International Journal of Coal Geology 112 (2013) 173–179

Contents lists available at SciVerse ScienceDirect

International Journal of Coal Geology

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

Spatiotemporal risk assessment of soil pollution in a lignite mining region using a Bayesian maximum entropy (BME) approach

K. Modis [a,*], K.I. Vatalis [b], Ch. Sachanidis [b]

[a] *School of Mining and Metallurgical Engineering, National Technical University of Athens, Greece*
[b] *Dept of Geotechnology & Environmental Engineering, Technological Educational Institution of W. Macedonia, Kozani, Greece*

ARTICLE INFO

Article history:
Received 3 July 2012
Received in revised form 16 November 2012
Accepted 16 November 2012
Available online 7 December 2012

Keywords:
Spatiotemporal geostatistics
Soil pollution
Risk assessment
BME

ABSTRACT

The present paper aims to map pollution and assess the risk for agricultural soils in a wider lignite opencast mining and industrial area. Geochemical data related to environmental studies show that the waste characteristics favor solubilization and mobilization of inorganic contaminants and in some cases the generation of acidic leachates.

The spatiotemporal distribution of soil contamination is studied by the application of the Bayesian Maximum Entropy (BME) theory which allows merging spatial and temporal estimations in a single model.

Results reveal a correlation range of contaminant concentrations up to 5000 m and indicate a potential forecasting range up to 4 years. Inspection of the produced spatiotemporal maps indicates that the whole study area is contaminated by As and various heavy metals, a situation which seems to be more or less stable over time.

1. Introduction

Lignite, which is a poor quality coal with properties intermediate to those of bituminous coal and peat, is the only significant domestic fossil fuel in Greece. There are significant resources of lignite, located mostly in the northern part of the country, in the region of Western Macedonia. Total exploitable lignite reserves are estimated at about 4 billion tons (Kavouridis, 2008).

Usually, activities associated with lignite and coal mining and beneficiation result in the production of huge volumes of reactive solid wastes and the subsequent generation of acidic leachates containing heavy metals and metalloids causing widespread contamination of soil, surface- and groundwater (Gläβer et al., 2011; Komnitsas et al., 1998; Schroeter and Gläβer, 2011). During lignite combustion, the major, minor and trace elements may concentrate in the fly ash while the mineral matter undergoes a series of physical and chemical changes (Filippidis and Georgakopoulos, 1992). During the past, solid wastes not only have been destroying and occupying productive lands in intensive mining areas, but also brought up a chain of environmental problems related to air and underground water pollution.

Today, the environmental protection of soils is based on critical limits or Environmental Quality Standards expressed as soil concentrations. There are many sorts of pollutants in soils of mining areas, such as heavy metals. Regarding soils, the degree of contamination and the resulting "hazard index" may vary when different thresholds, set only in a limited number of countries, are considered. For example, Canadian guidelines classify contaminated soils according to the land use (agricultural, industrial, and residential) whereas The Netherlands guidelines provide background, target and intervention values. Target values define protective levels while intervention values are indicative of serious soil contamination. UK guidelines of Interdepartmental Committee on the Redevelopment of Contaminated Land, (ICRCL) concern the restoration and aftercare of metalliferous mining sites for pasture and grazing. These guidelines focus on toxic elements and trigger concentrations and are acceptable only when soil contamination is derived from mine spoil (ICRCL, 1990). Based on the analysis of the soils in Russia, Snakin et al. (1996) developed a set of five representative soil quality indicators to assess the physical, chemical, and biological degradation of soil, concerning various anthropogenic uses of the land.

Taking into account that the magnitude of the hazard caused by certain chemical elements depends on parameters such as the concentration, chemical form, particle size, soil or water pH and extent of exposure to the element, a set of different ranging values or pollution classes has been setup in the last years by various researchers, in order to identify sources of soil pollution due to human influence, weathering process and industrial or mining activities (Adriano, 1986; Alloway, 1990; Clarke and Sloss, 1992; EPA, 1992; FAO, 1992; Kabata-Pendias and Pendias, 1992; Swaine, 1990; Swaine and Goodarzi, 1995).

The assessment of the degree of soil and water contamination in wider mining and waste disposal sites is in general a complex process and involves calculation of ecological or health risk in mining, waste disposal, industrial, agricultural or residential areas as well as in ecosystems. The conventional methodology used is based on the principle

* Corresponding author.
E-mail address: kmodis@mail.ntua.gr (K. Modis).

0166-5162/$ – see front matter
http://dx.doi.org/10.1016/j.coal.2012.11.015

"source–pathway–target" and accounts for spatial and temporal variability of contaminant patterns. A probabilistic assessment incorporates variability of parameters and uncertainty in measurements (Komnitsas and Modis, 2006). Geostatistics is used to predict the extent of soil and groundwater contamination as well as to calculate the risk in active or abandoned mining, waste disposal and urban sites, by accounting for the spatial distribution and uncertainty of the estimates. It facilitates quantification of the spatial features of soil parameters and enables spatial interpolation (Hristopulos, 2003; Journel and Huijbregts, 1978; Modis and Komnitsas, 2007; Modis et al., 2010). The application of BME theory (Christakos, 2000) allows the simultaneous study of spatial and temporal systematic correlations.

2. Related notions from BME theory

A Random Function (RF) can be viewed as a collection of correlated random variables, say, $\mathbf{x}_{map} = (x_1, \ldots, x_m, x_k)$ at the points $\mathbf{p}_{map} = (\mathbf{p}_1, \ldots, \mathbf{p}_m, \mathbf{p}_k)$, where the symbol $\mathbf{x}_{map}$ is adopted because the goal is to obtain maps displaying estimates at points $\mathbf{p}_k$ of the unknown values χ_k of the natural variable from its observed values $\chi_1, \ldots, \chi_m$. A realization of the RF at these points is denoted by the vector $\boldsymbol{\chi}_{map} = (\chi_1, \ldots, \chi_m, \chi_k)$. If available, the complete characterization of a RF is provided by the multivariate probability density function (pdf) f_G defined as

$$\text{Prob}[\chi_1 \le x_1 \le \chi_1 + d\chi_1, \ldots, \chi_m \le x_m \le \chi_m + d\chi_m, \chi_k \le x_k \le \chi_k + d\chi_k] = f_G(\chi_{map}) d\chi_{map}$$

where the subscript G denotes the general knowledge base (Christakos, 2000) used to derive the pdf.

A generally incomplete but in many practical applications satisfactory characterization of this RF is provided by a limited set of statistical moments g_α, which in general are defined as:

$$\overline{g_a(\mathbf{x}_{map})} = \overline{g_a}(p_{map}) = \int g_a(\boldsymbol{\chi}_{map}) f_G(\boldsymbol{\chi}_{map}) d\boldsymbol{\chi}_{map} \quad (1)$$

where a = 0, 1, …, N; the overbar denotes stochastic expectation.

At the Prior Stage of the BME analysis, the general knowledge base G is processed before any specificatory knowledge is taken into account. Given G, the information contained in the map $\mathbf{x}_{map}$ can be expressed as (Christakos, 2000):

$$\text{Info}_G[x_{map}] = -\log f_G(x_{map}). \quad (2)$$

Then, the expected information is given by:

$$\overline{\text{Info}_G[\mathbf{x}_{map}]} = -\int f_G(\boldsymbol{\chi}_{map}) \log f_G(\boldsymbol{\chi}_{map}) d\boldsymbol{\chi}_{map}. \quad (3)$$

The prior (G-based) pdf f_G of the map is obtained by maximizing the expected prior information in Eq. (3) subject to the physical constraints introduced by Eq. (1). In the usual case where only the first and second order moments are to be used (which is also the case in this work) the above process leads to the adoption of a Gaussian prior pdf.

At the Meta-Prior Stage, the specificatory knowledge base S is organized in appropriate quantitative forms that can be explicitly incorporated into the BME formulation. For purposes of modern geostatistics, the data sets $\boldsymbol{\chi}_{data}$ available as specificatory knowledge are usually divided into two main categories

$$S: \chi_{data} = (\chi_{hard}, \chi_{soft}) = (x_1, \ldots, x_m) \quad (4)$$

where $\boldsymbol{\chi}_{hard}$ denotes "hard" data (i.e., accurate measurements obtained from real-time observations) and $\boldsymbol{\chi}_{soft}$ denotes "soft" data (uncertain observations expressed in terms of interval values or probability statements in general).

At the Posterior Stage of the BME analysis, the updated pdf f_K of the map should incorporate both the general and specificatory knowledge bases, where $K = G \cup S$.

In the case of hard data, the posterior pdf at the estimation point χ_k is calculated by simply substituting the data values in f_G.

In case of soft or interval data, the posterior distribution is given by

$$f_K(\chi_k) = A^{-1} \int_I f_G(\boldsymbol{\chi}_{map}) d\boldsymbol{\chi}_{soft} \quad (5)$$

where $A = \int f_G(\boldsymbol{\chi}_{data}) d\boldsymbol{\chi}_{soft}$ is the normalization parameter and the I-domain corresponds to the subset of data points p_i at which interval (soft) data rather than actual observations are available.

Since the posterior pdf is rigorously determined through the previously described analysis, a large number of options become available depending on the characteristics of the application considered. The most usual choice is the map that maximizes the posterior pdf and referred to as the "BMEmode" map, $\widehat{\chi}_{k,\ mode}$, which formally is calculated at each point p_k by solving the equation

$$\frac{\partial}{\partial \chi_k} f_K(\chi_k)\Big|_{\chi_k = \widehat{\chi}_k} = 0. \quad (6)$$

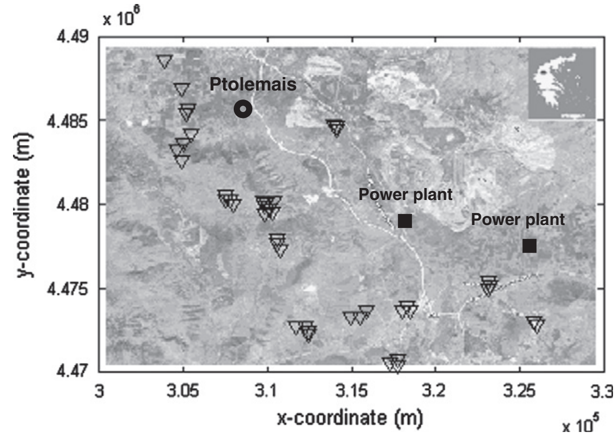

Fig. 1. Map of the study area in the Ptolemais lignite basin, showing the sampling locations and two power plants.

3. Site description

The area under study covers a 600 km^2 (30 × 20 km) field belonging to the wider lignite mining and waste disposal region of Ptolemais, 600 km north of Athens, Greece (Fig. 1).

The main industrial activities are situated in this basin. More than 150,000 people live and work in the area, mainly in the towns of Kozani and Ptolemais. There are also several villages with populations ranging from a few hundred to several thousands of inhabitants. The economic and industrial activity of the area is centered on its lignite

Table 1
Arithmetic means, ranges and tentative allowable concentrations (ppm) of chemical pollutants.

Pollutant	Mean	Stdev	Min	Median	Max	TAC
As	12.30	10.10	2.40	7.10	42.20	7.00
Cr	17.50	19.90	0.60	6.50	60.50	55.00
Ni	10.10	10.50	0.40	6.50	57.50	20.00
Bi	12.70	16.10	0.00	0.30	43.50	0.20