РАЗРАБОТКА ИНФОРМАЦИОННО-РЕГРЕССИОННОГО МЕТОДА МОДЕЛЬНОЙ ОЦЕНКИ СТЕПЕНИ ЗАГРЯЗНЕННОСТИ ПОЧВЫ НЕФТЕПРОДУКТАМИ

DEVELOPMENT OF AN INFORMATION-REGRESSION METHOD FOR A MODEL ASSESSMENT OF THE DEGREE OF SOIL CONTAMINATION BY OIL PRODUCTS

Х.Г. Асадов
Hikmat G. Asadov

НИИ аэрокосмической
информатики,
г. Баку,
Азербайджанская республика

Research Institute
of Aerospace Informatics,
Baku, Republic of Azerbaijan

Р.А. Эминов
Ramiz A. Eminov

Азербайджанский
государственный университет
нефти и промышленности,
г. Баку,
Азербайджанская Республика

Azerbaijan State Oil
and Industrial University,
Baku, Republic of Azerbaijan

Н.З. Мурсалов
Nemat Z. Mursalov

Азербайджанский
государственный университет
нефти и промышленности,
г. Баку,
Азербайджанская Республика

Azerbaijan State Oil
and Industrial University,
Baku, Republic of Azerbaijan

Метод дистанционного зондирования в ближнем и среднем инфракрасном диапазонах может быть использован для обнаружения следов загрязнения почвы нефтепродуктами прямыми и косвенными способами. При этом при применении прямых методов используют непосредственно спектральные признаки отражения углеводородов, а в косвенных методах используются спектральные признаки отражения растительности, подверженных влиянию углеводородов. Для оценки степени загрязненности широко используются статистические и информационные модели оценки. Статья посвящена исследованию возможности формирования нового информационно-регрессионного метода модельной оценки степени загрязненности почвы нефтепродуктами. Предлагается метод модельной информационно-регрессионной оценки в качестве комбинации известного метода информационной оценки верности статистической модели в несколько сокращенном виде и метода линейной регрессионной оценки. Показано, что согласно предлагаемому методу минимальная погрешность оценки достигается при равенстве разностей модельных оценок с исключением континуальной составляющей и наблюдаемых значений, также с исключением континуальной составляющей применительно ко всем спектральным областям, выбранным спектральным признакам загрязнения почвы нефтью и нефтепродуктами.

The method of remote sensing in the near and middle infrared bands can be used to detect traces of soil contamination by oil products by direct and indirect methods. In this case, when using direct methods, spectral signs of the reflection of hydrocarbons are used directly, while indirect methods use spectral signs of reflection of vegetation subjected to the influence of hydrocarbons. Statistical and informational assessment models are widely used to assess the degree of contamination. The article is devoted to the study of the possibility of the formation of a new information-regression method for the model assessment of the degree of soil contamination by oil products. A method of model information-regression assessment is proposed as a combination of the well-known method of informational assessment of the fidelity of the statistical model in a somewhat abbreviated form and the method of linear regression assessment. It is shown that according to the proposed method, the minimum estimate error is achieved when the differences between model estimates with the exception of the continual component and the observed values, also with the exception of the continual component with respect to all spectral regions, selected spectral signs of soil contamination by oil and oil products are equal.

Ключевые слова

загрязненность почвы,
модельная оценка,
информативность,
спектр поглощения,
углеводороды

Key words

soil contamination,
model assessment,
information content, absorption
spectrum, hydrocarbons

Введение

Хорошо известно, что метод дистанционного зондирования в ближнем и среднем инфракрасном (ИК) диапазоне может быть использован для обнаружения следов загрязнения почвы нефтепродуктами прямыми [1–5] и косвенными [6, 7] способами. При этом при применении прямых методов используют непосредственно спектральные признаки отражения углеводородов, а в косвенных методах используются спектральные признаки отражения растительности, подверженных влиянию углеводородов. Для оценки степени загрязненности широко используются статистические [8] и информационные [9] модели оценки. В настоящей статье исследуется возможность разработки нового комбинированного информационно-регрессионного метода для модельной оценки степени загрязнения почвы нефтепродуктами.

Теоретические основы предлагаемого метода

В таблице 1 приведены некоторые характеристики спутниковых сенсоров ИК диапазона, пригодных для дистанционного определения степени загрязнения почвы нефтепродуктами [5].

В работе [8] были исследованы статистические модели загрязнения почвы с использованием признака интегрированного участка спектра поглощения в зоне пиков поглощения в спектре на участках ближнего и среднего инфракрасного диапазона. Использовались такие модели, как линейная регрессия, многократная линейная регрессия, метод частичной регрессии с использованием способа наименьших квадратов и др.

Для оценки точности полученных статистических моделей были использованы такие методы, как вычисление коэффициента детерминации, среднеквадратичного отклонения, а также информационный критерий, предложенный в работе [9], вычисляемый по формуле

$$AIC = N \cdot \ln \left[\sum_{i=1}^N (y_i - y_0)^2 \right] + 2np, \quad (1)$$

где N — количество наблюдений; y_i — результат наблюдения; y_0 — предсказуемая величина; np — количество параметров, используемых для создания модели. Модель с меньшей величиной AIC рассматривается в качестве предпочтительной модели.

Как отмечается в работе [8], интенсивность пиков поглощения в среднем ИК диапазоне сильна по сравнению с пиками в ближнем ИК диапазоне, возникших за счет обертонов фундаментальных молекулярных вибрационных связей. В качестве примера на рисунке 1 приведены NIR спектры песочной (а) и глиняной (б) почв, загрязненных дизельным топливом. Исследованы пики в областях 2220–2460 нм и 1650–1800 нм.

В качестве другого примера на рисунке 2 приведены спектры поглощения песочной (а) и глиняной (б) почв, загрязненных дизельным топливом в различной степени.

Спектры сняты в среднем ИК диапазоне, а для вычисления концентрации загрязнителя использовалась область 2990–2810 см⁻¹. Согласно [8], для вычисления концентрации загрязнителя почвы использовались спектры с удаленной континуальной составляющей. Общий вид таких спектров показан на рисунке 3.

Таблица 1. Некоторые характеристики спутниковых сенсоров ИК диапазона

Название сенсора	Головная организация	Спектральный диапазон	Кол-во спектральных каналов	Спектральное разрешение
SEBASS	Aerospace Corporation, USA	7,5–13,5 μm 741–1333 cm ⁻¹	128	10 nm
АHI	University of Hawaii, USA	7,5–11,5 μm 870–1333 cm ⁻¹	256	125 nm
TASI-600	ITRES, Canada	8,0–11,5 μm 870–1250 cm ⁻¹	64	125 nm
AISA Owl	SPECIM, Finland	8,0–12,0 μm 833–1250 cm ⁻¹	84	100 nm
MAKO	Aerospace Corporation, USA	7,8–13,4 μm 746–1282 cm ⁻¹	128	40 nm
MAGI	Aerospace Corporation, USA	7,0–12,7 μm 787–1429 cm ⁻¹	32	≈ 190 nm
HyTES	JPL, USA	7,5–12,0 μm 833–1333 cm ⁻¹	256	35 nm
SIELETERS	ONERA, France	8,0–11,5 μm 870–1250 cm ⁻¹	n/a	<230 nm

Рисунок 1. Спектры поглощения песочной (а) и глиняной (б) почвы, загрязненной дизельным топливом в различной степени. Исследуются пики на участках 2220–2400 нм и 656–1800 нм для вычисления концентрации загрязнителя

Рисунок 2. Спектры поглощения в среднем ИК диапазоне песочной (а) и глиняной (б) почв, загрязненных дизельным топливом в различной степени. Для вычисления концентрации загрязнителя исследуется спектральный участок 2990–2810 cm^{-1}

Рисунок 3. Спектры поглощения песочной почвы, загрязненной дизельным топливом, в концентрациях 5000 ppm (a), 1000 ppm (b), а также загрязненной моторным маслом в концентрациях 5000 ppm (c) и 10000 ppm (d)

Спектры сняты в среднем ИК диапазоне, в области 2990–2810 см^{-1} [8]. Приняты следующие обозначения: (a) загрязнение дизельным топливом концентрацией 5000 ppm; (b) загрязнение дизельным топливом концентрацией 10000 ppm; (c) загрязнение моторным маслом концентрацией 5000 ppm; (d) загрязнение моторным маслом концентрацией 10000 ppm.

Согласно [8], с использованием спектров с удаленной континуальной составляющей, концентрация углеводородного загрязнителя может быть вычислена, используя линейную регрессионную модель

$$TRH(ppm) = 2328 - 9739A + 1108B - 812C + 1061D, \quad (2)$$

где A — площадь под пиком в области 1630–1580 см^{-1} ; B — площадь под пиком в области 1930–1840 см^{-1} ; C — площадь под пиком в области 2060–1930 см^{-1} ; D — площадь под пиком в области 2990–2810 см^{-1} .

Предлагаемый метод

Целью проводимого исследования является разработка нового метода проверки достоверности оценок с комбинированием критериев (1) и (2).

Считаем, что известны следующие оценки:

1) Оценки площадей под кривой спектра в заданной спектральной области с вычетом континуальной составляющей для исходного спектра, составленного после статистической обработки измерительных данных $x_{i,об}$; $i = \overline{1, m}$;

2) Оценки площадей под кривой спектра в заданной спектральной области с вычетом континуальной составляющей без статистической обработки измерительных данных $x_{i,m}$; $i = \overline{1, m}$.

Предлагаемый метод проверки достоверности оценок данных $\{x_{i,m}\}$; $i = \overline{1, m}$ предусматривает решение нижеуказанной оптимизационной задачи:

$$F = F_1 + \lambda(F_2 - F_{20}), \quad (3)$$

$$\text{где } F_1 = N \cdot \ln \left[\sum_{i=1}^4 (x_{i,m} - x_{i,об})^2 \right], \quad (4)$$

F_2 — модельная регрессионная оценка концентрации загрязнителя; F_{20} — известная величина концентрации загрязнителя:

$$F_2 = a_1 + a_2 x_{1,m} + a_3 x_{2,m} + a_4 x_{3,m} + a_5 x_{4,m}, \quad (5)$$

где λ — множитель Лагранжа; $a_j = \text{const}$; $j = \overline{1,5}$.

Сравнивая выражения (1) и (4), нетрудно увидеть, что критерий F_1 является сокращенной формой записи критерия AIC . При этом критерий F_1 совпадает по форме с оценкой (2).

Задачей исследования является вычисление таких модельных величин $x_{i,m}$, при которых $F \rightarrow \min$.

Задача исследования решается методом множителей Лагранжа [Л], согласно которому должны быть вычислены следующие производные с дальнейшим приравнением их к нулю:

$$\left. \begin{aligned} \frac{dF}{dx_{1,m}} &= \frac{2N(x_{1,m} - x_{1,об})}{\sum_{i=1}^4 (x_{i,m} - x_{i,об})^2} + \lambda a_2 = 0 \\ \frac{dF}{dx_{2,m}} &= \frac{2N(x_{2,m} - x_{2,об})}{\sum_{i=1}^4 (x_{i,m} - x_{i,об})^2} + \lambda a_3 = 0 \\ &\dots\dots\dots \\ \frac{dF}{dx_{4,m}} &= \frac{2N(x_{4,m} - x_{4,об})}{\sum_{i=1}^4 (x_{i,m} - x_{i,об})^2} + \lambda a_5 = 0 \end{aligned} \right\}. \quad (6)$$

Из выражений (6) нетрудно получить следующее равенство

$$F = F_{1,C} = N \ln \left[\left(\frac{x_{1,m} - x_{1,об}}{a_2} \right)^2 + \left(\frac{x_{2,m} - x_{2,об}}{a_3} \right)^2 \right]. \quad (7)$$

Модельные исследования

Покажем, что при решении (7) функционал (3) достигает минимальной величины. Для упрощения выкладок ограничимся двумя членами сумм в выражениях (4) и (5). Имеем

$$F = F_{1,C} = N \ln \left[\left(\frac{x_{1,m} - x_{1,об}}{a_2} \right)^2 + \left(\frac{x_{2,m} - x_{2,об}}{a_3} \right)^2 \right], \quad (8)$$

$$F_{2,C} = [(a_1 + a_2 x_{1,m} + a_3 x_{2,m}) - F_{20,C}], \quad (9)$$

$$F = F_{1,C} + \lambda \cdot F_{2,C}, \quad (10)$$

$$F = N \ln \left[\left(\frac{x_{1,m} - x_{1,об}}{a_2} \right)^2 + \left(\frac{x_{2,m} - x_{2,об}}{a_3} \right)^2 \right] + \lambda \cdot [(a_1 + a_2 x_{1,m} + a_3 x_{2,m}) - F_{20,C}]. \quad (11)$$

Получим полный дифференциал критерия (8). Имеем:

$$dF_{1,C} = \frac{\frac{2N}{a_2}(x_{1,m} - x_{1,об})dx_{1,m} + \frac{2N}{a_3}(x_{2,m} - x_{2,об})dx_{2,m}}{\left[\left(\frac{x_{1,m} - x_{1,об}}{a_2} \right)^2 + \left(\frac{x_{2,m} - x_{2,об}}{a_3} \right)^2 \right]}, \quad (12)$$

из выражения (9) получим:

$$\frac{dx_{1,m}}{dx_{2,m}} = -\frac{a_3}{a_2} \quad (13)$$

Из выражений (12) и (13) находим

$$\frac{dF_{1,C}}{dx_{1,m}} = \frac{\frac{2N}{a_2}(x_{1,m} - x_{1,об}) - \frac{2N}{a_3}(x_{2,m} - x_{2,об}) \cdot a_2/a_3}{\left[\left(\frac{x_{1,m} - x_{1,об}}{a_2} \right)^2 + \left(\frac{x_{2,m} - x_{2,об}}{a_3} \right)^2 \right]}. \quad (14)$$

С учетом (7) из выражения (14) ясно, что при $a_2 = a_3$ имеем нулевую величину $\frac{dF_{1,C}}{dx_{1,m}}$, т.е. при (7) условие экстремума выполняется.

Вычислим вторую производную $F_{1,C}$ по $x_{1,m}$ при $a_2 = a_3$.

$$\frac{d^2 F_{1,C}}{dx_{1,m}^2} = \frac{2N[M] - 2(x_{1,m} - x_{1,об}) \cdot [2N(x_{1,m} - x_{1,об}) - 2N(x_{2,m} - x_{2,об})]}{[M^2]} = 2N, \quad (15)$$

$$\text{где } M = [(x_{1,m} - x_{1,об})^2 + (x_{2,m} - x_{2,об})^2]. \quad (16)$$

Как видно из выражения (15), $\frac{d^2 F_{1,C}}{dx_{1,m}^2}$ принимает положительную величину. Следовательно, при решении (7) функционал (3) достигает минимальную величину, что указывает на правомерность предлагаемого информационно-регрессионного критерия оценки степени загрязненности почвы нефтепродуктами.

Вывод

Таким образом, предлагаемый метод информационно-регрессионной оценки является комбинацией известного метода информационной оценки верности статистической

модели [9] в несколько сокращенном виде и метода линейной регрессионной оценки иско́мой концентрации загрязнителя. Показано, что согласно предлагаемому методу оценки минимальная погрешность оценки достигается при равенстве разностей модельных оце-

нок с исключением континуальной составляющей и наблюдаемых значений, также с исключением континуальной составляющей применительно ко всем спектральным областям, выбранных спектральных признаков загрязнения почвы нефтью и нефтепродуктами.

СПИСОК ИСПОЛЪЗУЕМЫХ ИСТОЧНИКОВ

1. Lammoglia T., Filho C.R.D.S. Spectroscopic Characterization of Oils Yielded from Brazilian Offshore Basins: Potential Applications of Remote Sensing // *Remote Sensing of Environment*. 2011. No. 115 (10). P. 2525–2535.
2. Bihong F., Zheng G., Ninomiya Y., Wang C., Sun G. Mapping Hydrocarbon Induced Mineralogical Alteration in the Northern Tian Shan Using ASTER Multispectral Data // *Terra Nova*. 2007. No. 19 (4). P. 225–231.
3. van der Werff H., van der Meijde M., Jansma F., van der Meer F., Groothuis G.-J. A Spatial-Spectral Approach for Visualization of Vegetation Stress Resulting from Pipeline Leakage // *Sensors*. 2008. No. 8. P. 3733–3743.
4. Winkelmann K. On the Applicability of Imaging Spectroscopy for the Detection and Investigation of Contaminated Sites with Particular Consideration Given to the Detection of Fuel Hydrocarbon Contamination in Soil: Ph.D. Thesis. Brandenburgischen Technischen Universität Cottbus. 2005.
5. van der Meijde M., Knox N.M., Cundill S.L., Noomen M.F., van der Werff H.M.A., Hecker C. Detection of Hydrocarbons in Clay Soils: A Laboratory Experiment Using Spectroscopy in The Mid- and Thermal Infrared // *International Journal of Applied Earth Observation and Geoinformation*. 2013. No. 23. P. 384–388.
6. van der Meijde M., van der Werff H., Jansma F., van der Meer F., Groothuis G.-J. A Spectral-Geophysical Approach for Detecting Pipeline Leakage // *International Journal of Applied Earth Observation and Geoinformation*. 2009. No. 11. P. 77–82.
7. Noomen M., Skidmore A., van der Meer F., Prins H. Continuum Removed Band Depth Analysis for Detecting the Effects of Natural Gas, Methane, and Ethane on Maize Reflectance // *Remote Sensing of the Environment*. 2006. No. 105. P. 262–270.
8. Ng W., Malone B.P., Minasny B. Rapid Assessment of Petroleum-Contaminated Soils with Infrared Spectroscopy. Centre for Carbon, Food & Water, Faculty of Agriculture & Environment, University of Sydney. NSW 2006, Australia.
9. Akaike H. A New Look at the Statistical Model Identification // *IEEE Trans. Autom. Control*. 1974. No. 19. P. 716–723.

СВЕДЕНИЯ ОБ АВТОРАХ

ABOUT THE AUTHORS

Асадов Хикмет Гамид оглы, д-р техн. наук, начальник отдела, НИИ аэрокосмической информатики, г. Баку, Азербайджанская республика

Hikmat G. Asadov, Doctor of Engineering Sciences, Head of Division, Research Institute of Aerospace Informatics, Baku, Republic of Azerbaijan

e-mail: asadzade@rambler.ru

REFERENCES

1. Lammoglia T., Filho C.R.D.S. Spectroscopic Characterization of Oils Yielded from Brazilian Offshore Basins: Potential Applications of Remote Sensing. *Remote Sensing of Environment*, 2011, No. 115 (10), pp. 2525–2535.
2. Bihong F., Zheng G., Ninomiya Y., Wang C., Sun G. Mapping Hydrocarbon Induced Mineralogical Alteration in the Northern Tian Shan Using ASTER Multispectral Data. *Terra Nova*, 2007, No. 19 (4), pp. 225–231.
3. van der Werff H., van der Meijde M., Jansma F., van der Meer F., Groothuis G.-J. A Spatial-Spectral Approach for Visualization of Vegetation Stress Resulting from Pipeline Leakage. *Sensors*, 2008, No. 8, pp. 3733–3743.
4. Winkelmann K. *On the Applicability of Imaging Spectroscopy for the Detection and Investigation of Contaminated Sites with Particular Consideration Given to the Detection of Fuel Hydrocarbon Contamination in Soil: Ph.D. Thesis*. Brandenburgischen Technischen Universität Cottbus. 2005.
5. van der Meijde M., Knox N.M., Cundill S.L., Noomen M.F., van der Werff H.M.A., Hecker C. Detection of Hydrocarbons in Clay Soils: A Laboratory Experiment Using Spectroscopy in The Mid- and Thermal Infrared. *International Journal of Applied Earth Observation and Geoinformation*, 2013, No. 23, pp. 384–388.
6. van der Meijde M., van der Werff H., Jansma F., van der Meer F., Groothuis G.-J. A Spectral-Geophysical Approach for Detecting Pipeline Leakage. *International Journal of Applied Earth Observation and Geoinformation*, 2009, No. 11, pp. 77–82.
7. Noomen M., Skidmore A., van der Meer F., Prins H. Continuum Removed Band Depth Analysis for Detecting the Effects of Natural Gas, Methane, and Ethane on Maize Reflectance. *Remote Sensing of the Environment*, 2006, No. 105, pp. 262–270.
8. Ng W., Malone B.P., Minasny B. *Rapid Assessment of Petroleum-Contaminated Soils with Infrared Spectroscopy*. Centre for Carbon, Food & Water, Faculty of Agriculture & Environment, University of Sydney. NSW 2006, Australia.
9. Akaike H. A New Look at the Statistical Model Identification. *IEEE Trans. Autom. Control*, 1974, No. 19, pp. 716–723.

Эминов Рамиз Ахмед оглы, канд. техн. наук, доцент кафедры «Поиск и разведка месторождений нефти и газа», Азербайджанский государственный университет нефти и промышленности, г. Баку, Азербайджанская Республика

Ramiz A. Eminov, Candidate of Engineering Sciences, Assistant Professor of Search and Survey of Oil and Gas Deposits Department, Azerbaijan State Oil and Industrial University, Baku, Republic of Azerbaijan

e-mail: Eminovramiz@mail.ru

Мурсалов Немат Закир оглы, инженер кафедры «Поиск и разведка месторождений нефти и газа», докторант Института водных проблем, Азербайджанский государственный университет нефти и промышленности, г. Баку, Азербайджанская Республика

Nemat Z. Mursalov, Engineer of Search and Survey of Oil and Gas Deposits Department, Doctorant of Institute of Water Problems, Azerbaijan State Oil and Industrial University, Baku, Republic of Azerbaijan

e-mail: mursalovnemat@mail.ru