Journal of Petroleum Science and Engineering 78 (2011) 396–406

Contents lists available at ScienceDirect

Journal of Petroleum Science and Engineering

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

The example of background determination and mathematical processing of data from surface geochemical survey for the purposes of petroleum exploration

Henryk Sechman *, Marek Dzieniewicz

AGH University of Science & Technology, Faculty of Geology, Geophysics and Environment Protection, Department of Fossil Fuels, Al. Mickiewicza 30, 30-059 Kraków, Poland

ARTICLE INFO

Article history:
Received 15 February 2011
Accepted 6 June 2011
Available online 15 June 2011

Keywords:
surface geochemical survey
soil gas
background values
normal distribution
anomalous values
Poland

ABSTRACT

The paper reports on the methods of mathematical processing of the results of surface geochemical survey applied to petroleum exploration. The results include concentrations of methane analyzed in 1251 soil gas samples collected from the area over the Wierzchowice gas field (SW Poland). The methods were presented of determining the geochemical background, the anomalous threshold as well as of data normalization and filtration. The advantages were presented of iterative method for determination of background values proposed by the authors. In the authors' opinion, selection of the background should be preceded by statistical analysis of each set received during the iteration process. The method presented relies on the assumption of normal distribution of this dataset. The spatial distributions were compared of raw (measured) concentrations and mathematically process values. The applied mathematical data processing enables the researcher to eliminate the "noise" and to select objectively the anomalous values. Thus, such an attempt facilitates the comprehensive interpretation of geochemical data referred to geological model of a field. Results of the studies have demonstrated that the distribution of surface anomalies of methane over the Wierzchowice gas field has been disturbed by long-lasting exploitation. The anomalies appear mainly over dislocation zones, and in the SE part of the area they take the "halo" form around exploited portions of the field due to unexploited areas off the field.

1. Introduction

The fundamental assumption of surface geochemical methods is that petroleum fields are sources of dispersion of accumulated hydrocarbons. Light hydrocarbons escape from the reservoir in small amounts. Due to concentration and pressure gradients, these hydrocarbons migrate vertically towards the surface. Hence, the increased concentrations of gaseous hydrocarbons measured in the soil gas should disclose the presence of hydrocarbon accumulations at depths (see e.g., Abrams, 2005; Jones et al., 2000; Klusman, 1993; Matthews, 1996; McDermott, 1940; Schumacher, 1999; Sokolov and Grigoriev, 1962; Tedesco, 1995). However, the nature of gaseous hydrocarbons migration from deep accumulations to the surface is a complicated matter and depends on the properties of migrating compounds, on the character of the overburden and on the physiography of studied area. Consequently, the near-surface distribution pattern of gaseous hydrocarbons is a combined effect of deep accumulations and various disturbing factors. Hence, the proper geochemical data processing plays crucial role. This processing aims to select the sites of increased concentrations of gaseous hydrocarbons originating from petroleum fields at depths. The application of basic statistical methods enables the researcher to characterize quantitatively and qualitatively the measured values, i.e., to calculate basic statistical parameters for populations of measured concentrations and to analyze their distribution patterns. Such analysis aims to select and evaluate the identified, characteristic data populations. However, the anomalous results are defined by the reference level, which is represented by geochemical background.

The definition of geochemical background and the evaluation of its applicability to geochemical exploration and to environment protection were comprehensively discussed by many authors (see e.g., Cheng et al., 2000; Dzieniewicz and Mościcki, 1983; Galuszka, 2007; Garrett, 1991; Jones et al., 2000; Klusman, 1993; Lepeltier, 1969; Levinson, 1974; Li et al., 2003; Matschullat et al., 2000; Reimann and Garrett, 2005; Rice et al., 2002; Rose et al., 1979; Saunders et al., 1991; Sinclair, 1976; Sokolov and Grigoriev, 1962; Tedesco, 1995). It must be emphasized that, despite various methods of geochemical background determination and various applications of such data, all these authors underline the necessity of background determination if spatial variability of measured values is to be evaluated.

According to the definition of the State Geological Institute in Poland, geochemical background is the natural abundance of an element or a chemical compound in given environment. It is also named the "reference level" (of concentration) or the "average content". Hence, in traditional meaning, the geochemical background determines the range of values of an element or a chemical compound characteristic for given geological structure, region, province and/or

* Corresponding author. Tel.: +48 12 6173889; fax: +48 12 6336504.
E-mail address: sechman@agh.edu.pl (H. Sechman).

0920-4105/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.petrol.2011.06.001

country. Similar definitions can be found in the world literature (see e.g., Bates and Jackson, 1984; Galuszka, 2007; Garrett, 1991; Klusman, 1993; Matschullat et al., 2000; Matthews, 1996; Reimann and Garrett, 2005).

The positive or negative anomalies are concentrations which fall outside the lower or the upper threshold values (anomalous values). In the case of surface geochemical survey applied to petroleum exploration the subject of interest is increased concentrations (i.e., "positive anomalies") of hydrocarbons, which are understood as those exceeding the upper limit of fluctuations around the background value. The results of geochemical survey presented as distribution of anomalous values are commonly the output of digital processing of measurement results. Moreover, the determination of background value enables the researcher to refine the surface geochemical pattern by elimination of some disturbances which result from the effects of near-surface factors (Saunders et al., 1991; Sechman and Dzieniewicz, 2007).

The paper presents a method of mathematical processing of the results of surface geochemical survey based on dataset obtained from the measurements around the Wierzchowice gas field (SW Poland). Special attention was paid to the methods of determination of geochemical background and anomalous threshold values as well as to data normalization and filtration. The advantages were discussed of the iterative method of background determination modified by the authors. Finally, the spatial distributions of raw (measured) concentrations were compared with such distributions of mathematically processed results.

2. Analytical materials and methodology of field work and laboratory analyses

The methods of determination of background and anomalous values were presented using the dataset obtained from the surface geochemical survey of the Wierzchowice gas field (SW Poland). The studies were run in 1995 when the field was at final stage of its exploitation. Totally, 1251 samples of soil gases were collected from the sampling depth 1.2 m. The sampling pattern was a 200 × 200 m square grid (Fig. 1). The patented sampling procedure ensured the "sterile" sampling of soil gases, free of contamination from the air (Dzieniewicz and Sechman, 2001, 2002).

Soil gas samples were analyzed chromatographically for concentrations of methane and succeeding light homologues (ethane, propane, i-butane, n-butane) as well as gaseous alkenes (ethylene, propylene, 1-butene). Analyses were carried out at the laboratory of the Department of Fossil Fuels, Faculty of Geology, Geophysics and Environment Protection, AGH University of Science and Technology in Kraków, Poland.

Analyses were carried on with the FISSONS Instruments GC 8160 gas chromatograph applying the flame-ionization detector. The following analytical conditions were applied: metal column (internal diameter 4 mm, length 1.3 m) filled with the Activated Alumina (mesh 100/120), carrier gas (helium) flow rate 60 ml/min, programmed column temperatures: 90 °C for 3 min, 90–200 °C increment at 30 °C/min, 200 °C for 3 min, FID working temperature 270 °C, injection chamber working temperature 120 °C, volume of analyzed sample 1 ml.

The FID readings were processed with the WINNER software. Chromatograph was calibrated with the gas standards "Scott II" supplied by the Supelco and the Alltech. Analytical error was estimated as ± 2% of measured values.

The obtained results were applied to determination of geochemical background for the future underground gas storage (Gorski et al., 1999; Mularczyk and Kretschmar, 1999).

3. Methods of statistical procedure

Among the population of measured hydrocarbon concentrations the methane dataset was selected for a test of methods of mathematical data processing and visualization of the results.

3.1. Preliminary evaluation of measured concentrations

The preliminary evaluation included both the qualitative and quantitative characterization of measured concentrations. Practically, the histogram was constructed for measured methane concentrations and principal statistical parameters were determined: minimum and maximum values, arithmetic mean, standard deviation, median, first and third quartiles (Q_1, Q_3), and skewness.

The histogram is the simplest visualization method of distribution of obtained results. It allows the researcher to determine basic and anomalous populations of values. Analyses of hydrocarbon concentrations measured up to date for the purposes of surface geochemical survey revealed that most of data populations show the log-normal distributions with positive (right) skew (Dzieniewicz and Mościcki, 1983; Klusman 1993). Thus, for such data population the arithmetic mean does not reflect the average value of measured concentrations. Much better results are obtained by using the location parameters such as e.g., the median defined as a numeric value which separates the higher and the lower halves of values within given population. The median is the middle value of measured concentrations and is regarded by some authors as the background value (Reimann and Garrett, 2005). Other location parameters are 75th (third quartile — Q_3), 95th or 97, 5th percentile also accepted by some authors as a representation of anomalous threshold (Klusman, 1993; Prikle et al., 1983).

3.2. Method used to determine "background", "threshold" and "anomaly"

Determination methods for background and anomalies can be divided into graphic and statistical (mathematical).

The graphic methods include the construction of plots and histograms. Their interpretation enables the researcher to distinguish concentrations representing the background, the anomalous threshold and the range (ranges) of anomalous values.

The statistical (analytical) methods apply relevant calculation procedures. Usually, the mean value is calculated from the population of results representing the background values. Then, the relative values of concentrations are calculated in relation to the background value taken as a reference level.

3.2.1. Background determination with the Lepeltier's method

One of graphic methods of background determination is the plotting of measured concentrations against the two-cycle log probability plot. Such method was invented and applied for the first time by Lepeltier (1969). It is based upon an assumption that measured concentrations match the log-normal distribution. Hence, the grouping of samples according to the increasing values and their presentation as cumulative probability plot enables us to identify the range of values deviated from the log-normal distribution. In the plot such deviation is seen as a distinct change in the slope of the curve. According to Lepeltier (1969), this point is the upper threshold of undisturbed range of concentrations. The arithmetic mean for this range is the calculated background value and the upper threshold of the "noise" is calculated as "threshold level = arithmetic mean + 2σ" (Hawkes and Webb, 1962).

Somewhat simplified version of this method is in use, as well (Klusman, 1993, 2002; Matschullat et al., 2000). Precisely, the slope change point in the probability plot is taken as a boundary between the populations of background and anomalous values. Moreover, such an attempt enables the researcher to distinguish the sub-populations within the anomalous dataset. Their limits are defined by point at which the slope changes appear. Interpolated values of concentrations corresponding to these points can be applied as values of contour lines in the maps.

Unfortunately, this method cannot be applied for populations with low number of data due to interpretation problems. For such populations the precise positioning of changes in slopes of particular curves bears a significant error (Matschullat et al., 2000).