

АППРОКСИМАЦИЯ ТРЕХМЕРНЫХ СВЯЗЕЙ НА ОСНОВЕ ТЕХНОЛОГИЙ МНОГОВАРИАНТНОГО МОДЕЛИРОВАНИЯ

Кобрунов А.И., Дорогобед А.Н., Кожевникова П.В.

ФГБОУ ВПО «Ухтинский государственный технический университет»,

Ухта, e-mail: aira_dark@list.ru

В данной статье рассмотрена аппроксимация трехмерных связей, с использованием технологий многовариантного моделирования на основе теории нечетких множеств. Показан способ представления данных в форме нечетких отношений, а измеренных значений параметра – в форме нечетких величин. Описан выбор функции принадлежности при прогнозировании параметров, а также выбор элемента аппроксимации. В данном случае функция принадлежности рассматривается как диффузионное рассеяние в бесконечном однородном пространстве параметров точечных источников. Проведен подробный анализ расчета количества точечных источников. Приведен эксперимент расчета аппроксимации функции принадлежности на примере трехпараметрических связей. Поэтапно приведены результаты расчета с их описанием. В конечном итоге приведен результат представления трехмерных связей в виде нечеткого множества и для лучшего понимания представлены сечения данного множества, а именно сечения прямой плоскости и наклонной.

Ключевые слова: трехмерные связи, нечеткое моделирование, нечеткие множества, поле рассеяния, функция принадлежности, прогноз параметров

APPROXIMATION OF THREE-DIMENSIONAL RELATION BASED ON TECHNOLOGY OF MULTIPLE MODELING

Kobrunov A.I., Dorogobed A.N., Kozhevnikova P.V.

Ukhta state technical university, Ukhta, e-mail: aira_dark@list.ru

In this article the approximation of the three-dimensional relationships using multiple modeling techniques based on the theory of fuzzy sets. Illustrates a method of data presentation in the form of fuzzy relations and the measured values in the form of fuzzy variables. Described selection of membership function parameters in the prediction and selection of the element of approximation. In this case the membership function is considered as a diffuse scattering in an infinite homogeneous parameter space point source. The detailed analysis of the calculation of the number of point sources. An experiment approximation calculation of the membership function on the example of three-parameter relations. Incrementally shows the results of the calculation with their description. Ultimately, the result is a representation of three-dimensional relationships as a fuzzy set, and for a better understanding of this section presents a plurality of, namely, a straight section and an inclined plane.

Keywords: three-dimensional relation, fuzzy modeling, fuzzy sets, field scattering, membership function, forecast parameters

В 3-мерном пространстве S подмножество $\mathcal{A} \subset S$ образуют экспериментально измеренные значения $s^j \in S, j = 1 \dots M$ параметров $\mathbf{s} = \{s_1, s_2, s_3\}$. Экспертное правило прогноза одних параметров $\boldsymbol{\eta}$ из числа $s_p, i = 1 \dots N$, по измеренным другим, обозначаемым $\boldsymbol{\xi}$: $\mathbf{s} = \boldsymbol{\xi} \times \boldsymbol{\eta}$ определяется данными $\mathcal{A} \subset S$, являющимися экспериментальным основанием. В задачах промышленной геофизики петрофизические зависимости традиционно являются таким экспертным правилом [1], представляя данные \mathcal{A} в виде регрессионных зависимостей одних параметров от других. Данный способ обширно используется в настоящее время, но он не принимает во внимание объективные свойства неоднородностей в изучаемых параметрах.

Правило $\mathcal{A}[\bar{\boldsymbol{\xi}}]$ вывода результата о прогнозируемых параметрах $\bar{\boldsymbol{\eta}}$ посредством использования информации об измеренных параметрах $\bar{\boldsymbol{\xi}}$, связанное с обучающи-

ми данными $\mathcal{A} \subset S$, реализуется посредством: определения нечеткого отношения $\mu_{\mathcal{A}}(\boldsymbol{\xi}, \boldsymbol{\eta}) = \mu_{\mathcal{A}}(\mathbf{s})$ по обучающей выборке \mathcal{A} [5] между параметрами $\boldsymbol{\eta}$ и $\boldsymbol{\xi}$; представления конкретных измерений $\bar{\boldsymbol{\xi}}$ в виде нечеткой величины и последующего использования алгоритма нечеткого логического вывода [3] о нечеткой величине $\boldsymbol{\eta}$.

Поле рассеяния является основанием для отображения данных \mathcal{A} в виде нечетких отношений, а измеренных значений параметра $\bar{\boldsymbol{\xi}}$, с помощью которых выполняется прогноз, в виде нечетких величин.

Поле рассеяния для данных $s^j \in S, j = 1 \dots M$ является такая функция $\mathcal{A}^\varepsilon(\mathbf{s})$ в пространстве, что для каждой подобласти ΔS пространства S :

$$\max_{\Delta S \in S} |\mathcal{A}^\varepsilon(\mathbf{s}) \Delta S - \mathcal{A}(\Delta S)| \leq \varepsilon, \quad (1)$$

где $\mathcal{A}(\Delta S)$ – число значений из $\mathcal{A} \subset S$, целиком лежащее в ΔS .

Функцией принадлежности $\mu_{\mathfrak{A}}(\mathbf{s})$ для измеренных значений параметров $\mathbf{s} \in S$ в виде нечетких величин является нормированное к единице поле рассеяния $\mathfrak{A}^e(\mathbf{s})$.

Для того чтобы построить поле рассеяния $\mathfrak{A}^e(\mathbf{s})$, необходимо пространство S разбить сеткой ΔS^k , $k = 1 \dots K$, покрывающей всю область значений параметров из \mathfrak{A} , $\sum_{k=1}^K |\mathfrak{A}(\Delta S^k)| = N$, и поставить задачу нахождения функции рассеяния по (1):

$$\max_{\Delta S^k \in S} |\mathfrak{A}^e(\mathbf{s}) \Delta S^k - \mathfrak{A}(\Delta S^k)| \leq \varepsilon. \quad (2)$$

Необходимо выбрать базисную систему функций $K(\mathbf{h}, \mathbf{s})$, параметризованную вектором параметров \mathbf{h} , и поставить задачу нахождения $\mathfrak{A}^e(\mathbf{s})$ в виде

$$\mathfrak{A}^e(\mathbf{s}) = \sum_{k=1}^K \varphi(\mathbf{h}^k) K(\mathbf{h}^k, \mathbf{s}). \quad (3)$$

Функция $\varphi(\mathbf{h}^k)$ требует нахождения из (2). Выражение (3) показывает, что поле рассеяния рассчитывается в виде линейной комбинации базовых функций $K(\mathbf{h}, \mathbf{s})$, определяющих метод для аппроксимации поля рассеяния $\mathfrak{A}^e(\mathbf{s})$, а также функции принадлежности $\mu_{\mathfrak{A}}(\mathbf{s})$.

Для аппроксимации функции принадлежности принято использовать треугольные трапециевидные и другие ранее вводимые зависимости [2]. В противоположность данным принципам для элемента аппроксимации $K(\mathbf{h}, \mathbf{s})$ введем метод максимальной энтропии. Согласно данному методу в качестве элемента аппроксимации была использована функция нормального закона распределения [4], параметром \mathbf{h} в котором является математическое ожидание:

$$K(\mathbf{h}, \mathbf{s}) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{|\mathbf{h}-\mathbf{s}|^2}{2\sigma^2}\right], \quad (4)$$

где σ^2 – дисперсия нормального распределения.

Из выражения (3) получаем

$$\mathfrak{A}^e(\mathbf{s}) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{k=1}^K \varphi(\mathbf{h}^k) \exp\left[-\frac{|\mathbf{h}^k - \mathbf{s}|^2}{2\sigma^2}\right]. \quad (5)$$

Выражение (5) поясняется как диффузионное рассеяние в бесконечном однородном пространстве параметров точечных источников, расположенных в \mathbf{h}^k . По описанному принципу расчета аппроксимации функции принадлежности произведем вычислительный эксперимент.

В результате измерений получено следующее отношение между характеристиками, где X – абсолютная глубина; Y – двойное время ВСП (вертикальное сейсмическое профилирование), а Z – средняя скорость.

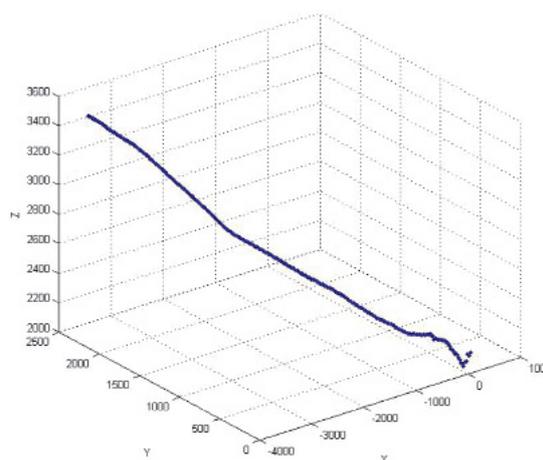


Рис. 1. Зависимость между абсолютной глубиной, двойным временем ВСП и средней скоростью

Для более наглядного отображения представлены проекции на плоскости YZ , XZ и XY (табл. 1).

Таблица 1

Проекция на плоскости

Проекция на плоскость YZ	Проекция на плоскость XZ	Проекция на плоскость XY

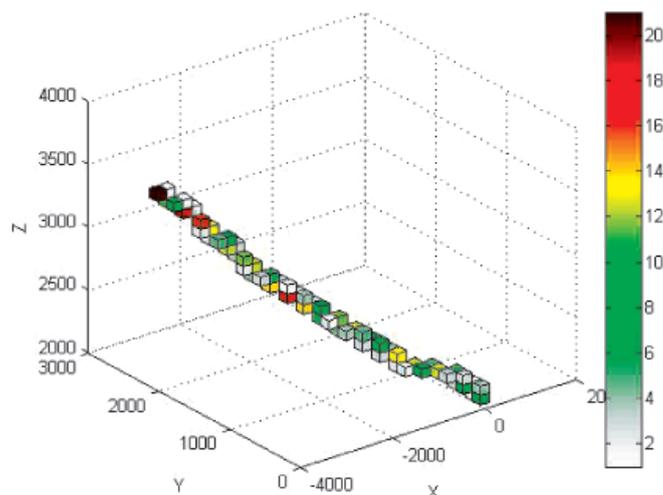


Рис. 2. Карта плотности данных
(справа легенда по параметру количества данных, попавших в ячейку)

На основе данного отношения была получена карта плотности данных (см. рис. 2). Картой плотности называется отношение числа данных для каждого выделенного интервала к общему числу данных.

На основе приведенной карты плотности данных необходимо построить карту источников информации. Очевидно, что чем больше источников на карте, тем меньше погрешность при аппроксимации функции принадлежности, отсюда следует задача поиска оптимального количества источников.

Метод состоит в следующем: задается уровень погрешности и по невязке определяется необходимое количество источников. Далее проведем расчеты для поставленной задачи.

На рис. 3 представлен график исходных данных. Погрешность считалась по формуле (1).

На основе данного графика мы можем посмотреть отношение приращения количества источников к приращению погрешности по формуле (6), что покажет скорость изменения количества источников относительно предыдущего шага (табл. 2).

$$\frac{\Delta n}{\Delta \epsilon} = \frac{N_i - N_{i+1}}{\epsilon_{i+1} - \epsilon_i}, \quad (6)$$

где N_i – это значение количества источников при текущем значении погрешности ϵ_i .

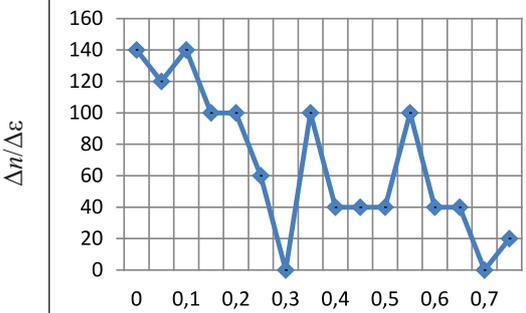
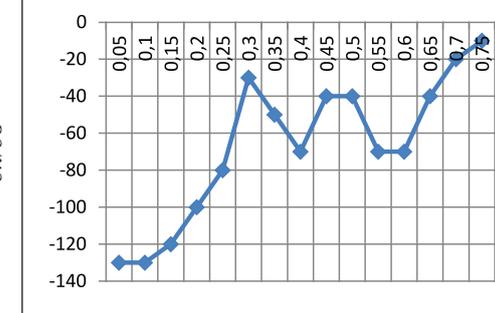
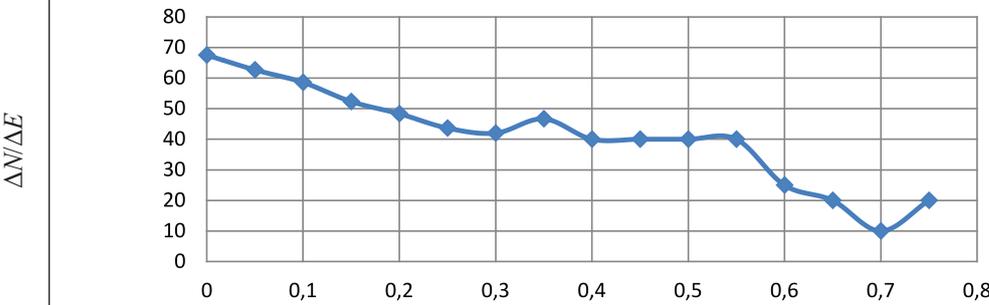
Из чего мы можем сделать вывод: для данного эксперимента оптимальной погрешностью являются значения 0,3 и 0,7, при этом значения количества источников соответственно равны 26 и 6, так как отношение приращения количества источников к приращению погрешности на данном отрезке дает наименьший результат.



Рис. 3. Зависимость погрешности от количества источников

Таблица 2

Поиск оптимального количества источников

1	<p>Скорость изменения</p> 	2	<p>Скорость нарастания количества источников с изменением погрешности</p> 
Погрешность, ε		Погрешность, ε	
3	<p>Изменение количества источников относительно всего объема данных</p> 		
Погрешность, ε			

Далее была рассчитана зависимость скорости нарастания количества источников от увеличения погрешности, которая рассчитывается как производная в каждой точке (табл. 2). Так как производная в точке равна углу наклона касательной, проведенной через данную точку, в качестве касательной была рассчитана прямая проходящая через прилежащие точки (7):

$$\frac{x - N_{i-1}}{N_{i+1} - N_{i-1}} = \frac{y - \epsilon_{i-1}}{\epsilon_{i+1} - \epsilon_{i-1}}. \quad (7)$$

Как видно на графике, скорость нарастания количества источников ближе к нулю при значениях погрешности 0,3 и 0,7, при этом количество источников равно 26 и 6.

Также рассчитываем изменение количества источников относительно всего объема данных по формуле (8), что позволит найти оптимальное значение погрешности, учитывая весь объем выборки (табл. 2).

$$\frac{\Delta N}{\Delta E} = \frac{N_1 - N_i}{\epsilon_i - \epsilon_1}. \quad (8)$$

Минимумы изменения количества источников относительно всего объема данных (59 возможных источников) при значениях 0,3; 0,4 и 0,7, при этом количество источников равны 26, 21 и 6.

Вывод: оптимальное количество источников для данного эксперимента – не менее 26. Погрешность при данном количестве источников равна 0,3.

После нахождения оптимального количества источников была построена карта источников информации (рис. 4), построенное поле рассеяния по которой, наиболее соответствует наблюдаемым данным.

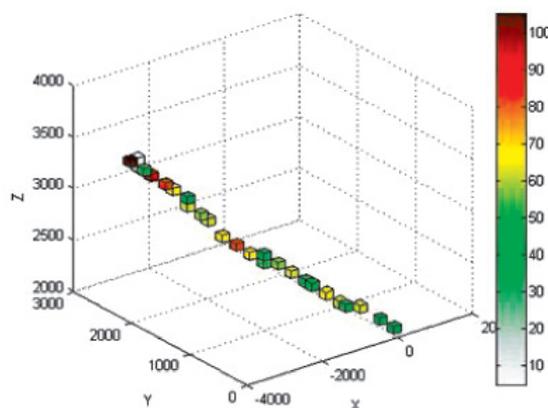


Рис. 4. Карта источников (справа легенда по параметру значений в источниках)

Таблица 3

Поле рассеяния, дифференцированное по уровню доверия

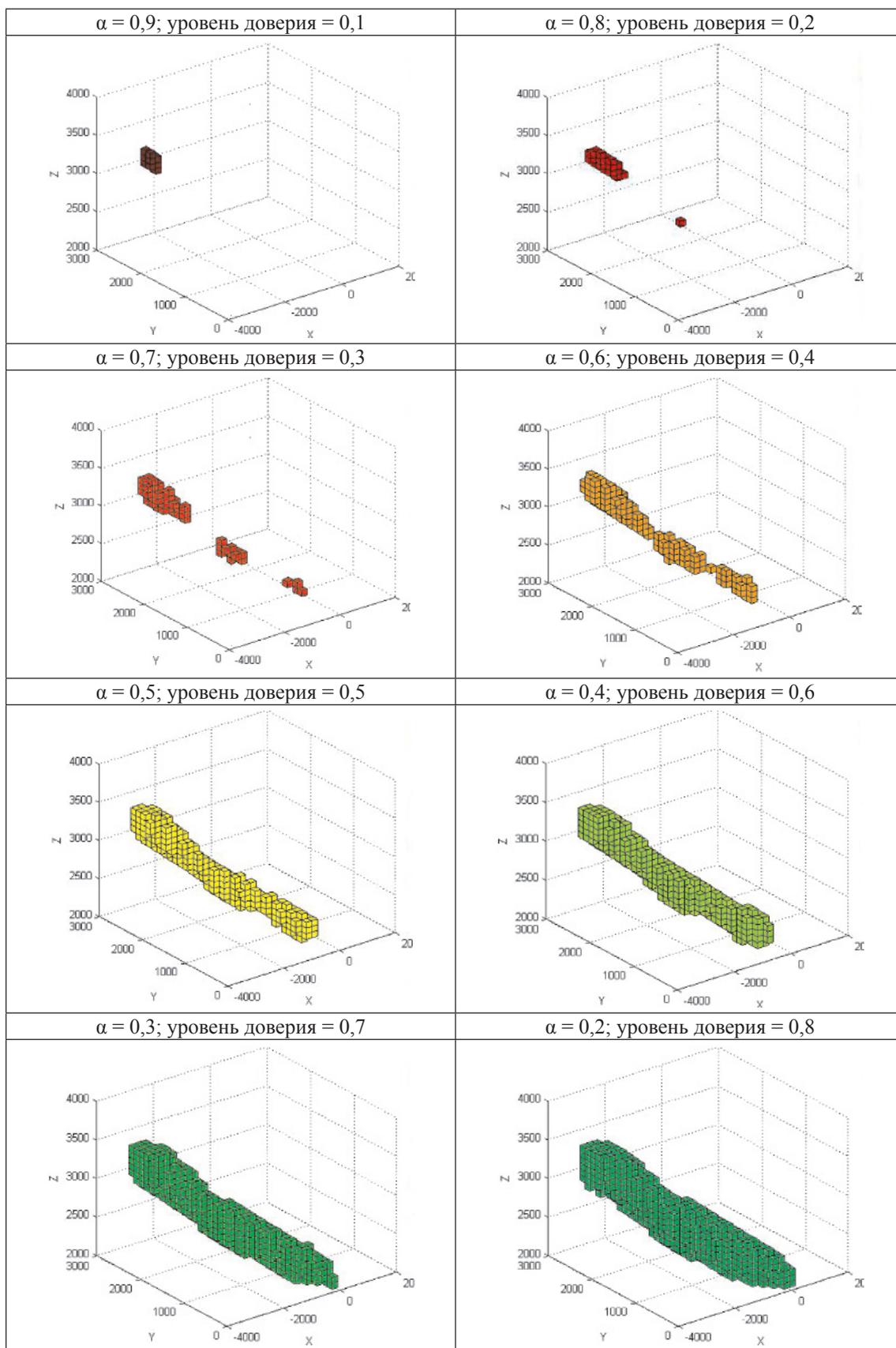
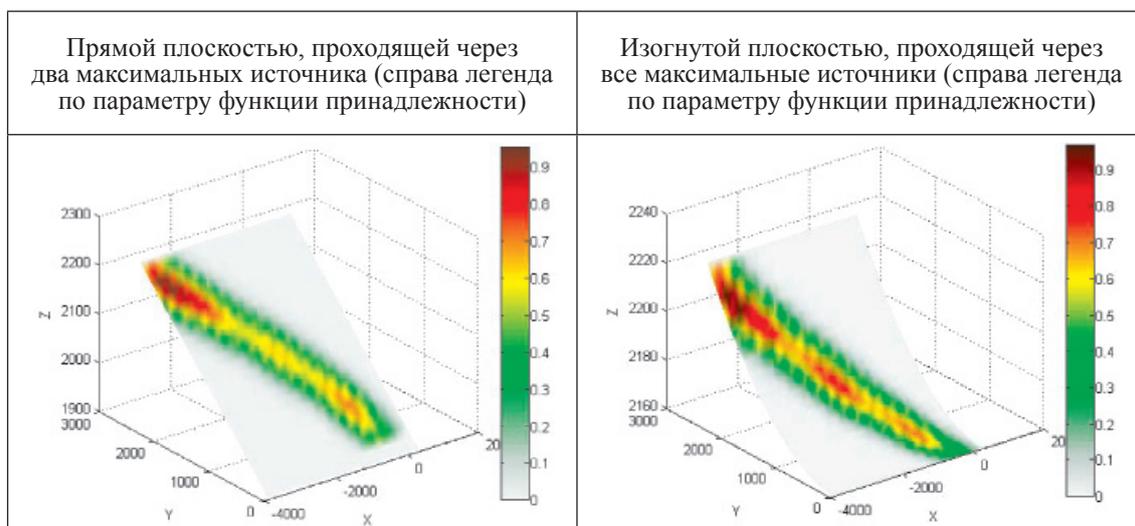


Таблица 4

Сечение поля рассеяния плоскостями



На основании карты были рассчитаны поля рассеяния. Ниже приведен пример построенного поля рассеяния, дифференцированный по уровню доверия (табл. 3).

Чтобы лучше понять данное поле рассеяния, представлены сечения (табл. 4).

Список литературы

1. Вендельштейн Б.Ю., Резванов Р.А. Геофизические методы определения параметров нефтегазовых коллекторов: при подсчете запасов и проектировании разработки месторождений. – М.: Недр, 1978. – 318 с.
2. Жирабок А.Н. Нечеткие множества и их использование для принятия решений // Соросовский образовательный журнал. – 2001. – Т. 7. – № 2. – С. 109–115.
3. Mamdani E.H. Application of fuzzy algorithms for control of simple dynamic plant. *Electrical Engineers // Proceedings of the IEE.* – 1974. – № 121(12). – P. 1585–1588.
4. Park S.Y., Bera A.K. Maximum entropy autoregressive conditional heteroskedasticity model // *Journal of Econometrics.* – 2009. – Т. 150. – № 2. – С. 219–230.
5. Zadeh L.A. Fuzzy sets. – *Information and Control.* – 1965. – Vol. 8, № 3. – P. 338–353.

References

1. Vendelshtejn B.Ju., Rezvanov R.A. *Geofizicheskie metody opredelenija parametrov neftegazovyh kollektorov: pri podschete zapasov i proektirovanii razrabotki mestorozhdenij.* M.: Nedr, 1978. 318 p.
2. Zhirabok A.N. Nечеткие множества и их использование для принятия решений // *Sorosovskij obrazovatelnyj zhurnal.* 2001. T. 7. no. 2. pp. 109–115.
3. Mamdani E.H. Application of fuzzy algorithms for control of simple dynamic plant. *Electrical Engineers // Proceedings of the IEE.* 1974. no. 121(12). pp. 1585–1588.
4. Park S.Y., Bera A.K. Maximum entropy autoregressive conditional heteroskedasticity model // *Journal of Econometrics.* 2009. T. 150. no. 2. pp. 219–230.
5. Zadeh L.A. Fuzzy sets. *Information and Control.* 1965. Vol. 8, no. 3. R. 338–353.

Рецензенты:

Бурмистрова О.Н., д.т.н., профессор, заведующая кафедрой ТМЛ, Ухтинский государственный технический университет, г. Ухта;

Некучаев В.О., д.ф.-м.н., профессор, заведующий кафедрой физики, Ухтинский государственный технический университет, г. Ухта.