Geoderma 148 (2008) 189–199

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

Heavy metals in European soils: A geostatistical analysis of the FOREGS Geochemical database

Luis Rodríguez Lado [a,*], Tomislav Hengl [b], Hannes I. Reuter [a]

[a] European Commission, Directorate General JRC, Institute for Environment and Sustainability, TP 280, Via E. Fermi 1, I-21020 Ispra (VA), Italy
[b] Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 20 March 2008
Received in revised form 16 September 2008
Accepted 22 September 2008

Keywords:
Soil mapping
Regression-kriging
MODIS
Night lights image
Geochemical database
Pan-European monitoring

ABSTRACT

This paper presents the results of modeling the distribution of eight critical heavy metals (arsenic, cadmium, chromium, copper, mercury, nickel, lead and zinc) in topsoils using 1588 georeferenced samples from the Forum of European Geological Surveys Geochemical database (26 European countries). The concentrations were mapped using regression-kriging (RK) and accuracy of predictions evaluated using the leave-one-out cross validation method. A large number of auxiliary raster maps (topographic indexes, land cover, geology, vegetation indexes, night lights images and earth quake magnitudes) were used to improve the predictions. These were first converted to 36 principal components and then used to explain spatial distribution of heavy metals. The study revealed that this database is suitable for geostatistical analyses: the predictors explained from 21% (Cr) to 35% (Pb) of variability; the residuals showed spatial autocorrelation. The Principal Component Analysis of the mapped heavy metals revealed that the administrative units (NUTS level3) with highest overall concentrations are: (1) Liege (Arrondissement) (BE), Attiki (GR), Darlington (UK), Coventry (UK), Sunderland (UK), Kozani (GR), Grevena (GR), Hartlepool & Stockton (UK), Huy (BE), Aachen (DE) (As, Cd, Hg and Pb) and (2) central Greece and Liguria region in Italy (Cr, Cu and Ni). The evaluation of the mapping accuracy showed that the RK models for As, Ni and Pb can be considered satisfactory (prediction accuracy 45–52% of total variance), marginally satisfactory for Cr, Cu, Hg and Zn (36–41%), while the model for Cd is unsatisfactorily accurate (30%). The critical elements limiting the mapping accuracy are: (a) the problem of sporadic high values (hot-spots); and (b) relatively coarse resolution of the input maps. Automation of the geostatistical mapping and use of auxiliary spatial layers opens a possibility to develop mapping systems that can automatically update outputs by including new field observations and higher quality auxiliary maps. This approach also demonstrates the benefits of organizing standardized joint European monitoring projects, in comparison to the merging of several national monitoring projects.

1. Introduction

The generalized mobilization and dispersion of pollutants from their natural reservoirs to the atmosphere, soil and water is one of the most significant negative impacts of human activities on terrestrial and aquatic ecosystems (Salemaa et al., 2001; Lin et al., 2002; Koptsik et al., 2003). Mining, iron and steel industry, road transport, waste incineration, and the use of fertilizers and agrochemicals are identified as the main human sources of heavy metals in soils and water in the superficial ecosystems (Hutton and de Meeûs, 2001; Hansen et al., 2002). In addition, emissions from volcanoes, degassing processes in the Earth's crust, forest fires or the chemical composition of the parent material can be also important sources of heavy metals in soils (Løkke et al., 1996; Palumbo et al., 2000).

A large proportion of soils in industrialized countries contain higher levels of several elements and compounds considered as pollutants than the corresponding natural background values in a pristine situation (Hijmans et al., 2005). Pollution caused by heavy metals is especially problematic in areas where synergy with other types of polluting agents exists. This is true in the case of industrial areas receiving large inputs of acidifying compounds, which creates optimal conditions for increased mobilization, bioavailability and thus toxicity of the metals stored in soils (Salemaa et al., 2001; Lin et al., 2002; Clemente et al., 2003; Cappuyns et al., 2004; de Vries et al., 2005).

In Europe, mandatory reductions on the annual emissions of cadmium, lead and mercury to avoid significant adverse effects on ecosystems have been put in the Heavy Metals Protocol (UN/ECE, 1998). This will have an increasing importance because, according to the most recent report of the Coordination Center for Effects (Posch et al., 2005), the distribution and magnitude of the deposition of these elements puts large areas of European ecosystems at risk both in 2000 and 2020. Apart from cadmium, lead and mercury, six additional

* Corresponding author. Tel.: +39 0332 789977; fax: +39 0332 786394.
E-mail addresses: luis.rodriguez-lado@jrc.it (L.R. Lado), hengl@science.uva.nl (T. Hengl), hannes.reuter@jrc.it (H.I. Reuter).

0016-7061/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.geoderma.2008.09.020

heavy metals (chromium, nickel, copper, zinc, arsenic and selenium) present an increasing threat for human health and the environment.

Recently, the European Commission (EC) has been preparing a proposal for a framework Directive (European Communities, 2006) that sets out common principles for the protection of soils across the EC. From the aspect of soil pollution, three main questions need to be clarified: (i) which threshold values should be used to classify soil as polluted? (ii) at which locations can high natural background values of heavy metals be expected? and (iii) which methods should be used to explain the spatial distribution of these elements in soils of Europe?

Threshold values for soils are difficult to evaluate since the toxicity and bioavailability of heavy metals is not only dependent on the total content in soils but also on many other environmental variables. Although some European countries (Netherlands, UK) developed their own soil quality standards for some heavy metals, at European level only threshold values related to the application of sewage sludge in agricultural soils have been defined (EU Directive 86/278/EC). The determination of natural background values is controversial because they can decrease the responsibility of human activities for the overall pollution on soils (Baize and Sterckeman, 2001).

At pan-European scale, it is often difficult to determine the background values that would correspond to a pristine situation since the geochemistry of most of our ecosystems is greatly influenced by a long history of anthropic activities, and even the concept of a background value is often fuzzily defined (Reimann and de Caritat, 2005; Reimann and Garrett, 2005). To improve this situation, in 1993 the International Union of Geological Sciences and the International Association of Geochemistry and Cosmochemistry (IUGS/IAGC) initiated the Global Geochemical Baselines Programme with an aim to establish the geochemical reference baselines for a number of elements at a global scale (Salminen, 2006). The European contribution to this International Programme started in 1997 and it has been carried out by governmental institutions of 26 European countries under the advisement of the Forum of European Geological Surveys (FOREGS). The final product of this collaboration was the "Geochemical Atlas of Europe" (http://www.gsf.fi/publ/foregsatlas/) that includes a database of about three thousand samples for solid media: topsoil, floodplain sediment, stream sediment and humus (ranging from 385 samples of sub soil to 852 samples of stream sediment) in addition to a number of maps of the element contents in soils of Europe based on the former database. The term "geochemical baseline" in this context is not equal to the "background value" since it represents a measurement that does not correspond to a pristine situation. However, these geochemical baselines can provide a perspective on the present status of pollution of the Europeans soils, and serve as a model for future pan-European monitoring projects. Some recent analyses of the FOREGS database can be seen in Imrie et al. (2008).

So far, several attempts were made to determine the spatial distribution of the concentrations of heavy metals in European soils. Reimann et al. (2003) created maps of heavy metal contents in soils using an Inverse Distance Weighted interpolator on about 740 samples from agricultural fields. The Geochemical Atlas of Eastern Barents region (Salminen et al., 2004) also includes interpolated maps of heavy metals from 1358 sampling sites in the northern part of Europe. Gawlik and Bidoglio (2006) produced maps for Cd, Cr, Cu, Hg, Ni, Pb and Zn in 11 European countries by using empirical relationships with soil parent material and land use. The European Environmental Agency (2006) merged the sampling points from three different soil databases to create a map of the concentration of lead in topsoils across Europe (http://dataservice.eea.europa.eu).

In principle, the only official maps of heavy metals in soils of Europe are those presented in the "Geochemical Atlas of Europe". This atlas contains maps for 85 variables for five different media: floodplain sediment, humus, soil, stream sediment and stream water, all produced using the Alkemia Smooth interpolation on a 6 km grid (de Vos and Tarvainen, 2006). Although this work represents an important contribution, we identified several points that could be considered to improve the reliability and accuracy of the final maps produced. Firstly, the authors of the FOREGS atlas did not consider merging the samples taken for different media or running more sophisticated geostatistical analyses to generate the maps. Secondly, they did not mask out areas that were not represented in a specific sample (histosols, water bodies), which limits the further uses of such maps. Moreover, the FOREGS atlas shows multiple pan-European maps of the same variable, which might be confusing for decision makers and spatial modelers. For example, there are 10 maps of lead distribution in the book: in subsoil (two laboratory techniques), in topsoil (two laboratory techniques), in humus, in stream water, in stream sediment (two laboratory techniques) and in floodplain sediment (two laboratory techniques). All 10 maps show values of the same element (often with different patterns) over the whole of Europe. Thirdly, because no transformation has been applied to the original data (most of the heavy metal concentrations show a skewed distribution), the resulting patterns were very much influenced by locally high values, which questions the overall reliability of the produced maps (see e.g. Papritz et al., 2005; Romić et al., 2007 for discussion). Fourthly, the interpolation method used to build the maps did not provide an estimate of its uncertainty.

With this background, we decided to test a geostatistical, digital soil mapping framework to interpolate the concentration of heavy metals over broad areas and to provide information about the uncertainty of the produced maps. Our assumption was that auxiliary predictors can be used to improve the detail and accuracy of the existing FOREGS atlas maps and extend the spatial analysis of the point data using *state-of-the-art* geostatistical techniques. We also wanted to demonstrate the benefits of producing geoinformation through a joint pan-European project and to suggest ways to design soil monitoring networks for the permanent monitoring of this soil threat.

2. Materials and methods

2.1. The FOREGS dataset

The point dataset was obtained from the FOREGS website, courtesy of the Association of the Geological Surveys of The European Union. For the purpose of this study, we used only the laboratory measurements of extractable heavy metal concentrations (HMC) in top-soil and floodplains determined for As, Cd, Cr, Cu, Ni, Pb and Zn (mg kg^{-1}) by ICP-AES using the Aqua Regia method (ISO, 1995) and for Hg (mg kg^{-1}), directly on the soil solid samples using a cold vapour absorption technique within an Advanced Mercury Analyzer (AMA-254, ALTEC). We did not include the stream sediment measurements in the analyses because the geochemical processes and accumulation in the water environment are different. At five locations the precision of the coordinates was not sufficient for separation of sampling locations. At these locations values for topsoil and floodplain were averaged and replaced the original value. The FOREGS database is rather extensive (85 variables measured in five media), so we have focused on mapping a selection of variables in the database. In the near future, we anticipate that also other variables from this database will be mapped with an improved spatial and thematic detail.

The total number of points used in this exercise was 1588, although not all values were available at all locations, ranging from 17 missing values for As, Cd, Cu, Ni, Pb and Zn to 79 missing values for Hg. The original coordinates of the sampling points have been provided in degrees, rounded to 0.01°, which corresponds to a horizontal precision of 850 m at 40° North latitude, and transformed to European Terrestrial Reference System (see www.euref.eu). The reorganized point database and all input and derived maps shown in this paper can be accessed from the http://eusoils.jrc.it website.