МЕТАЛЛУРГИЯ

The mathematical model of multidimensional regressive analysis is presented. Its practical application for rolling production is treated.

А. Н. ЧИЧКО, БНТУ, Л. А. ФЕКЛИСТОВА, В. И. ЩЕРБАКОВ, А. В. ВЕДЕНЕЕВ, РУП «БМЗ»

УДК 629.113

МНОГОМЕРНЫЙ РЕГРЕССИОННЫЙ АНАЛИЗ ТЕХНОЛОГИИ ПРОЦЕССА ПРОКАТКИ

Известно, что свойства катанки формируются на этапах выплавки стали, получения литой заготовки и заготовки в процессе прокатки. На свойства катанки влияет множество факторов, например, химический состав стали, наличие дефектов после выплавки и в процессе получения литой заготовки, микроструктура перлита катанки после сорбитизации с проката, технология изготовления катанки. Присутствие многоступенчатости технологических переделов приводит к усложнению поиска причин ухудшения качества продукции при массовом производстве. Непрерывность технологического процесса не позволяет останавливать технологические агрегаты и линии для проведения исследований причин и поиска оптимальных технологических режимов. В таких ситуациях, как правило, используют различные методы статистического анализа. Проработка литературных данных показала, что для сложных технологических циклов отсутствуют готовые методы регрессионного анализа. А учитывая тот факт, что в условиях реального производства и в борьбе за рынки сбыта происходит постоянная смена сортамента выпускаемой продукции при прокатке, требуется всегда проводить статистический анализ технологических параметров и характеристик используемых материалов для улучшения качества продукции. В данной работе предлагается собственная схема проведения анализа с использованием специальных характеристик. *Специальная характеристика* – это характеристики продукции и процессов, назначенные потребителем, описывающие безопасность и правительственные нормы и/или назначенные поставщиком благодаря знаниям о продукции и процессе, требующие мониторинга и внесения в планы управления и другие технологические документы.

При определении влияния на зависимую переменную нескольких факторов можно использовать многофакторный дисперсионный анализ, с помощью которого можно определить степень зависимости между характеристиками и в результате последовательного анализа выбрать специальные характеристики. Главное преимущество этого метода в том, что он позволяет исследователю изучать взаимодействие факторов. Взаимодействия (interaction) имеют место, когда эффекты одного фактора на зависимую переменную зависят от уровня других факторов.

Цель проведения многомерного регрессионного анализа – исследование влияния химического состава стали на величину прочности, относительного сужения и относительного удлинения стали 65К, предназначенной для изготовления металлокорда; выбор специальных характеристик катанки, которые могут повлиять на безопасность или соответствие правительственным нормам, установку, функцию, работоспособность или последующую переработку продукции.

Задачей анализа является построение математической модели и выбор одного или нескольких оптимальных параметров из многих выходных параметров, при этом другие служат ограничениями, которые бы минимизировали сумму квадратов отклонений наблюдаемых точек от поверхности (регрессионная поверхность выражает наилучшее предсказанное значение зависимой переменной y для заданных значений независимых переменных x). Всегда полезно исследовать возможность уменьшения выходных параметров, так как невозможно построить математические модели, одновременно оптимизировав несколько функций [1, 2].

Анализ зависимости химического состава стали от величины прочности, относительного сужения и относительного удлинения проволоки из стали 65К, предназначенной для изготовления металлокорда по ЗТУ 840-03

Для многомерного регрессионного анализа воспользуемся моделью черного ящика (рис. 1).

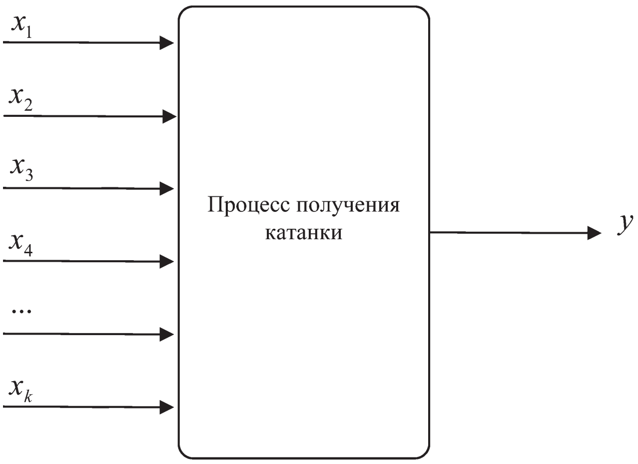

Рис. 1. Модель черного ящика для производства катанки в условиях РУП «БМЗ»

Под моделью будем понимать вид функции отклика. При построении многомерных моделей наглядность представления теряется и приходится выражаться на языке алгебры. Модель должна быть адекватной, т. е. предсказанное с помощью модели значение отклика не должно отличаться от фактического больше, чем на некоторую заданную величину. Выбранная линейная модель будет всегда адекватна по причине аналитичности функции отклика, и всегда существует окрестность любой точки, в которой линейная модель адекватна. Математическая модель нужна для предсказания направления, в котором величина параметра оптимизации улучшается быстрее, чем в любом другом направлении. В исследуемом случае уже известен допуск на рассматриваемые характеристики и, следовательно, известна область факторного пространства. Эта модель имеет конечное число опытов, позволяющее получить выборочные оценки для коэффициентов уравнения [3–5]. Их точность и надежность зависят от свойств выборки и нуждаются в статистической проверке. Задачей здесь является вычисление значений коэффициентов модели.

В табл. 1 приведены химический состав и механические свойства стали по применяемой на РУП «БМЗ» марке стали для изготовления металлокорда и проволоки в соответствии с заводскими техническими условиями, в табл. 2 – статистические данные по механическим и химическим характеристикам катанки марки стали 65К для производства метизной продукции.

Таким образом, функция является линейной и система N линейных уравнений (алгебраических полиномов первой степени) и уравнение множественной регрессии (число независимых переменных ≥ 2) имеет следующий вид [6]:

$$y_i = B_0 + B_1 x_1 + B_2 x_2 + \ldots + B_k x_k \,,$$

где B_0, B_1, B_2, …, B_k – свободные коэффициенты; x_1, x_2, …, x_k, y_i – переменные, представляющие собой два массива чисел. Применительно к исследуемым характеристикам переменные x_1, x_2, …, x_k – это процентное содержание химических элементов в стали, а y_1, y_2, y_3 – это значение временного сопротивления разрыву σ_b, относительного сужения ψ и относительного удлинения δ соответственно. Тогда система линейных уравнений примет вид

$$\begin{cases} y_1 = B_{10} + B_{11}x_{11} + B_{12}x_{12} + \ldots + B_{19}x_{19}, \\ y_2 = B_{20} + B_{21}x_{21} + B_{22}x_{22} + \ldots + B_{29}x_{29}, \\ y_3 = B_{30} + B_{31}x_{31} + B_{32}x_{32} + \ldots + B_{39}x_{39}. \end{cases}$$

На рис. 2 показан алгоритм построения модели множественной регрессии. Используя этот алгоритм, рассмотрим пошагово его практическое применение на примере прокатного производства.

Шаг 1. Построим входную расширенную матрицу для входных переменных (x_1, x_2, x_3, …, x_9) и одной выходной (y) для n = 89 наблюдений. Для данного исследования матрица имеет вид

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} & x_{15} & x_{16} & x_{17} & x_{18} & x_{19} & y_1 \\ x_{21} & x_{22} & x_{23} & x_{24} & x_{25} & x_{26} & x_{27} & x_{28} & x_{29} & y_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ x_{91} & x_{92} & x_{93} & x_{94} & x_{95} & x_{96} & x_{97} & x_{98} & x_{99} & y_9 \end{bmatrix}.$$

По исходным данным (табл. 2) проведем вычисления средних значений и дисперсии для всех переменных.

Шаг 2. Преобразуем исходную матрицу к стандартизированному виду, используя средние ариф-

Таблица 1. **Химический состав и механические свойства стали по ЗТУ 840-03-2006**

<table>
<tr><td rowspan="3">Марка стали</td><td colspan="10">Массовая доля элементов, %</td></tr>
<tr><td rowspan="2">C</td><td rowspan="2">Mn</td><td>Si</td><td>P</td><td>S</td><td>Cr</td><td>Ni</td><td>Cu</td><td>Al</td><td>N_2</td></tr>
<tr><td colspan="8">не более</td></tr>
<tr><td>65К</td><td>0,67–0,71</td><td>0,45–0,55</td><td>0,30</td><td>0,015</td><td>0,015</td><td>0,06</td><td>0,06</td><td>0,07</td><td>0,004</td><td>0,007</td></tr>
<tr><td colspan="11">Требования по механическим свойствам стали</td></tr>
<tr><td colspan="3">$\sigma = 910 - 1090$, Н/мм^2</td><td colspan="4">ψ , не менее 38%</td><td colspan="4">δ , не менее 12%</td></tr>
</table>

Т а б л и ц а 2. **Исходная матрица для исследуемых характеристик катанки диаметром 5,5 мм марки стали 65К**

№ п/п	Значение переменных											
	x_1(C)	x_2(Si)	x_3(Mn)	x_4(P)	x_5(S)	x_6(Cr)	x_7(Ni)	x_8(Cu)	x_9(N_2)	y_1(σ_b)	y_2(ψ)	y_3(δ)
1	0,682	0,1979	0,495	0,00483	0,0112	0,02673	0,02898	0,04798	0,0053	970	47,0	15,5
2	0,6878	0,2144	0,502	0,00394	0,00996	0,02572	0,0266	0,05044	0,0053	980	49,0	18,0
3	0,6845	0,2123	0,5179	0,00498	0,01103	0,02813	0,02888	0,05188	0,00495	990	47,0	16,0
4	0,6923	0,2046	0,50105	0,00548	0,01253	0,02663	0,0289	0,04448	0,0053	970	44,0	15,5
5	0,677	0,1989	0,5078	0,0061	0,0129	0,0216	0,0241	0,043	0,0053	980	52,0	17,5
6	0,678	0,205	0,5098	0,0064	0,0093	0,0367	0,027	0,0395	0,0053	960	44,0	17,0
7	0,679	0,2033	0,51277	0,00603	0,00983	0,03823	0,03123	0,04547	0,00557	950	46,0	17,5
8	0,687	0,2041	0,5051	0,0067	0,0133	0,0368	0,0287	0,0471	0,0057	980	45,0	16,0
9	0,6730	0,1941	0,4718	0,0040	0,0113	0,0265	0,0326	0,0501	0,0053	960	51,0	18,0
10	0,6900	0,1992	0,4830	0,0038	0,0122	0,0355	0,0349	0,0484	0,0055	990	48,0	18,5
11	0,6820	0,1975	0,4869	0,0038	0,0115	0,0265	0,0319	0,0480	0,0047	960	48,0	19,0
12	0,6820	0,2057	0,4899	0,0035	0,0110	0,0328	0,0297	0,0471	0,0059	980	48,0	18,5
13	0,6720	0,2029	0,4894	0,0035	0,0114	0,0328	0,0299	0,0471	0,0059	950	48,0	17,5
14	0,6960	0,2021	0,4878	0,0037	0,0105	0,0351	0,0313	0,0474	0,0050	980	49,0	18,5
–	–	–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–	–	–
86	0,6880	0,2234	0,4949	0,0064	0,0096	0,0328	0,0314	0,0475	0,0047	970	46,0	15,5
87	0,6850	0,1955	0,5347	0,0053	0,0134	0,0328	0,0306	0,0537	0,0059	980	49,0	16,0
88	0,6840	0,1964	0,5378	0,0058	0,0133	0,0305	0,0321	0,0459	0,0059	1010	49,0	16,0
89	0,6700	0,1848	0,5494	0,0064	0,0103	0,0197	0,0281	0,0370	0,0050	940	50,0	15,5
Сумма $\left(\sum_{i=1}^{n} x_i\right)$	0,68417	0,20673	0,5027676	0,005407	0,011142	0,027526	0,028561	0,046072	0,005341	991,4606	0,684177	0,20673
Среднее значение $(\bar{x})$	0,00775	0,01219	0,0198072	0,001339	0,001811	0,005436	0,003209	0,007280	0,000572	26,04849	0,007755	0,01219
Среднеквадратическое отклонение (S)	60,8918	18,3993	44,74632	0,48126	0,99165	2,44984	2,54199	4,10045	0,47539	88240	60,8918	18,3993

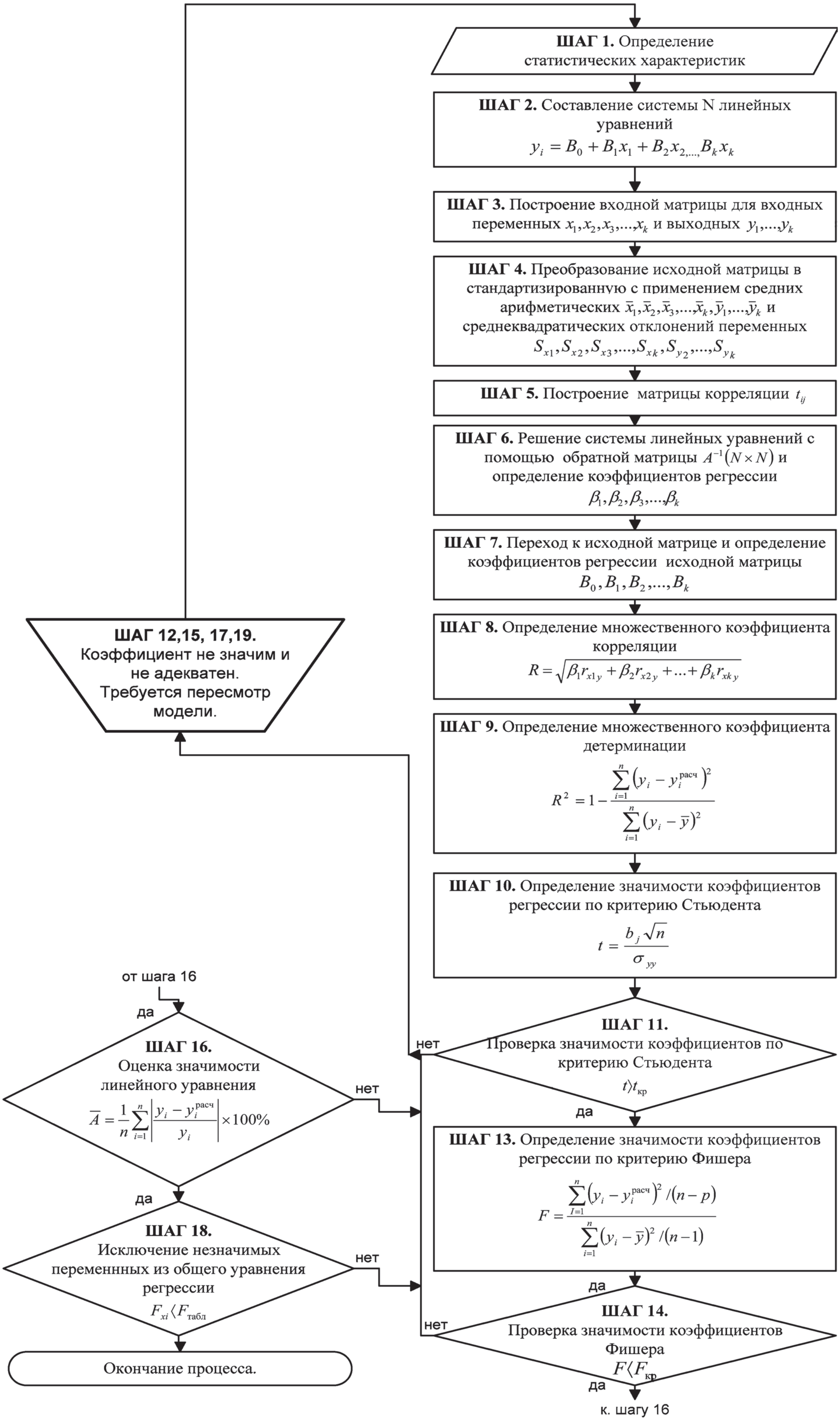

Рис. 2. Алгоритм определения специальных характеристик с помощью математической модели

метические $\overline{x}_1, \overline{x}_2, \overline{x}_3, ..., \overline{x}_9, \overline{y}_1, \overline{y}_2, \overline{y}_3$ и среднеквадратические отклонения переменных S_{x1}, S_{x2}, S_{x3}, …, S_{x9}, S_{y1}, S_{y2}, S_{y3} (табл. 3) по следующей формуле[6–8]:

$$t_{11} = \frac{x_{11} - \overline{x}_1}{S_{x1}} ... t_{91} = \frac{x_{91} - \overline{x}_1}{S_{x1}},$$

$$\dots \dots \dots$$

$$t_{19} = \frac{x_{19} - \overline{x}_9}{S_{x9}} ... t_{99} = \frac{x_{99} - \overline{x}_9}{S_{x9}}.$$

Полученная матрица будет иметь следующий вид:

$$\begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14}\, t_{15} & t_{16} & t_{17} & t_{18}\, t_{19} & t_{20} \\ t_{21} & t_{22} & t_{23} & t_{24}\, t_{25} & t_{26} & t_{27} & t_{28}\, t_{29} & t_{21} \\ ... & ... & ... & ...\ ... & ... & ... & ...\ ... & ... \\ t_{91} & t_{92} & t_{93} & t_{94}\, t_{95} & t_{96} & t_{97} & t_{98}\, t_{99} & t_{100} \end{bmatrix}.$$

В новых переменных уравнение имеет вид

$$t = \beta_1 t_1 + \beta_2 t_2 + \beta_3 t_3 + ... + \beta_k t_k .$$

Здесь в качестве переменной y применяется переменная t, а в качестве переменных x_1, x_2, …, x_k – переменные t_1, t_2, …, t_k.

Шаг 3. Представим эту систему линейных уравнений в матричном виде. Вычислим частные коэффициенты корреляции между переменными x_i и x_j для новой стандартизированной матрицы t_{ij} (табл. 4).

Составим систему линейных уравнений, введя новые переменные β_1, β_2, β_3, …, β_9, являющиеся неизвестными.

Применительно к рассматриваемому случаю уравнение имеет следующий вид:

$$\begin{cases} r_{x1y} = r_{x1x1}\beta_1 + r_{x1x2}\beta_2 + r_{x1x3}\beta_3 + ... + r_{x1x9}\beta_9, \\ r_{x2y} = r_{x2x1}\beta_1 + r_{x2x2}\beta_2 + r_{x2x3}\beta_3 + ... + r_{x2x9}\beta_9, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ r_{x9y} = r_{x9x1}\beta_1 + r_{x9x2}\beta_2 + r_{x9x3}\beta_3 + ... + r_{x9x9}\beta_9. \end{cases}$$

В качестве основной формулы для расчета коэффициента корреляции r_{xy} для двух переменных x и y используют формулу [1]:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}},$$

где $\overline{x}, \overline{y}$ – соответственно средние значения переменных x_i и y_i.

Линейный коэффициент корреляции находится в пределах: $-1 \le r_{xy} \le 1$. Система уравнений формируется из матрицы корреляции (табл. 5):

$$\begin{bmatrix} r_{x1x1} & r_{x1x2} & r_{x1x3} & ...r_{x1x9} & r_{x1y} \\ r_{x2x1} & r_{x2x2} & r_{x2x3} & ...r_{x2x9} & r_{x2y} \\ ... & ... & ... & ...\ ... & ... \\ r_{x9x1} & r_{x9x2} & r_{x9x3} & ...r_{x9x9} & r_{x9y} \end{bmatrix}.$$

Шаг 4. Решим систему линейных уравнений с помощью обратной матрицы [9, 10].

В результате получены значения коэффициентов стандартизированного уравнения β_1, β_2, β_3, …, β_9 (табл. 5) путем перемножения массивов обратной матрицы и коэффициентов корреляции.

Шаг 5. Далее решая систему, определяем регрессионные коэффициенты $\beta_1, \beta_2, \beta_3, ..., \beta_9$, по которым осуществляется переход к исходным коэффициентам модели $B_0, B_1, B_2, ..., B_9$ (табл. 5). При этом используются формулы [6]

$$B_1 = \beta_1 \frac{\sigma_y}{\sigma_{x1}}, ..., \ B_9 = \beta_9 \frac{\sigma_y}{\sigma_{x9}} \text{ и } B_0 = \overline{y} - \sum_{j=1}^{9} B_9 \overline{x}_9 .$$

Шаг 6. Определение множественного коэффициента корреляции.

Множественный коэффициент корреляции определяется как:

$$R = \sqrt{\beta_1 r_{x1y} + \beta_2 r_{x2y} + ... + \beta_9 r_{x9y}} .$$

Полученные в результате исследований множественные коэффициенты корреляции приведены в табл. 6. Множественный коэффициент корреляции считается значительным, т. е. имеет место статистическая зависимость между y и остальными факторами x_1, x_2, …, x_k, если $F_{\text{набл}} \rangle F_{\text{кр}} \left(\alpha, k-1, n-k\right)$, где $F_{\text{кр}}$ определяется по таблице F-распределения. Сила связи между химическим составом и пределом прочности, относительным удлинением является умеренной, а с относительным сужением – слабой.

Шаг 7. Определение множественного коэффициента детерминации (табл. 6).

Множественный коэффициент детерминации при линейной зависимости определяется как [1]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - y_i^{\text{расч}}\right)^2}{\sum_{i=1}^{n}\left(y - \overline{y}\right)^2},$$

где $\sum_{i=1}^{n}\left(y_i - y_i^{\text{расч}}\right)^2$ – остаточная сумма квадратов отклонений; $\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2$ – общая сумма квадратов отклонений. Коэффициент детерминации R^2 является суммарной мерой общего качества уравнения

Т а б л и ц а 3. **Стандартизированная матрица t_{ij}**

№ п/п	Значения переменных											
	$t_i(\mathrm{C})$	$t_i(\mathrm{Si})$	$t_i(\mathrm{Mn})$	$t_i(\mathrm{P})$	$t_i(\mathrm{S})$	$t_i(\mathrm{Cr})$	$t_i(\mathrm{Ni})$	$t_i(\mathrm{Cu})$	$t_i(\mathrm{N_2})$	$t_j(\sigma_b)$	$t_j(\psi)$	$t_j(\delta)$
1	–0,2807718	–0,7245985	–0,3921613	–0,4309911	0,0319405	–0,1464777	0,1303363	0,2620227	–0,0723918	–0,8238739	–0,4540516	–0,7774914
2	0,4670838	0,6288395	–0,0387555	–1,0952995	–0,6525165	–0,3322668	–0,6112119	0,5999343	–0,0723918	–0,4399746	0,2806864	0,8700499
3	0,0415797	0,4565837	0,7639805	–0,3190290	–0,0618963	0,1110519	0,0991788	0,7977363	–0,6835045	–0,0560752	–0,4540516	–0,4479831
4	1,0473166	–0,1750207	–0,0867177	0,0541779	0,7660758	–0,1648726	0,1054103	–0,2187458	–0,0723918	–0,8238739	–1,5561588	–0,7774914
5	–0,9254750	–0,6425720	0,2540664	0,5169545	0,9703089	–1,0901395	–1,3901491	–0,4220423	–0,0723918	–0,4399746	1,3827936	0,5405416
6	–0,7965343	–0,1422101	0,3550395	0,7408787	–1,0168242	1,6875005	–0,4865820	–0,9028108	–0,0723918	–1,2077733	–1,5561588	0,2110333
7	–0,6675937	–0,2816552	0,5049845	0,4647056	–0,7242741	1,9689435	0,8313798	–0,0827570	0,3990379	–1,5916726	–0,8214206	0,5405416
8	0,3639313	–0,2160340	0,1177528	0,9648029	1,1911015	1,7058955	0,0430954	0,1411438	0,6260226	–0,4399746	–1,1887897	–0,4479831
9	–1,4412375	–1,0362994	–1,5634491	–1,0505147	0,0871386	–0,1887861	1,2582374	0,5532311	–0,0723918	–1,2077733	1,0154245	0,8700499
10	0,7507532	–0,6179640	–0,9979998	–1,1997975	0,5839219	1,4667609	1,9748597	0,3197149	0,2768154	–0,0560752	–0,0866825	1,1995582
11	–0,2807718	–0,7574091	–0,8011023	–1,1997975	0,1975349	–0,1887861	1,0401350	0,2647700	–1,1200135	–1,2077733	–0,0866825	1,5290665
12	–0,2807718	–0,0847915	–0,6496427	–1,4237216	–0,0784558	0,9700968	0,3546702	0,1411438	0,9752298	–0,4399746	–0,0866825	1,1995582
13	–1,5701781	–0,3144658	–0,6748859	–1,4237216	0,1423368	0,9700968	0,4169852	0,1411438	0,9752298	–1,5916726	–0,0866825	0,5405416
14	1,5243969	–0,3800870	–0,7556644	–1,2744389	–0,3544465	1,3931810	0,8531901	0,1823525	–0,5962027	–0,4399746	0,2806864	1,1995582
15	–0,5386531	–1,6268905	–1,1090702	0,3676718	–0,1336539	–0,8326099	–0,7046844	–0,5868772	1,3244371	1,8634216	–3,0256350	–0,4479831
–	–	–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–	–	–
83	1,1375750	–0,1914260	–1,1545081	2,4576307	–0,3544465	–1,0901395	–0,2996370	–1,1225907	0,6260226	0,3278242	–1,1887897	–1,1069997
84	–0,5386531	0,0792616	–0,3012855	1,2633684	–1,0168242	–1,5132237	–0,6112119	–0,2984161	–0,0723918	2,2473210	0,2806864	–0,7774914
85	–1,0544156	1,0061616	–0,0892421	1,5619340	–0,9616261	–1,4948287	–0,3307945	–0,9028108	0,2768154	0,7117235	–1,1887897	–1,1069997
86	0,4928719	1,3670784	–0,3972100	0,7408787	–0,8512298	0,9700968	0,8843476	0,1960887	–1,1200135	–0,8238739	–0,8214206	–0,7774914
87	0,1060500	–0,9214622	1,6121544	–0,0801766	1,2462997	0,9700968	0,6350877	1,0477359	0,9752298	–0,4399746	0,2806864	–0,4479831
88	–0,0228906	–0,8476383	1,7686627	0,2930304	1,1911015	0,5470126	1,1024500	–0,0236912	0,9752298	0,7117235	0,2806864	–0,4479831
89	–1,8280593	–1,7991463	2,3543066	0,7408787	–0,4648428	–1,4396438	–0,1438496	–1,2462169	–0,5962027	–1,9755720	0,6480555	–0,7774914

Т а б л и ц а 4. **Матрица корреляции для стандартизированной матрицы t_{ij}**

r	Значение коэффициентов корреляции											
	$r_{xixj}(C)$	$r_{xi\,xj}(Si)$	$r_{xi\,xj}(Mn)$	$r_{xi\,xj}(P)$	$r_{xi\,xj}(S)$	$r_{xi\,xj}(Cr)$	$r_{xi\,xj}(Ni)$	$r_{xi\,xj}(Cu)$	$r_{xi\,xj}(N_2)$	$r_{xi\,yi}(\sigma_b)$	$r_{xi\,yi}(\psi)$	$r_{xi\,yi}(\delta)$
x_1	1	0,21808823	0,08849629	–0,0905135	0,07637028	–0,0870567	0,00526168	–0,1207405	–0,0637799	–0,0529359	–0,1848610	–0,0657490
x_2	0,21808823	1	–0,0701401	0,04526514	0,29436751	–0,2487202	–0,1216076	0,04946099	0,07434263	0,18971320	0,03273380	–0,2095558
x_3	0,08849629	–0,0701401	1	–0,1275236	0,01915328	0,23394851	–0,0309083	0,03718370	0,02872161	–0,2974568	–0,0998303	–0,1416919
x_4	–0,0905135	0,04526514	–0,1275236	1	0,20316276	0,03591094	0,22139016	–0,0817133	0,04920917	0,26578481	–0,0233879	–0,0183268
x_5	0,07637028	0,29436751	0,01915328	0,20316276	1	–0,1164692	0,13321099	0,06792275	0,14658344	0,12307906	0,01766306	0,02666579
x_6	–0,0870567	–0,2487202	0,23394851	0,03591094	–0,1164692	1	0,44012821	0,16307263	0,11316065	–0,1136324	–0,0735458	0,15334291
x_7	0,00526168	–0,1216076	–0,0309083	0,22139016	0,13321099	0,44012821	1	0,52318285	0,03690367	–0,0672579	–0,02317276	0,05103699
x_8	–0,1207405	0,04946099	0,03718370	–0,0817133	0,06792275	0,16307263	0,52318285	1	0,01203145	–0,0629696	–0,0653732	–0,2547838
x_9	–0,0637799	0,07434263	0,02872161	0,04920917	0,14658344	0,11316065	0,03690367	0,01203145	1	0,02880022	–0,0771234	0,06860365

Т а б л и ц а 5. **Обратная матрица $A^{-1}(9\times9)$**

Значение оцениваемых характеристик (k)									Коэффициенты регрессии стандартизированной системы линейных уравнений			Коэффициенты регрессии исходной системы линейных уравнений			
$x_1(C)$	$x_2(\mathrm{Si})$	$x_3(\mathrm{Mn})$	$x_4(\mathrm{P})$	$x_5(\mathrm{S})$	$x_6(\mathrm{Cr})$	$x_7(\mathrm{Ni})$	$x_8(\mathrm{Cu})$	$x_9(\mathrm{N}_2)$	β_σ	β_ψ	β_δ	B_σ	B_ψ	B_δ	B_0
1,15238	–0,29396	–0,15302	0,18708	–0,01369	0,13098	–0,31267	0,31678	0,08545	–0,0346	–0,2417	–0,0670	–116,06	–84,8257	–13,108	–
–0,2939	1,26457	0,09394	–0,09503	–0,32248	0,15587	0,29653	–0,26688	–0,08888	0,1392	0,0889	–0,1649	297,342	19,8386	–20,525	–
–0,1530	0,09394	1,13579	0,09374	–0,12820	–0,37458	0,27846	–0,13357	–0,00147	–0,2706	–0,0565	–0,1850	–355,84	–7,7700	–14,171	–
0,18708	–0,09503	0,09374	1,20395	–0,18832	0,05189	–0,46796	0,37144	–0,00841	0,2429	–0,0951	–0,1439	4723,52	–193,3106	–162,97	–
–0,0136	–0,32248	–0,12820	–0,18832	1,22189	0,23235	–0,27571	0,02896	–0,15953	0,0637	0,0305	0,1259	916,471	45,7825	105,461	–
0,13098	0,15587	–0,37458	0,05189	0,23235	1,49734	–0,78104	0,17701	–0,17183	0,0427	–0,0692	0,1499	204,781	–34,6350	41,8410	–
–0,3126	0,29653	0,27846	–0,46796	–0,27571	–0,78104	2,07311	–1,03999	0,03785	–0,1524	0,1203	0,1738	–1237,1	101,9911	82,1674	–
0,31678	–0,26688	–0,13357	0,37144	0,02896	0,17701	–1,03999	1,59979	0,02046	0,0243	–0,1572	–0,3841	86,9553	–58,7854	–80,068	–
0,08545	–0,08888	–0,00147	–0,00841	–0,15953	–0,17183	0,03785	0,02046	1,05370	0,0032	–0,0920	0,0518	146,489	–437,3666	137,169	–
–	–	–	–	–	–	–	–	–	–	–	–	1177,46	–	–	$B_{0(\sigma)}$
–	–	–	–	–	–	–	–	–	–	–	–	–	109,6971	–	$B_{0(\psi)}$
–	–	–	–	–	–	–	–	–	–	–	–	–	–	36,1809	$B_{0(\delta)}$

Т а б л и ц а 6. **Расчетная таблица для вычислений критерия Стьюдента и Фишера**

<table>
<tr><td colspan="12">Значение коэффициентов Стьюдента t</td></tr>
<tr><td>$t_1(\mathrm{C})$</td><td>$t_2(\mathrm{Si})$</td><td>$t_3(\mathrm{Mn})$</td><td>$t_4(\mathrm{P})$</td><td>$t_5(\mathrm{S})$</td><td>$t_6(\mathrm{Cr})$</td><td>$t_7(\mathrm{Ni})$</td><td>$t_8(\mathrm{Cu})$</td><td>$t_9(\mathrm{N_2})$</td><td colspan="3">$t_{0,05;88}=2,06$</td></tr>
<tr><td>–0,0651</td><td>0,2548</td><td>–0,553</td><td>7,6532</td><td>1,471</td><td>–0,2162</td><td>–1,4632</td><td>0,3267</td><td>4,2146</td><td>$t_{\text{кр}}(\sigma_b)$</td><td>–</td><td>–</td></tr>
<tr><td>0,1110</td><td>0,1675</td><td>–0,3079</td><td>–2,9940</td><td>1,1343</td><td>–0,29</td><td>1,1628</td><td>–0,59</td><td>–8,2492</td><td>–</td><td>$t_{\text{кр}}(\psi)$</td><td>–</td></tr>
<tr><td>–0,0970</td><td>–0,6658</td><td>–0,2921</td><td>–3,8370</td><td>2,5132</td><td>0,064</td><td>3,3498</td><td>–2,06</td><td>6,4932</td><td>–</td><td>–</td><td>$t_{\text{кр}}(\delta)$</td></tr>
<tr><td colspan="6">Значение коэффициентов множественной корреляции R_i</td><td colspan="6">Значение коэффициентов Фишера $F_{\text{кр}}$ = 1,43</td></tr>
<tr><td colspan="2">$R(\sigma_b)$</td><td colspan="2">$R(\psi)$</td><td colspan="2">$R(\delta)$</td><td colspan="3">$F_{\text{набл}}(\sigma_b)$</td><td colspan="2">$F_{\text{набл}}(\psi)$</td><td>$F_{\text{набл}}(\delta)$</td></tr>
<tr><td colspan="2">$R_1=0,43$</td><td colspan="2">$R_2=0,28$</td><td colspan="2">$R_3=0,33$</td><td colspan="3">$F_1=0,9$</td><td colspan="2">$F_2=1,08$</td><td>$F_3=0,89$</td></tr>
<tr><td colspan="6">Значение коэффициентов множественной детерминации R_i^2</td><td colspan="3">–</td><td colspan="2">–</td><td>–</td></tr>
<tr><td colspan="2">$R^2(\sigma_b)$</td><td colspan="2">$R^2(\psi)$</td><td colspan="2">$R^2(\delta)$</td><td colspan="3">–</td><td colspan="2">–</td><td>–</td></tr>
<tr><td colspan="2">$R_1^2=0,18$</td><td colspan="2">$R_2^2=0,08$</td><td colspan="2">$R_3^2=0,11$</td><td colspan="3">–</td><td colspan="2">–</td><td>–</td></tr>
</table>

регрессии (его соответствия статистическим данным).

Шаг 8. Вычислим значимость коэффициентов регрессии по критерию Стьюдента, который вычисляется как [1, 6] и приведен в табл. 6:

$$t=\frac{b_j\sqrt{n}}{\sigma_{yy}},$$

где σ_{yy} – ошибка воспроизводимости данных:

$$\sigma_{yy}=\sqrt{\frac{\sum\limits_{i=1}^{n}y_i^2-n\sum\limits_{j=0}^{k}b_j^2}{n-k-1}},$$

где k – число регрессионных коэффициентов без нулевого; b_j – коэффициент регрессии; n – число наблюдений.

Проверка значимости коэффициентов проводится по неравенству $|t|\geq t_{\text{кр}}$. Она показала, что

$$t_{0,05;88}\langle t\langle t_{0,05;88}\ ,\ 2,06\langle t\langle 2,06\ ,$$

где $t_{0,05;88}=2,06$ взято из таблицы [9, 10]. Так как условие $t_i\rangle t_{0,05;88}$ выполняется (см. табл. 5), то все коэффициенты, влияющие на прочность, являются значимыми кроме химического элемента азота Р; для относительного сужения все коэффициенты являются значимыми; для относительного удлинения – все рассмотренные кроме S, Ni, N_2.

Шаг 9. Вычислить значение критерия Фишера, используя формулу [1, 6, 9, 10]:

$$F=\frac{\sum\limits_{I=1}^{n}\left(y_i-y_i^{\text{расч}}\right)^2/(n-k)}{\sum\limits_{i=1}^{n}\left(y_i-\overline{y}\right)^2/(n-1)}.$$

Критерий Фишера при α = 0,05 для числа степеней свободы k_1 = 79 (число степеней свободы для числителя) и k_2 = 88 (число степеней свободы для знаменателя) находится из таблицы и равен $F_{\text{кр}}$ = 1,43. Наблюдаемые значения коэффициента Фишера $F_{\text{набл}}$ приведены в табл. 6. В этом случае для проверки адекватности используется неравенство $F_{\text{набл}}\langle F_{\text{кр}}$. При выполнении неравенства модель считается значимой и адекватной.

Тогда полученные значения $F_{\text{набл}}\langle$ 1,43 с вероятностью $\gamma=0,95$ могут свидетельствовать об адекватности построенной регрессии. Таким образом, регрессионные уравнения адекватны и имеют вид

$$y(\sigma_B)=1177,5-116,1x_1+297,3x_2-355,8x_3+4723,5x_4+916,5x_5+204,8x_6-1237,2x_7+86,9x_8+146,5x_9, \quad (1)$$

$$y(\psi)=109,7-84,8x_1+19,8x_2-7,8x_3-193,3x_4+45,8x_5-34,6x_6+101,9x_7-58,8x_8-437,4x_9, \quad (2)$$

$$y(\delta)=36,1-13,1x_1-20,5x_2-14,2x_3-162,9x_4+105,5x_5+41,8x_6+82,2x_7-80,1x_8+137,2x_9. \quad (3)$$

Свободные коэффициенты при независимых переменных указывают на силу влияния факторов. Чем больше численная величина коэффициента, тем большее влияние оказывает фактор. Если коэффициент имеет знак плюс, то с увеличением значения фактора параметр оптимизации увеличивается, а если минус, то уменьшается. Исходя из этого, можно сделать следующее заключение: в уравнении (1) существенное влияние на изменение прочности σ_B оказывает фактор x_4 (усиливает), которому соответствует химический элемент фосфор (Р), и x_7 – Ni (уменьшает); в уравнении (2) существенное влияние на изменение относительного сужения ψ оказывает фактор x_7 – Ni (усиливает), а x_9 – N_2 (уменьшает); в уравнении (3) существенное влияние на изменение относительного удлинения δ оказывает фактор x_9 – N_2 (усиливает), а x_4 – Р (уменьшает).

Шаг 10. После того, как найдено уравнение линейной регрессии, проводится оценка значимости как уравнения в целом, так и отдельных его параметров. Чтобы иметь общее суждение о качестве модели из относительных отклонений по каждому наблюдению, определяют среднюю ошибку аппроксимации [11, 12]:

$$\overline{A} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - y_i^{\text{расч}}}{y_i}\right| \cdot 100\%,$$

где $y_i^{\text{расч}}$ – теоретическое значение результативного признака, найденное исходя из уравнения регрессии. Средняя ошибка аппроксимации не должна превышать 8–10%.

Проверим уравнения регрессии $y = b_0 + b_1x_1 + b_2x_2 + ... + b_kx_k$, подставив в формулы средние значения переменных $x_1, x_2, ..., x_k$ и получив при этом средние значения y_1, y_2, y_3:

$y(\sigma_B) = 1177{,}5 - 116{,}1x_1 + 297{,}3x_2 - 355{,}8x_3 + 4723{,}5x_4 + 916{,}5x_5 + 204{,}8x_6 - 1237{,}2x_7 + 86{,}9x_8 + 146{,}5x_9 = 991{,}5$ Н/мм^2. При этом расчетное значение прочности $y^{\text{расч}} = 991{,}46$ Н/мм^2. Средняя ошибка аппроксимации составляет 0,001%. Следовательно, модель адекватна;

$y(\psi) = 109{,}7 - 84{,}8x_1 + 19{,}8x_2 - 7{,}8x_3 - 193{,}3x_4 + 45{,}8x_5 - 34{,}6x_6 + 101{,}9x_7 - 58{,}8x_8 - 437{,}4x_9 = 48{,}24\%$, среднее значение относительного сужения $y^{\text{расч}} = 48{,}236\%$. Здесь средняя ошибка аппроксимации составляет 0,001%. Модель адекватна;

$y(\delta) = 36{,}2 - 13{,}1x_1 - 20{,}5x_2 - 14{,}2x_3 - 162{,}9x_4 + 105{,}5x_5 + 41{,}8x_6 + 82{,}2x_7 - 80{,}1x_8 + 137{,}2x_9 = 16{,}68\%$, среднее значение относительного удлинения $y^{\text{расч}} = 16{,}679\%$. Средняя ошибка аппроксимации составляет 0,001%. Модель адекватна.

Таким образом, получены общие уравнения исследуемых характеристик. Задачей следующего этапа является нахождение специальных характеристик методом исключения незначимых из общего уравнения.

Шаг 11. Применяя метод исключения незначимых переменных из общего уравнения регрессии, можно выделить значимую и сконцентрировать внимание на ее управлении. Чтобы проверить переменные на этом шаге, необходимо определить частный F-критерий для каждой переменной $x_1, x_2, ..., x_k$ в предположении как будто данная переменная была включена в регрессионное уравнение последней.

В общем виде для фактора x_i частный F-критерий определится как [4]:

$$F_{xi} = \frac{R^2_{yx_1...x_i...x_k} - r^2_{yx_1...x_{i-1}x_{i+1}...x_k}}{1 - R^2_{yx_1...x_i...x_k}} \cdot \frac{n-k-1}{1},$$

где $R^2_{yx_1...x_i...x_k}$ – коэффициент множественной детерминации для модели с полным набором факторов; $R^2_{yx_1...x_{i-1}x_{i+1}...x_k}$ – тот же показатель, но без включения в модель фактора x_i; n – число наблюдений; k – число параметров в модели (без свободного члена).

Фактическое значение частного F-критерия сравнивается с табличным при уровне значимости α и числе степеней свободы: 1 и $n - k - 1$. Если фактическое значение F_{xi} превышает $F_{\text{табл}}(\alpha, k_1, k_2)$, то дополнительное включение фактора x_i в модель статистически оправдано и коэффициент чистой регрессии b_i при факторе x_i статистически значим. Если же фактическое значение F_{xi} меньше табличного, то дополнительное включение в модель фактора x_i не увеличивает существенно долю объясненной вариации признака y, следовательно, нецелесообразно его включение в модель; коэффициент регрессии при данном факторе в этом случае статистически незначим.

1. Определим частные F-критерии для уравнения (1):

$$y_{1.1}(\sigma_B) = 1104{,}5 - 238{,}5x_1 + 345{,}2x_2 + 14{,}4x_4 + 477{,}3x_5 - 222{,}8x_6 - 698{,}8x_7 - 26{,}9x_8 + 130{,}6x_9. \quad (1.1)$$

В этом уравнении исключена переменная x_3(Mn), так как значение F_{x3} оказалось меньше других значений в этом уравнении и составило 3,2, но больше $F_{\text{табл}} = 1{,}43$. Следовательно, требуется дальнейший перерасчет уравнения.

$$y_{1.2}(\sigma_B) = 1180{,}7 - 387{,}3x_1 + 393{,}9x_2 + 1044{,}1x_5 - 310{,}8x_6 + 184{,}8x_7 - 330{,}4x_8 + 214{,}2x_9. \quad (1.2)$$

Здесь исключена переменная x_4(P), так как значение F_{x4} оказалось меньше других значений в этом уравнении и составило 1,8, но больше $F_{\text{табл}} = 1{,}43$.

$$y_{1.3}(\sigma_B) = 1167{,}8 - 256{,}2x_1 + 1738{,}9x_5 - 448{,}1x_6 - 92{,}5x_7 - 212{,}5x_8 + 815{,}3x_9. \quad (1.3)$$

Здесь исключена переменная x_2(Si), так как значение F_{x2} оказалось меньше других значений в этом уравнении и составило 1,5, но больше $F_{\text{табл}} = 1{,}43$. Следовательно, процедура исключения переменной из уравнения (1.3) продолжается.

$$y_{1.4}(\sigma_B) = 1161{,}7 - 230{,}1x_1 - 582{,}1x_6 + 137{,}5x_7 - 217{,}4x_8 + 1741{,}2x_9. \quad (1.4)$$

Исключение переменной x_5(S) привело к конечному решению системы уравнений, так как при перерасчете уравнения частный F-критерий переменной x_6(Cr) = 0,6, подлежащей исключению, стал меньше $F_{табл}$ = 1,43.

Отсюда следует, что специальными характеристиками, влияющими на прочность стали марки 65К, являются x_1(C), x_6(Cr), x_7(Ni), x_8(Cu), x_9(N_2).

2. Определим частные F-критерии для уравнения (2):

$$y_{2.1}(\psi) = 55,9 + 6,1x_2 - 12,2x_3 - 113,6x_4 + 41,5x_5 - 20,9x_6 + 46,4x_7 - 33,9x_8 - 352,2x_9. \quad (2.1)$$

В этом уравнении исключена переменная x_1(C), так как значение F_{x1} оказалось меньше других значений в этом уравнении и составило 3,4, но больше $F_{табл}$ = 1,43. Следовательно, требуется дальнейший перерасчет уравнения.

$$y_{2.2}(\psi) = 49,7 + 7x_2 - 94,5x_4 + 25,9x_5 - 35,1x_6 + 62,3x_7 - 36,7x_8 - 348,5x_9. \quad (2.2)$$

Исключение переменной x_3(Mn) привело к конечному решению системы уравнений, так как при перерасчете уравнения частный F-критерий переменной x_9(N_2) = 0,9, подлежащей исключению, стал меньше $F_{табл}$ = 1,43.

Следовательно, дальнейший перерасчет уравнения не требуется, все оставшиеся переменные можно отнести к специальным характеристикам, влияющим на относительное сужение стали марки 65К.

3. Определим частные F-критерии для уравнения (3):

$$y_{3.1}(\delta) = 27,6 + 1,8x_1 - 28,5x_2 - 16,6x_3 - 61,9x_4 + 111,3x_5 + 53,7x_6 - 35,9x_7 + 150,2x_9. \quad (3.1)$$

Переменная x_8(Cu), так как значение F_{x8} = 10,9 оказалось меньше других значений в этом уравнении, но больше $F_{табл}$ = 1,43.

$$y_{3.2}(\delta) = 27,4 - 7,1x_1 - 15,6x_3 - 68,9x_4 + 61,4x_5 + 63,4x_6 - 24,9x_7 + 107,7x_9. \quad (3.2)$$

Здесь переменная x_2(Si), так как значение F_{x2} = 5,6 оказалось меньше других значений в этом уравнении, но больше $F_{табл}$ = 1,43.

$$y_{3..3}(\delta) = 28,3 - 11,1x_1 - 11,1x_3 - 71,8x_4 + 24,9x_5 + 25,8x_7 + 174,6x_9. \quad (3.3)$$

Переменная x_6(Cr), так как значение F_{x6} = 3,9 оказалось меньше других значений в этом уравнении, но больше $F_{табл}$ = 1,43.

$$y_{3.4}(\delta) = 24,2 - 13,3x_1 - 50,8x_4 + 20,6x_5 + 26,4x_7 + 161,1x_9. \quad (3.4)$$

Исключение переменной x_3(Mn) = 1,1 в уравнении (3.3) привело к конечному решению системы уравнений, так как при перерасчете уравнения (3.4) частный F-критерий переменной x_9(N_2) = 0,7, подлежащей исключению, стал меньше $F_{табл}$ = 1,43.

Следовательно, дальнейший перерасчет уравнения не требуется, все оставшиеся переменные можно отнести к специальным характеристикам, влияющим на относительное сужение стали марки 65К.

Таким образом, по результатам проведенного регрессионного анализа можно судить о значимости тех или иных компонентов в общей совокупности рассматриваемой модели. Разработанная и проверенная математическая модель, в основу которой положен метод Гаусса (наиболее удобный способ решения систем линейных уравнений), позволяет решить задачу выбора специальных характеристик продукции и параметров процесса, влияющих на качество конечного продукта. Преимуществом модели является ее научная основа и универсальность, а также возможность проведения проверки адекватности модели, что снижает степень ошибочности принятия решения. Кроме того, модель множественного линейного регрессионного анализа подразумевает поиск показателей (обозначаемых x), определяющих значение отдельной количественной переменной, обозначаемой y. Систематизация проведения подобного рода анализа позволяет максимально точно устанавливать зависимости между наблюдаемыми характеристиками и регулировать их с помощью рассчитанной математической модели. В случае автоматизации алгоритма проведения многомерного регрессионного анализа затраты времени на поиск решения задачи сокращаются в несколько раз. При этом не потребуется специальное обучение персонала.

В рассмотренном регрессионном анализе проведено исследование влияния химического состава катанки диаметром 5,5 мм стали марки 65К (C, Si, Mn, P, S, Cr, Ni, Cu, N_2) на механические свойства (прочность (σ_b), относительное сужение (ψ), относительное удлинение (δ)). По результатам определены специальные характеристики (C, Si, Mn, P, S, Cr, Ni, Cu) для данной марки стали, оказывающие влияние на механические свойства. Так, в случае определения прочности (σ_b) как специальной характеристики следует выделять следующие специальные параметры C, Cr, Ni, Cu, N_2: относительного сужения (ψ) – Si, Cr, Ni, Cu, N_2, S, P; относительного удлинения (δ) – C, Ni, N_2, S, P. Сила связи определена с помощью множественного коэффициента корреляции. Полученные значения показывают умеренную корреляцию между химическим составом, прочностью и относительным удлинением и слабую степень корреляции с относительным сужением (табл. 6).

Суммарную меру общего качества уравнения регрессии (его соответствия статистическим данным) показывает коэффициент детерминации. Как показывают расчеты, приведенные в табл. 6, линейная связь между химическим составом и пределом прочности, относительным удлинением является слабой, а с относительным сужением практически отсутствует. Поэтому следует провести исследования влияния других факторов на зависимую переменную *y*.

Литература

1. К у м э Х. Статистические методы повышения качества / Пер. с англ. М.: Финансы и статистика, 1990.
2. Ш и ш к и н И. В., С т а н я к и н В. М. Квалиметрия и управление качеством. М.: Изд-во ВЗПИ, 1992.
3. А д л е р Ю. П., М а р к о в а Е. В., Г р а н о в с к и й Ю. В. Планирование эксперимента при поиске оптимальных условий. Изд. 2-е. М.: Наука, 1976.
4. Д р е й п е р Н., С м и т Г. Прикладной регрессионный анализ. В 2-х кн. Кн. 2. пер. с англ., 2-е изд., перераб. и доп. М.: Финансы и статистика, 1987.
5. Н е у й м и н Я. Г. Модели в науке и технике. История, теория, практика. Л.: Наука, 1984.
6. Ч и ч к о А. Н., С о б о л е в В. Ф., Ч и ч к о О. И. Статистические методы регулирования качества продукции в литейном производстве. Мн.: БНТУ, 2006.
7. Рекомендации. Прикладная статистика. Методы обработки данных. Основные требования и характеристики. М.: ВНИИС, 1987.
8. Теория статистики / Под ред. Р. А. Шмойловой. М.: Финансы и статистика, 1998.
9. И в а ш о в А. Линейная алгебра. Матрицы: Учеб. пособ. М.: ВНИИС, 2004.
10. О р л о в А. И. Эконометрика: Учеб. пособ. М.: Изд-во «Экзамен», 2002.
11. Е ф и м о в В. В. Улучшение качества проектов и процессов: Учеб. пособ. Ульяновск: УлГТУ, 2004.
12. Е ф и м о в В. В. Статистические методы в управлении качеством: Учеб. пособ. Ульяновск: УлГТУ, 2003.