

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ТЕХНИКО-ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ РАБОТЫ УГОЛЬНЫХ ШАХТ МЕТОДОМ НЕЙРОННЫХ СЕТЕЙ

В. П. ПОТАПОВ, С. Н. ЛАЗАРЕНКО, М. В. АЛФЕРОВ
Институт угля и углехимии СО РАН, Кемерово, Россия
e-mail: m.alferov@kuzbass.net

This paper is devoted to system engineering of forecasting of significances of technical and economic parameters of work of the coal enterprises on a short-term perspective, based on the artificial neural network technology. Authors studied the forecast of technical and economic parameters of work of the coal enterprises earlier, and the obtained conclusions lie in the basis of the present work.

1. Постановка задачи нейросетевого прогнозирования

Имеются статистические данные технико-экономических показателей работы угольной шахты за определенные периоды ее деятельности, представленные в виде матрицы Q размерностью $m \times n$:

$$Q = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ q_{m1} & q_{m2} & \cdots & q_{mn} \end{pmatrix}, \quad (1)$$

где m — количество показателей; n — количество периодов статистической отчетности. Столбцы матрицы (1) представляют собой векторы показателей одного периода статистической отчетности предприятия. Элементы строк матрицы (1) образуют временные ряды исследуемых показателей. Для прогнозирования выбираются один из показателей и строка матрицы (1), соответствующая выбранному показателю. Используя аппарат нейронных сетей, требуется найти значения прогнозируемого показателя в перспективе из h периодов, которые образуют вектор

$$\bar{p} = (p_1, p_2, \dots, p_h). \quad (2)$$

2. Реализация нейросетевого прогнозирования

Выберем для прогнозирования один показатель и соответствующую ему строку матрицы Q , индекс которой обозначим v . Обозначим \bar{x}_i i -й столбец матрицы Q , а \bar{y}_i — часть строки

с индексом v матрицы Q , состоящей из h идущих последовательно элементов:

$$\bar{x}_i = \begin{pmatrix} q_{1i} \\ q_{2i} \\ \vdots \\ q_{mi} \end{pmatrix}, \quad i = \overline{1, n}, \quad (3)$$

$$\bar{y}_i = (q_{vi+1}, q_{vi+2}, \dots, q_{vi+h}), \quad i = \overline{1, n-h}. \quad (4)$$

Предположим, что h идущих подряд значений прогнозируемого параметра зависят от вектор-столбца показателей, идущего перед первым прогнозируемым значением:

$$\bar{y}_i = F \{ \bar{x}_i \}, \quad i = \overline{1, n-h}. \quad (5)$$

Тогда в качестве неизвестного оператора F будем использовать нейронную сеть, размерность входного и выходного пространства которой соответственно равна m и h , а пары векторов (\bar{x}_i, \bar{y}_i) , $i = \overline{1, n-h}$, будут составлять обучающее множество A для нейронной сети:

$$A = \bigcup_{i=1}^{n-h} (\bar{x}_i, \bar{y}_i). \quad (6)$$

После обучения и тестирования нейронной сети прогнозирование исследуемого параметра сводится к подстановке в обученную сеть векторов (3) при $i = \overline{n-h+1, n}$. Вычисленные нейронной сетью строки формируют матрицу R , i -я строка которой образована подстановкой в нейронную сеть вектора из множества (3) с индексом $j = n-h+i$. Для элемента p_1 вектора \bar{p} получаются h прогнозируемых значений, от каждого вектора (3) при $i = \overline{n-h+1, n}$. Для элемента p_2 получаются $h-1$ прогнозируемых значений и т.д., для элемента p_h получается одно прогнозируемое значение:

$$\begin{aligned} p_1 &: r_{ih-i+1}, & i = \overline{1, h}, \\ p_2 &: r_{i+1h-i+1}, & i = \overline{1, h-1}, \\ &\dots & \\ p_h &: r_{i+h-1h-i+1}, & i = 1. \end{aligned} \quad (7)$$

Логично предположить, что элементы диагоналей (7), сформированные векторами (3) с меньшим индексом, имеют большую вероятность прогноза, чем элементы диагоналей, сформированные векторами с большим индексом. Тогда для вычисления значений вектора \bar{p} целесообразно ввести коэффициенты взвешивания $k_i, i = \overline{1, h}$, и вычислить компоненты вектора \bar{p} как средневзвешенную сумму элементов диагоналей (7). Коэффициенты $k_i, i = \overline{1, h}$, придают больший вес элементам диагоналей (7), сформированным векторами (3) с меньшим индексом, и удовлетворяют следующим условиям:

$$\begin{aligned} \sum_{i=1}^h k_i &= 1, \\ k_i &\in [0, 1], \quad i = \overline{1, h}, \\ k_1 &> k_2 > \dots > k_h. \end{aligned} \quad (8)$$

Умножая элементы диагоналей (7) на соответствующие коэффициенты и складывая полученные произведения, получаем суммы, которые после умножения на нормировочные

коэффициенты дают компоненты вектора \bar{p} :

$$\begin{aligned} p_1 &= (k_1 r_{1h} + k_2 r_{2h-1} + \dots + k_h r_{h1}) \frac{1}{k_1 + k_2 + \dots + k_h}, \\ p_2 &= (k_2 r_{2h} + k_3 r_{3h-1} + \dots + k_h r_{h2}) \frac{1}{k_2 + k_3 + \dots + k_h}, \\ &\dots \\ p_h &= k_h r_{hh} \frac{1}{k_h}. \end{aligned} \quad (9)$$

В общем виде формула для расчета компонент вектора \bar{p} записывается в виде

$$p_j = \frac{1}{\sum_{i=j}^h k_i} \sum_{i=1}^{h-j+1} k_{j+i-1} r_{j+i-1, h-i+1}, j = \overline{1, h}. \quad (10)$$

3. Вычислительный эксперимент

Апробация алгоритмов системы прогнозирования выполнялась с использованием статистических данных технико-экономических показателей работы 34 угольных шахт Кузбасса. Данные взяты из ежегодных статистических отчетов о деятельности этих предприятий за 1982, 1984, 1985, 1988, 1989, 1995, 1999 и 2001 гг., при этом значения за пропущенные годы получены при помощи кусочно-линейной интерполяции. Для исследования использованы следующие показатели:

- годовая добыча угля, млн т;
- себестоимость 1 т угля, руб.;
- отношение производственных затрат к трудовым затратам, млн руб.;
- средняя мощность пласта, м;
- марка угля, бал.;
- число очистных забоев, шт.;
- средняя нагрузка на забой, т/сут.;
- глубина разработки, м;
- категория по газу, бал.;
- зольность угля, %.

Каждый показатель нормировался в интервале $[-1, 1]$. Статистические показатели, измеряемые в рублях, приведены к ценам 2002 г. при помощи расчетных коэффициентов приведения. В табл. 1 перечислены коэффициенты приведения стоимостных показателей по годам.

В качестве главных объектов прогнозирования выбраны два показателя — годовая добыча угля и себестоимость добычи 1 т угля. Период выполнения прогноза составлял с 2002 по 2008 гг. Коэффициенты взвешивания k_i , $i = \overline{1, h}$, определялись по формуле

$$k_i = \frac{i}{\sum_{j=1}^h j} = \frac{2i}{h(h+1)}, i = \overline{1, h}. \quad (11)$$

Для проверки адекватности моделирования нейронной сети алгоритм прогнозирования на этапе ее обучения предусматривал наличие тестового подмножества обучающего

Таблица 1
Коэффициенты приведения к ценам 2002 г.

Год	Коэффициент приведения
1982	27.25
1984	27.25
1985	27.25
1988	27.25
1989	27.25
1995	0.0055
1999	1.85
2001	1.12

Таблица 2
Результаты прогнозирования показателей годовой добычи угля и себестоимости 1 т угля

Название шахты	2003	2004	2005	2006	2007	2008
Годовая добыча угля, млн т						
Сибирская	0.82	0.90	0.96	0.99	0.97	0.90
Первомайская	1.44	1.17	0.96	1.09	1.66	1.68
им. Кирова	2.22	2.11	2.60	3.27	3.66	3.86
им. 7 Ноября	2.13	2.35	2.43	2.37	2.35	2.51
Себестоимость 1 т угля, руб.						
Сибирская	489	524	537	568	582	612
Первомайская	340	287	306	356	341	375
им. Кирова	301	323	333	328	327	342
им. 7 Ноября	339	272	263	277	279	264

множества A , количество элементов которого составляло $\approx 25\%$ от количества элементов множества A . Элементы тестового подмножества не участвовали в процессе обучения нейросети. Процесс обучения нейросети выполнялся итерационным методом, при этом на каждой итерации рассчитывалась погрешность прогнозирования на всем множестве A , определяемая как средняя норма разности между векторами прогнозируемых и фактических значений исследуемого показателя. Если полученная на определенной итерации погрешность превосходила порог максимального значения погрешности, то выполнялась следующая итерация обучения нейросети с другими заново сгенерированными начальными параметрами обучения нейросети. Экспериментально установлено, что эффективная погрешность прогнозирования, рассчитанная в среднем по всем исследовавшимся угольным предприятиям, не превосходит значения 0.04. Зависимость погрешности прогнозирования от количества итераций обучения нейросети оказалась такой, при которой уменьшение погрешности прогнозирования ниже уровня 0.04 влечет за собой значительное увеличение количества итераций и соответственно машинного времени. Именно поэтому в алгоритме прогнозирования значение порога максимальной погрешности выбрано 0.04. На рис. 1 показан график зависимости погрешности прогнозирования от количества итераций алгоритма прогнозирования.

Количество элементов входного слоя нейронной сети определяется количеством исследуемых показателей и равно десяти, количество элементов в выходном слое нейронной сети равно семи и соответствует периоду прогнозирования. Расчеты по определению вну-

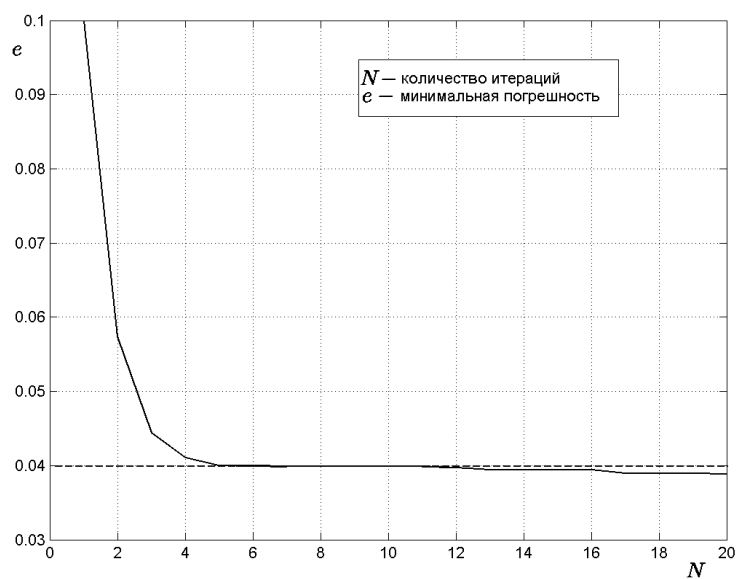


Рис. 1. График зависимости погрешности прогнозирования от количества итераций.

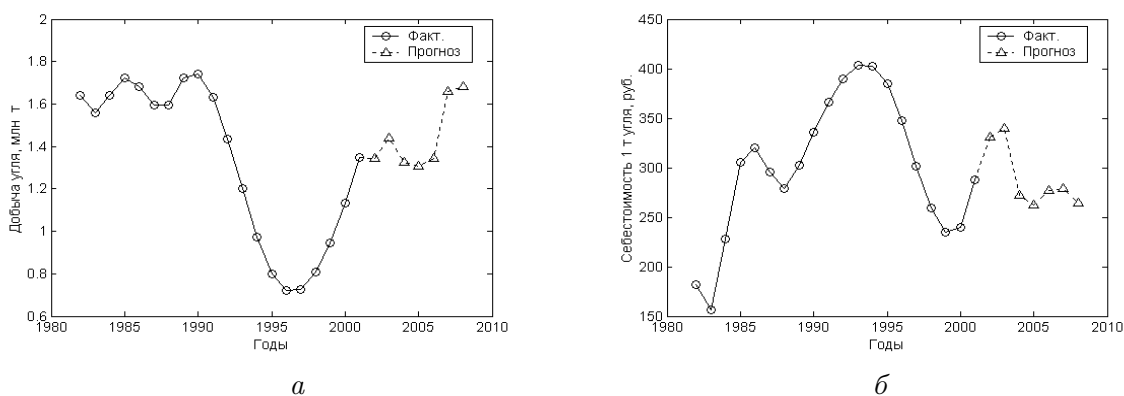


Рис. 2. Результаты прогнозирования показателей годовой добычи угля шахты Первомайская (а) и себестоимости 1 т угля шахты им. 7 Ноября (б).

тренних структур нейросетей проводились в соответствии с методиками, изложенными в [1, 2]. Программная реализация алгоритмов прогнозирования осуществлялась в среде MATLAB [3]. Результаты прогнозирования технико-экономических показателей шахт приведены в табл. 2.

Для количественной оценки результатов прогнозирования показателей годовой добычи угля и себестоимости 1 т угля и проведения сравнительного анализа с другими методами выполнен прогноз при $h = 3$ фактических данных на период 1999 — 2001 гг. Усредненные значения по всем исследованным угольным шахтам приведены в табл. 3. По результатам сравнения методов можно сделать вывод о том, что предложенный метод может использоваться для решения подобного рода задач.

При сравнении результатов прогноза технико-экономических показателей работы угольных шахт на 2008 г. со статистическими данными за 2001 г. можно установить, что максимальное значение темпов прироста показателя годовой добычи угля было получено для

Таблица 3

Результаты сравнения методов на тестовом периоде

Название метода	Среднеквадратическая погрешность	Средняя относительная погрешность, %
Годовая добыча угля		
Метод экспоненциального сглаживания	0.39 млн т	18.0
Кусочно-кубическая эрмитовая интерполяция	0.36 млн т	20.2
Нейронная сеть	0.19 млн т	11.1
Себестоимость 1 т угля		
Метод экспоненциального сглаживания	58.3 руб.	22.0
Кусочно-кубическая эрмитовая интерполяция	62.3 руб.	18.1
Нейронная сеть	47.9 руб.	18.0

шахты Первомайская, составившее 35,8 %, а минимальное значение темпов прироста показателя себестоимости 1 т угля — для шахты им. 7 Ноября, которое составило 8,2 %. На рис. 2 приведены результаты прогнозирования изучавшихся показателей для двух угольных шахт.

Заключение

Разработана оригинальная система прогнозирования технико-экономических показателей работы угольных предприятий на краткосрочную перспективу, основанная на технологии искусственных нейронных сетей. Выполнена апробация алгоритмов прогнозирования по четырем шахтам Кузбасса на период 2002 — 2008 гг.

Полученные результаты прогноза годовой добычи угля и себестоимости 1 т угля не противоречат объективно существующим тенденциям в развитии предприятий угольной промышленности, что в свою очередь позволяет считать правомерным использование метода нейронных сетей для решения задач подобного рода.

Список литературы

- [1] Головкин В. А. Нейронные сети: обучение, организация и применение. Кн. 4. М.: ИПРЖР, 2001. 256 с.
- [2] Круглов В. В., Борисов В. В. Искусственные нейронные сети. Теория и практика. М.: «Торжачая линия-Телеком», 2001. 382 с.
- [3] Потемкин В. Г. Введение в MATLAB. М.: Диалог-Мифи, 1999. 247 с.

*Поступила в редакцию 24 апреля 2003 г.,
в переработанном виде — 18 июня 2003 г.*