop>.
- Filzmoser, P., 2015. Package 'StatDA' Version 1.6.9. The Comprehensive R Archive Network (CRAN). Available at: <https://cran.r-project.org/web/packages/StatDA/index.html>.
- Filzmoser, P., Hron, K., Reimann, C., 2010. Bivariate statistical analysis of environmental (compositional) data. *Sci. Total Environ.* 408, 4230–4238.
- Frank, A., 1998. 'Mysterious' moose disease in Sweden. Similarities to copper deficiency and/or molybdenosis in cattle and sheep. Biochemical background of clinical signs and organ lesions. *Sci. Total Environ.* 209, 17–26.
- Hawkes, H.E., 1957. Principles of geochemical prospecting. *U.S. Geol. Surv. Bull.* 100-F, 225–355. Available at: <https://pubs.usgs.gov/bul/1000f/report.pdf>.
- Hawkes, H.E., Webb, J.S., 1962. *Geochemistry in Mineral Exploration*. Harper, New York, 415 pp.
- ISO (International Organization for Standardization), 2005. Soil quality — guidance on the determination of background values. *Int. Stand. ISO 19258 2005 (E)*, 24.
- Jarva, J., Tarvainen, T., Reinikainen, J., Eklund, M., 2010. TAPIR – Finnish national geochemical baseline database. *Sci. Total Environ.* 408, 4385–4395.
- Johnson, C.C., 2011. Understanding the quality of chemical data from the urban environment – Part 1: quality control procedures. Chapter 5. In: Johnson, C.C., Demetriades, A., Locutura, J., Ottesen, R.T. (Eds.), *Mapping the Chemical Environment of Urban Areas*. Wiley-Blackwell, John Wiley & Sons Ltd., Chichester, U.K., pp. 61–76.
- Johnson, C.C., Demetriades, A., Locutura, J., Ottesen, R.T. (Eds.), 2011. *Mapping the Chemical Environment of Urban Areas*. Wiley-Blackwell, Chichester, p. 516.
- Johnson, C.C., Ander, E.L., Cave, M.R., Palumbo-Roe, B., 2012. Normal background concentrations (NBCs) of contaminants in english soils: final project report. *Br. Geol. Surv. Comm. Rep. CR/12/035*, 40 pp. Available at: <http://nora.nerc.ac.uk/19946/>.
- Kloke, A., 1980. Orientierungsdaten für tolerierbare Gesamtgehalte einiger Elemente in Kulturböden. *Mitteilungen VDLUFA, Heft 1–3/1980* (in German).
- Kriete, C., 2011. Results of the GEMAS Ring Test as Supplied to the Participating Laboratories. BGR Report on the GEMAS Project Proficiency Test. Laboratory Assessment, Hannover, p. 57.
- Kürzl, H., 1988. Exploratory data analysis: recent advances for the interpretation of geochemical data. *J. Geochem. Explor.* 30, 309–322.
- Lepeltier, C., 1969. A simplified statistical treatment of geochemical data by graphical representation. *Econ. Geol.* 64, 538–550.
- Mackovych, D., Lucivjanský, P., 2014. Preparation of GEMAS project samples and standards. Chapter 4. In: Reimann, C., Birke, M., Demetriades, A., Filzmoser, P., O'Connor, P. (Eds.), *Chemistry of Europe's Agricultural Soils – Part a: Methodology and Interpretation of the GEMAS Data Set*. Geologisches Jahrbuch (Reihe B 102). Schweizerbart, Hannover, pp. 37–40.
- Mali, M., Dell'Anna, M.M., Mastroianni, P., Damiani, L., Ungaro, N., Belviso, C., Fiore, S., 2015. Are conventional statistical techniques exhaustive for defining metal background concentrations in harbour sediments? A case study: the coastal area of Bari (southeast Italy). *Chemosphere* 138, 708–717.
- Marin, A.R., Masscheleynm, P.H., Patrick Jr., W.H., 1993. Soil redox-pH stability of arsenic species and its influence on arsenic uptake by rice. *Plant Soil* 152, 245–253.
- Matschullat, J., Ottenstein, R., Reimann, C., 2000. Geochemical background – can we calculate it? *Environ. Geol.* 39, 990–1000.
- MEF (Ministry of the Environment, Finland), 2007. Government Decree on the Assessment of Soil Contamination and Remediation Needs. 214/2007 (March 1, 2007) – in Finnish and Swedish.
- Oorts, K., Schoeters, I., 2014. Use of monitoring data for risk assessment of metals in soil under the European REACH regulation. In: Reimann, C., Birke, M., Demetriades, A., Filzmoser, P., O'Connor, P. (Eds.), *Chemistry of Europe's Agricultural Soils – Part B: General Background Information and Further Analysis of the GEMAS Data Set*. Geologisches Jahrbuch (Reihe B103). Schweizerbart, Hannover, pp. 189–202.
- Oorts, K., Smolders, E., McGrath, S.P., Van Gestel, C.A.M., McLaughlin, M.J., Carey, S., 2016. Derivation of ecological standards for risk assessment of molybdate in soil. *Environ. Chem.* 13, 168–180.
- Pedersen, F.D., 1986. An outline of the geology of the Hurdal area and the Nordli granite molybdenite deposit. In: Olerud, S., Ihlen, P.M. (Eds.), *Metallogeny Associated with the Oslo Paleorift*, vol. 59. Sveriges Geologiska Undersökning, pp. 18–25.
- Piteå Kommun, 2005. Arsenik I Dricksvatten I Piteå Kommun. Lägerapport, Augusti 2005, p. 7 (in Swedish).
- Provoost, J., Cornelis, C., Swartjes, F., 2006. Comparison of soil clean-up standards for trace elements between countries: why do they differ? *J. Soils Sediments* 6, 173–181.
- Rassios, A., Smith, A.G., 2000. Constraints on the formation and emplacement age of western Greek ophiolites (Vourinos, Pindos and Othris) inferred from deformation structures in peridotites. In: Dilek, Y., Moores, E.M., Elthon, D., Nicolas, A. (Eds.), *Ophiolites and Ocean Crust: New Insights from Field Studies and the Ocean Drilling Program*. Geological Society of America Special Papers, vol. 349, pp. 473–483.
- Reimann, C., 2007. Hintergrund und Schwellenwert: unbestimmbare Größen? *Jahrb. Geol. Bundesanst. Wien* 146, 243–253 (in German with English summary).
- Reimann, C., Caritat, P. de, 2017. Establishing background variation and threshold values for 59 elements in Australian surface soil. *Sci. Total Environ.* 578, 633–648.
- Reimann, C., Filzmoser, P., 2000. Normal and lognormal data distribution in geochemistry: death of a myth. Consequences for the statistical treatment of geochemical and environmental data. *Environ. Geol.* 39, 1001–1014.
- Reimann, C., Garrett, R.G., 2005. Geochemical background – concept and reality. *Sci. Total Environ.* 350, 12–27.
- Reimann, C., Siewers, U., Tarvainen, T., Bityukova, L., Eriksson, J., Gilucis, A., Gregorauskiene, V., Lukashev, V.K., Matinian, N.N., Pasieczna, A., 2003. *Agricultural Soils in Northern Europe: a Geochemical Atlas*. Geologisches Jahrbuch, Sonderhefte, Reihe D, Heft SD 5. Schweizerbart'sche Verlagsbuchhandlung, Stuttgart, p. 279.
- Reimann, C., Garrett, R.G., Filzmoser, P., 2005. Background and threshold – critical comparison of methods of determination. *Sci. Total Environ.* 346, 1–16.

- Reimann, C., Filzmoser, P., Garrett, R.G., Dutter, R., 2008. *Statistical Data Analysis Explained. Applied Environmental Statistics with R*. Wiley, Chichester, p. 343.
- Reimann, C., Demetriades, A., Eggen, O.A., Filzmoser, P., the EuroGeoSurveys Geochemistry Expert Group, 2009a. The EuroGeoSurveys Geochemical Mapping of Agricultural and Grazing Land Soils Project (GEMAS) – Evaluation of Quality Control Results of Aqua Regia Extraction Analysis. Norges Geologiske Undersøkelse Report, 2009.049, 92 pp. Available at: http://www.ngu.no/upload/Publikasjoner/Rapporter/2009/2009_049.pdf.
- Reimann, C., Englmaier, P., Gough, L., Lamothe, P., Nordgulen, Ø., Smith, D., 2009b. Geochemical gradients in O-horizon soils of southern Norway: natural or anthropogenic? *Appl. Geochem.* 24, 62–76.
- Reimann, C., Matschullat, J., Birke, M., Salminen, R., 2009c. Arsenic distribution in the environment: the effects of scale. *Appl. Geochem.* 24, 1147–1167.
- Reimann, C., Birke, M., Filzmoser, P., 2011. Data analysis for urban geochemical data. Chapter 7. In: Johnson, C.C., Demetriades, A., Locutura, J., Ottesen, R.T. (Eds.), *Mapping the Chemical Environment of Urban Areas*. Wiley-Blackwell, John Wiley & Sons Ltd, Chichester, U.K, pp. 99–115.
- Reimann, C., Demetriades, A., Birke, M., Eggen, O.A., Filzmoser, P., Kriete, C., Euro-GeoSurveys Geochemistry Expert Group, 2012a. The EuroGeoSurveys Geochemical Mapping of Agricultural and Grazing Land Soils Project (GEMAS) – Evaluation of Quality Control Results of Particle Size Estimation by MIR Prediction, Pb-isotope and MMI[®] Extraction Analyses and Results of the GEMAS Ring Test for the Standards Ap and Gr. Geological Survey of Norway, Trondheim. NGU report 2012.051, 136 pp. Available at: http://www.ngu.no/upload/Publikasjoner/Rapporter/2012/2012_051.pdf.
- Reimann, C., Filzmoser, P., Fabian, K., Hron, K., Birke, M., Demetriades, A., Dinelli, E., Ladenberger, A., Gemas project team, 2012b. The concept of compositional data analysis in practise – Total major element concentrations in agricultural and grazing land soils of Europe. *Sci. Total Environ.* 426, 196–210.
- Reimann, C., Flem, B., Fabian, K., Birke, M., Ladenberger, A., Negrel, P., Demetriades, A., Hoogewerff, J., Gemas project team, 2012c. Lead and lead isotopes in agricultural soils of Europe – the continental perspective. *Appl. Geochem.* 27, 532–542.
- Reimann, C., Birke, M., Demetriades, A., Filzmoser, P., O'Connor, P., 2014a. Chemistry of Europe's Agricultural Soils, Part a: Methodology and Interpretation of the GEMAS Data Set, *Geologisches Jahrbuch Reihe B, Band B 102*. Schweizerbart Science Publishers, Stuttgart, p. 523.
- Reimann, C., Birke, M., Demetriades, A., Filzmoser, P., O'Connor, P., 2014b. Chemistry of Europe's Agricultural Soils, Part B: General Background Information and Further Analysis of the GEMAS Data Set, *Geologisches Jahrbuch Reihe B, Band B 103*. Schweizerbart Science Publishers, Stuttgart, p. 352.
- Reimann, C., Ladenberger, A., Birke, M., Caritat, P. de, 2016. Low density geochemical mapping and mineral exploration: application of the mineral system concept. *Geochem. Explor. Environ. Anal.* 16, 48–61.
- Rothwell, K.A., Cooke, M.P., 2015. A comparison of methods used to calculate normal background concentrations of potentially toxic elements for urban soil. *Sci. Total Environ.* 532, 625–634. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0048969715302783>.
- Saccani, E., Beccaluva, L., Photiades, A., Zeda, O., 2011. Petrogenesis and tectonomagmatic significance of basalts and mantle peridotites from the Albanian–Greek ophiolites and sub-ophiolitic mélanges. New constraints Triassic–Jurassic Evol. Neo-Tethys Dinaride Sect. *Lithos* 124, 227–242.
- Salminen, R., Batista, M.J., Bidovec, M., Demetriades, A., De Vivo, B., De Vos, W., Gilucis, A., Gregorauskiene, V., Halamic, J., Heitzmann, P., Lima, A., Jordan, G., Klaver, G., Klein, P., Lis, J., Locutura, J., Marsina, K., Mazreku, A., Mrnkova, J., O'Connor, P.J., Olsson, S., Ottesen, R.-T., Petersell, V., Plant, J.A., Reeder, S., Salpeteur, I., Sandström, H., Siewers, U., Steenfelt, A., Tarvainen, T., 2005. FOREGS Geochemical Atlas of Europe, Part 1 – Background Information, Methodology, and Maps. Geological Survey of Finland, Espoo, p. 690. Available at: <http://weppi.gtk.fi/publ/foregsatlas/>.
- Scealy, J.L., Caritat, P. de, Grunsky, E.C., Tsagris, M.T., Welsh, A.H., 2015. Robust principal component analysis for power transformed compositional data. *J. Am. Stat. Assoc.* 110, 136–148.
- Scheib, A.J., Birke, M., Dinelli, E., GEMAS project team, 2014. Geochemical evidence of aeolian deposits in European soils. *Boreas* 43 (1), 175–192.
- Sinclair, A.J., 1974. Selection of threshold values in geochemical data using probability graphs. *J. Geochem. Explor.* 3, 129–149.
- Sinclair, A.J., 1976. Application of Probability Graphs in Mineral Exploration, Special vol. 4. Association of Exploration Geochemists, Toronto, p. 95.
- Sinclair, A.J., 1983. Univariate analysis. Chapter 3 In: R.J. Howarth (Editor), *Statistics and Data Analysis in Geochemical Prospecting*. Vol. 2 In: G.J.S. Govett (Series Editor), *Handbook of Exploration Geochemistry*. Elsevier, Amsterdam, 59–81.
- Sinclair, A.J., 1986. Statistical interpretation of soil geochemical data. In: W.K. Fletcher, S.J. Hoffman, M.B. Mehrtens, A.J. Sinclair, I. Thompson (Editors), *Exploration Geochemistry: Design and Interpretation of Soil Surveys*. Reviews in Economic Geology, Vol. 3, J.M. Robertson (Series Editor). Society of Economic Geologists, Chelsea, MI, USA, 97–115.
- Sinclair, A.J., 1991. A fundamental approach to threshold estimation in exploration geochemistry: probability plots revisited. *J. Geochem. Explor.* 41, 1–22.
- Smith, D.B., Wang, X., Reeder, S., Demetriades, A., 2012. The IUGS/IAGC task Group on global geochemical Baselines. *Earth Sci. Front.* 19 (3), 1–6.
- Smolders, E., Oorts, K., Van Sprang, P., Schoeters, I., Janssen, C.R., McGrath, S.P., McLaughlin, M.J., 2009. Toxicity of trace metals in soil as affected by soil type and aging after contamination: using calibrated bioavailability models to set ecological soil standards. *Environ. Toxicol. Chem.* 28, 1633–1642.
- Steinnes, E., Hanssen, J.E., Rambæk, J.P., Vogt, N.B., 1994. Atmospheric deposition of trace elements in Norway: temporal and spatial trends studied by moss analysis. *Water, Air, Soil Pollut.* 74, 121–140.
- Tennant, C.B., White, M.L., 1959. Study of the distribution of some geochemical data. *Econ. Geol.* 54, 1281–1290.
- Tóth, G., Jones, A., Montanarella, L., 2013. The LUCAS topsoil database and derived information on the regional variability of cropland topsoil properties in the European Union. *Environ. Monit. Assess.* 185, 7409–7425.
- Tóth, G., Hermann, T., Da Silva, M.R., Montanarella, L., 2016a. Heavy metals in agricultural soil of the European Union with implications for food safety. *Environ. Int.* 88, 299–309.
- Tóth, G., Hermann, T., Szatmári, G., Pásztor, L., 2016b. Maps of heavy metals in the soils of the European Union and proposed priority areas for detailed assessment. *Sci. Total Environ.* 565, 1054–1062.
- Tristán, E., Demetriades, A., Ramsey, M.H., Rosenbaum, M.S., Stavrakis, P., Thornton, I., Vassiliades, E., Vergou, K., 2000. Spatially resolved hazard and exposure assessments: an example of lead in soil at Lavrion, Greece. *J. Environ. Res.* 82 (1), 33–45.
- Tukey, J.W., 1977. *Exploratory Data Analysis*. Addison Wesley, Reading, p. 506.
- Van Meirvenne, M., Meklit, T., Verstraete, S., De Boever, M., Tack, F., 2008. Could shelling in the First World War have increased copper concentrations in the soil around Ypres? *Eur. J. Soil Sci.* 59, 372–379.
- Wolfram Research Inc, 2016. *Mathematica, Version 11.0* (Champaign, IL, USA).