

МНОГОКРИТЕРИАЛЬНАЯ ОПТИМИЗАЦИЯ ПОРОШКОВЫХ СТАЛЕЙ

ЭС является итогом проведенного анализа знаний и служит одной из возможностей его непосредственного практического использования. Результаты работы подтвердили, что использование ЭС позволяет сократить время принятия решений при производстве ИПМ, предоставляя возможность быстро и точно подобрать состав исходных смесей, чтобы получить требуемый результат.

Получение порошковых материалов представляет многостадийный процесс, связанный с исходными свойствами порошков, условиями приготовления шихты, формования, режимов спекания и последующей термической обработкой. Экспериментальные данные о свойствах изделий порошковой металлургии (ИПМ) были получены на основании испытаний, проведенных в лаборатории кафедры материаловедения и технологии материалов Оренбургского государственного университета. В качестве исходных материалов для порошковых смесей использовались:

- порошок железный марки ПЖ4М2;
- порошок медный ПМ1;
- порошок чугунный;
- порошок графита;
- порошок стальной Х30.

База данных представлена в виде таблицы, отражающей зависимость между свойствами ИПМ: твердость, предел прочности

на растяжение – и конструкторско-технологическими параметрами производства: материала порошковой смеси, содержания углерода, меди, никеля, хрома, режимами закалки, отпуска и скоростью охлаждения. Влияющие переменные варьировались в следующих пределах:

- 1) содержание углерода $C = 1,0..2\%$;
- 2) содержание меди $Cu = 0..3\%$;
- 3) содержание хрома $Cr = 0..3,0\%$;
- 4) содержание никеля $Ni = 0..3,0\%$;
- 5) пористость $Por = 10..20\%$;
- 6) температура нагрева под закалку $T_{нз} = 750..900^\circ C$;
- 7) скорость охлаждения $Vo_{хл} = 200..800^\circ C$;
- 8) температура отпуска $T_{отп} = 150..500^\circ C$.

Среди зависимых переменных исследовались:

- 1) твердость ИПМ – НВ;
- 2) предел прочности на растяжение – б.

Фрагмент экспериментальной таблицы представлен в таблице 1.

Таблица 1. Экспериментальные данные (фрагмент)

N1	МАТЕРИАЛ	C,%	CU,%	NI,%	Пористость	T _{нз}	Vo _{хл}	T _{отп}	НВ	б
1	ЖГр1.0Д2.5	1,0	2,5	0,0	12,0	750	600	450	123	65,87
2	ЖГр1.0Д2.5	1,0	2,5	0,0	12,0	850	200	150	236	57,60
3	ЖГр1.0Д2.5	1,5	2,5	0,0	12,0	850	600	150	467	29,63
4	ЖГр1.0Д2.5	1,0	2,5	0,0	22,0	800	800	300	185	24,13
5	ЖГр1.25Д2.5	1,2	2,5	0,0	18,0	800	800	300	226	36,40
6	ЖГр1.0Д2.5	1,0	2,5	0,0	11,0	800	800	300	191	75,33
7	ЖГр1.0Д2.5	1,5	2,5	0,0	11,0	740	800	300	269	57,45
8	ЖГр1.0Д2.5	1,5	1,5	0,0	11,0	750	200	450	185	59,60
9	ЖГр1.0Д2.5	1,0	1,5	0,0	11,0	750	600	150	388	49,50
10	ЖГр1Д2.5Н3	1,0	2,5	3,0	20,0	850	600	450	226	28,10
...										
528	ЖГр1.5Д2.5Н3	1,5	2,5	3,0	25,0	800	600	300	297	13,10
529	ЖГр1.5Д2.5Н3	1,5	2,5	3,0	25,0	800	600	300	297	21,60
530	ЖГр1.5Д2.5Н3	1,5	2,5	3,0	25,0	800	600	300	293	26,50
531	ЖГр1.5Д2.5Н3	1,5	2,5	3,0	25,0	800	600	300	297	16,50
532	ЖГр1.5Д2.5Н3	1,5	2,5	3,0	25,0	800	600	300	270	22,90
533	ЖГр1.5Д2.5Н3	1,5	2,5	3,0	25,0	800	600	300	292	21,80
534	ЖГр1.5Д2.5Н3	1,5	2,5	3,0	25,0	800	600	300	288	29,00

Таблица 2. Парная корреляция факторных признаков

	<i>C,%</i>	<i>CU,%</i>	<i>CR,%</i>	<i>NI,%</i>	<i>ПОРЫ</i>	<i>ТНЗ</i>	<i>VOXЛ</i>	<i>ТОТП</i>
<i>C,%</i>	1							
<i>CU,%</i>	0,472097	1						
<i>CR,%</i>	-0,61261	-0,76518	1					
<i>NI,%</i>	-0,06999	0,222701	-0,22123	1				
<i>ПОРЫ</i>	0,216219	0,375214	-0,38869	0,345398	1			
<i>ТНЗ</i>	0,14266	0,14078	-0,25974	0,149054	0,107496	1		
<i>VOXЛ</i>	0,193049	0,117707	-0,13658	-0,15371	-0,03078	0,020509	1	
<i>ТОТП</i>	0,156134	0,058632	-0,08438	-0,00266	0,080922	-0,02974	-0,04207	1

Для анализа табличных данных использовалась система Statistica 6.0 (StatSoft). Перед этим была проведена проверка данных на целостность и избыточность с помощью СУБД Access [1]. Также с помощью Access были сформированы выборочные БД, отражающие последствия изменения одного исходного параметра при постоянстве остальных. В системе Statistica были исследованы корреляционные зависимости между исходными и зависимыми параметрами ИПМ, выделены кластеры исходных параметров, исследовано их влияние на конечные свойства ИПМ, и на основании этого проведено их ранжирование, проведены регрессионные исследования и получены зависимости между зависимыми и независимыми параметрами. Результаты анализа представлены ниже.

Для определения тесноты связи между факторными и результативными и факторными признаками в системе Statistica были проведены корреляционные исследования.

Проверка парной корреляции между двумя конструкционно-технологическими параметрами ИПМ позволяет узнать, не коррелируют ли они между собой, т. е. не указывают ли некоторые поля таблицы экспериментальных данных, по сути, на одни и те же признаки (таблица 2).

Таблица 3. Парная корреляция факторных и результативных признаков

	<i>НВ</i>	<i>б</i>
<i>C</i>	0,105	-0,368
<i>Сu</i>	-0,115	-0,05
<i>Cr</i>	-0,121	0,229
<i>Ni</i>	-0,06	0,242
<i>Пор</i>	-0,192	-0,071
<i>Тнз</i>	0,23	-0,27
<i>Вох</i>	0,24	-0,16
<i>Тотп</i>	-0,458	-0,014

Результаты парной корреляции не выявили существенных зависимостей между факторными признаками. Поэтому все переменные должны быть учтены при построении множественных зависимостей [2].

Проверка парной корреляции между факторными и результативными признаками позволяет оценить близость зависимости к линейной, прямой или обратной (таблица 3).

Наиболее отчетливая обратная зависимость наблюдается между твердостью НВ и температурой отпуска Тотп и пределом прочности б и содержанием углерода С, но, тем не менее, ее величина позволяет говорить о нелинейной зависимости между признаками. Между остальными факторными и результативными признаками наблюдается значительная нелинейная зависимость.

Результаты корреляционного исследования указывают на полную самостоятельность независимых факторов при формировании свойств ИПМ. Следовательно, регрессионные модели должны учитывать все независимые переменные.

Для построения регрессионной модели процесса формирования свойств ИПМ проводим исследование данных в модуле системы Statistica – «множественная регрессия».

Проверим условия применимости регрессионного анализа. Зависимый параметр должен иметь нормальное распределение и являться случайной величиной. Гипотезу о нормальном распределении проверяем с помощью основных статистических описаний случайной величины. Для нормального распределения среднее значение и медиана выборки должны быть равны, эксцесс и асимметричность выборки должны быть равны нулю (таблица 4).

На основании представленных оценок можно сделать вывод о близости распределения результативных показателей к нормальному и о случайности выборки. Это соответствует утверждению, что при значительной величине выборки изменения случайной величины закон ее распределения соответствует нормальному. Следовательно, можно применять регрессионный анализ. Результаты анализа представлены в таблицах: твердости НВ (таблица 5, 6) и предела прочности (таблица 7, 8).

Аппроксимирующая функция зависимости НВ от параметров X1-X8 выглядит следующим образом:

$$НВ = 312,1 + 9,5 * X1 - 29,2 * X2 - 32 * X3 + 20,73 * X4 - 5,1 * X5 + 0,13 * X6 + 0,08 * X7 - 0,3 * X8$$

Аппроксимирующая функция зависимости предела прочности б от параметров X1-X8 выглядит следующим образом:

$$б = 11,41 - 16,2 * X1 + 2,2 * X2 - 7,6 * X3 + 2,7 * X4 - 2,7 * X5 - 0,02 * X6 - 0,01 * X7 + 0,01 * X8$$

Показатель R-квадрат оценивает коэффициент детерминации. Значение R-квадрат, равное 0,48 для твердости, показывает, что изменчивость значений переменной Y около линии регрессии составляет 1-0,48 от исходной дисперсии; другими словами, 48% от исходной изменчивости могут быть объяснены, а 60% остаточной изменчивости остаются необъясненными. В идеале желательно иметь

Таблица 4. Основные статистические показатели

НВ		б	
Среднее	240,6233766	Среднее	35,09685065
Стандартная ошибка	4,29836152	Стандартная ошибка	0,850240986
Медиана	239	Медиана	34,885
Мода	185	Мода	30,5
Стандартное отклонение	75,43593852	Стандартное отклонение	14,92166875
Дисперсия выборки	5690,58082	Дисперсия выборки	222,6561982
Эксцесс	-0,00506524	Эксцесс	0,003844639
Асимметричность	0,00385628	Асимметричность	0,002216924
Минимум	90	Минимум	10,3
Максимум	467	Максимум	94,1
Сумма	74112	Сумма	10809,83
Счет	308	Счет	308

Таблица 5. Оценка модели твердости

Регрессионная статистика	
Множественный R	0,69151686
R-квадрат	0,47819557
Нормированный R-квадрат	0,46423425
Стандартная ошибка	55,2161056
Наблюдения	308

объяснение если не для всей, то хотя бы для большей части исходной изменчивости. Значение R-квадрата является индикатором степени подгонки модели к данным (значение R-квадрата, близкое к 1,0, показывает, что модель объясняет почти всю изменчивость соответствующих переменных). Значение R-квадрат для предела прочности составляет 0,51, т. е. адекватность модели немного выше, но, тем не менее, остается довольно низкой.

Статистическая значимость результата P-значение представляет собой оцененную меру уверенности в его «истинности». P-уро-

Таблица 6. Коэффициенты регрессии для твердости ИПМ

	Коэффициенты	Стандартная ошибка	t-статистика	P-Значение
Y-пересечение	312,0759558	47,91066077	6,513705942	3,10085E-10
Переменная X 1	9,468777665	15,79671154	0,599414482	0,549350445
Переменная X 2	29,18188266	5,293312272	5,51297206	7,63117E-08
Переменная X 3	32,00050268	6,406229561	4,995216356	1,00122E-06
Переменная X 4	20,73412177	3,025447086	6,85324224	4,12846E-11
Переменная X 5	5,027048874	0,883300697	5,691208999	3,00888E-08
Переменная X 6	0,12600184	0,047161833	2,671690889	0,007960671
Переменная X 7	0,083577255	0,015091224	5,538136187	6,70064E-08
Переменная X 8	0,287743602	0,027081352	10,62515656	1,39149E-22

Таблица 7. Оценка модели предела прочности

Регрессионная статистика	
Множественный R	0,714064703
R-квадрат	0,509888399
Нормированный R-квадрат	0,496775046
Стандартная ошибка	10,58518565
Наблюдения	308

вень – это показатель, находящийся в убывающей зависимости от надежности результата. Более высокий р-уровень соответствует более низкому уровню доверия к найденной в выборке зависимости между переменными. Р-уровень представляет собой вероятность ошибки, связанной с распространением наблюдаемого результата на всю популяцию. Наихудшие значения для твердости: р-уровень = 0,549, имеет вторая переменная X2, это означает, что имеется 5,5% вероятность, что найденная в выборке связь между переменными является лишь случайной, и для предела прочности X7 р-уровень = 0,13 и X8 р-уровень = 0,14. Следовательно, зависимость между этими факторными признаками и результативными хуже всего описывается линейной моделью.

Построение линейной зависимости между факторными и результативными признаками показало низкую адекватность полученной модели, на это же указывают исследования о парной корреляции, описанные в предыдущем параграфе, поэтому необходимо провести нелинейное оценивание [3].

На основании данных о множественной регрессии строим несколько гипотез о виде зависимости между параметрами ИПМ. После проверки гипотез в модуле системы Statistica, «Нелинейное оценивание», была получена следующая регрессионная функция:

$$HB = 112,4 + 1,5 * X1 - 9,2 * X2 + 1,2 * X23 - 5,2 * X32 + 0,73 * X42 + 2,1 * X5 + 0,13 * X63 + 0,08 * X7 + 0,05 * X72 - 0,3 * X8 + 0,12 * X82$$

Коэффициент детерминации составил R2 = 0,82.

Наилучшие результаты были получены с помощью регрессионных моделей с точками разрыва. Это оценивание проводится в предположении, что вид зависимости между предикторами и переменной отклика различается в разных областях значений независимых переменных. Определенная зависимость имеет силу до некоторого значения результирующей переменной, после которого более точно зависимость отражает другая модель [4]. Аппроксимирующие функции с точкой разрыва представлены ниже.

$$HB = 187,3895 + 16,16783 * X + 15,2180 * X^2 + 4,3556 * X^3 + 13,63748 * X^4 + 1,11216 * X^5 + 0,037329 * X^6 + 0,02652 * X^7 - 0,04898 * X^8$$

(точка разрыва HB = 240)

$$\text{and } HB = 349,3814 + 17,18680 * X + 14,1733 * X^2 + 12,0331 * X^3 + 7,195041 * X^4 - 5,32342 * X^5 + 0,12559 * X^6 + 0,039114 * X^7 - 0,290733 * X^8$$

Значение коэффициента детерминации R = 0,88691 указывает на значительную адекватность представленной модели и возможность ее использования при формировании правил БЗ.

Для уточненной идентификации зависимости был проведен кластерный анализ, позволяющий разбить исследуемое множество

Таблица 8. Коэффициенты регрессии для предела прочности б

	Коэффициенты	Стандартная ошибка	t-статистика	P-Значение
Y-пересечение	11,4094557	9,184697709	12,4222441	7,52359E-29
Переменная X 1	16,23704984	3,028303471	-5,361764432	1,65124E-07
Переменная X 2	2,198218203	1,014752715	2,166259987	0,031081262
Переменная X 3	7,551753706	1,228104163	-6,149114978	2,49088E-09
Переменная X 4	2,663121039	0,579992354	4,591648529	6,48559E-06
Переменная X 5	-2,712414014	0,169332874	-16,01823644	4,00503E-42
Переменная X 6	-0,017376697	0,009041144	-1,921957842	0,055561221
Переменная X 7	-0,004399464	0,002893058	-1,520696533	0,129392777
Переменная X 8	0,007634306	0,005191622	1,470505048	0,142476553

Таблица 9. Кластеры и дистанции до центров кластеризации

Кластер 1, две переменные						
	Переменная	Переменная				
	X6	X7				
Дистанция	185,5006	185,5006				
Кластер 2, шесть переменных						
	Переменная	Переменная	Переменная	Переменная	Переменная	Переменная
	X1	X2	X3	X4	X5	X8
Дистанция	55,73336	55,17973	56,5735	56,13316	40,35888	263,4574

параметров процесса производства ИПМ на несколько кластеров. Отнесение точки к тому или иному кластеру происходит на основе меры близости, соответствующей расстоянию точки до центра кластера.

Выберем К произвольных центров в пространстве переменных, для разбиения записей на соответствующее количество кластеров.

Число разбиений К = 2 (таблица 9).

В первый кластер системой выделены переменные: X6 – температура нагрева под закалку Tнз = 750..900° С, X7 – скорость охлаждения Vох = 200..800° С.

Во второй объединены переменные: X1 – содержание углерода С = 1,0..1,5%, X2 – содержание меди Cu = 0..2,5%, X3 – содержание хрома Cr = 0..3,0%, X4 – содержание никеля Ni = 0..3,0%, X5 – пористость Пор = 10..20%, X8 – температура отпуска Тотп = 150..300° С.

При увеличении числа разбиений К = 3 выделяются следующие кластеры (таблица 10):

Первый кластер, переменные: X6 – температура нагрева под закалку Tнз = 750..900° С, X7 – скорость охлаждения Vох = 200..800° С.

Второй кластер: X8 – температура отпуска Тотп = 150..300° С.

Третий кластер: X1 – содержание углерода С = 1,0..1,5%, X2 – содержание меди Cu = 0..2,5%, X3 – содержание хрома Cr = 0..3,0%, X4 – содержание никеля Ni = 0..3,0%, X5 – пористость Пор = 10..20%, (рисунок 1).

На основе результатов разбиения множества пара-

метров процесса производства ИПМ на два кластера проведем регрессионное исследование и определим влияние параметров каждого кластера на формирование свойств ИПМ.

Линейная аппроксимация:

Кластер 1

$$НВ = 0,188364 * T_{нз} + 0,212258 * V_{ох} - 0,45803 * T_{отп}$$

Кластер 2

$$НВ = -0,0126 * C - 0,35432 * Cu - 0,44719 * Cr + 0,300255 * Ni - 0,31547 * Поры$$

Коэффициент детерминации при линейном оценивании:

Кластер 1 R = 0,55. Кластер 2 R = 0,44

Полученная зависимость при линейной аппроксимации малозначима.

Нелинейное оценивание:

Кластер 1

$$НВ = 118,7712 + 0,090258 * T_{нз} + 0,009623 * V_{ох} - 0,02894 * T_{отп} \text{ (Breakpt.240,6234) and}$$

$$НВ = 248,5278 + 0,116646 * T_{нз} + 0,048201 * V_{ох} - 0,25209 * T_{отп}$$

Кластер 2

$$НВ = 243,9059 - 6,4515 * C - 16,5482 * Cu - 21,2777 * Cr + 9,988771 * Ni - 0,99723 * Поры \text{ (Breakpt.240,6234) and}$$

Таблица 10. Кластеры и дистанции до центров кластеризации

Кластер 1 состоит из двух переменных					
	Переменная	Переменная			
	X6	X7			
Дистанция	185,5006	185,5006			
Кластер 2 состоит из пяти переменных					
	Переменная	Переменная	Переменная	Переменная	Переменная
	X1	X2	X3	X4	X5
Дистанция	55,73336	55,17973	56,5735	56,13316	40,35888
Кластер 3 состоит из одной переменной					
	Переменная				
	X8				
Дистанция	0				

$NB = 416,651 - 9,78042 * C - 10,2767 * Cu -$
 $- 15,9128 * Cr - 1,58751 * Ni - 4,15808 * \text{Поры}.$

Коэффициент детерминации при нелинейном оценивании:

Кластер 1 $R = 0,85$. Кластер 2 $R = 0,84$

Значение коэффициента детерминации при нелинейной аппроксимации указывает на значительную адекватность представленной модели. Следовательно, она может ис-

пользоваться при формировании базы знаний ЭС ИПМ.

Для получения прогнозов на основе проведенного анализа данных используется комплекс поддержки принятия решений, включающий оболочку экспертной системы и базу знаний, обеспечивающий в диалоговом режиме принятие решений на основе концентрации известных на данный момент времени знаний.

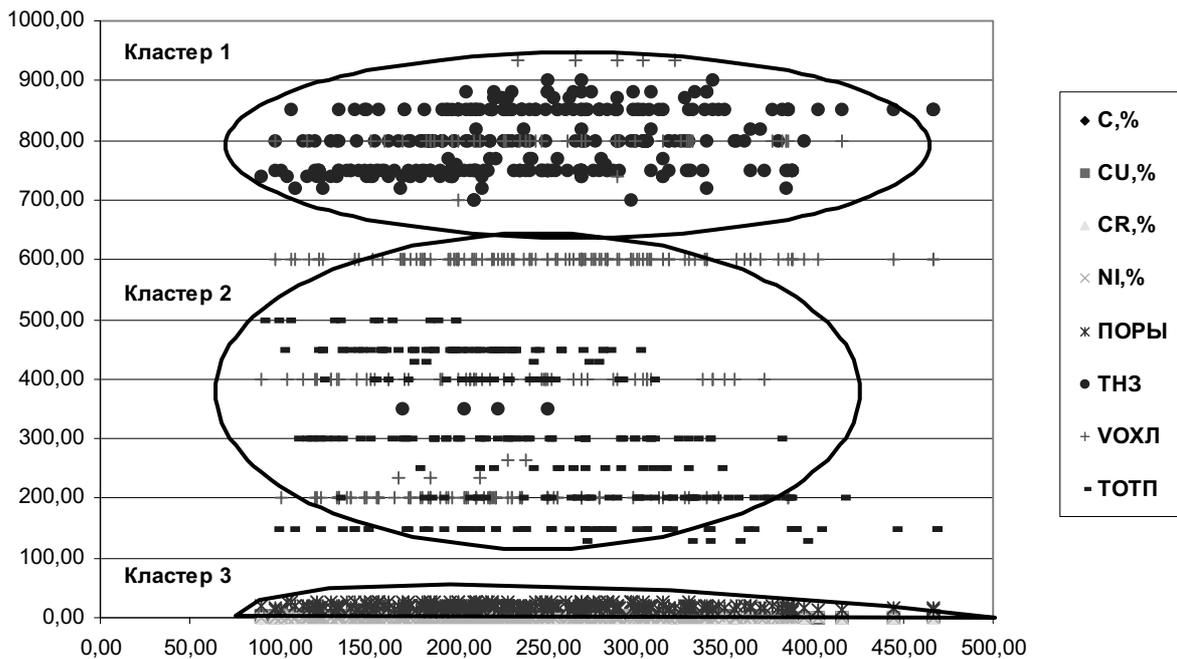


Рисунок 1. Результаты кластерного анализа.

Список использованной литературы:

1. STATISTICA (c) Copyright StatSoft, Inc., 1984-1998 StatSoft, Inc. – Спб.: БХВ-Петербург, 2001.
2. Таусенд К., Фохт Д. Проектирование и программная реализация экспертных систем на персональных ЭВМ: Пер. с англ. – М.: Финансы и статистика, 1996.– 320 с.
3. К. П. Нейлор. Как построить свою экспертную систему.– М.: Энергоатомиздат, 1991.
4. В.О. Сафонов. Экспертные системы – интеллектуальные помощники специалистов. – СПб: Санкт-Петербургская организация общества «Знания» России, 2002.