

Исследование причин снижения работоспособности насосных штанг

Саттаров И.Р., Ризванов Р.Р., Хакимов Т.А.

Ввиду многочисленности фонда скважин на месторождениях находящихся на поздней стадии разработки, оборудованных штанговыми установками, вопросы повышения эффективности их эксплуатации приобретают первостепенное значение с точки зрения технико-экономических показателей. Как правило, работа этого фонда отличается низкими показателями межремонтного периода (МРП). Этому способствуют сложные геолого-технические и физические условия, сверхнормативные пространственные параметры стволов, технологические параметры форсированной откачки и т. д. В данной статье проведены статистические исследования накопленного опыта и получено уравнение регрессии для прогнозирования наработки на отказ штанговых колонн и принятия решений, направленных на повышение работоспособности штанговых установок в целом.

Структура исходных данных по скважинам Контузлинского месторождения ОАО «Татнефть» за 7 лет их работы представлена в таблице (Табл. 1). Каждая запись которой характеризует обрыв штанговой колонны для одной скважины и технические параметры этой скважины, включая информацию о колоннах НКТ и штанг. Последняя заключена в колонках 9, 9.1-9.3.

Табл. 1 Колонки исходной таблицы

n/n	Название колонки	Ед. изм.	Обозначение	
1	№ скважины	б/р	Нскв	X1
2	Название месторождения	б/р	Нмст	X2
3	Марка насоса	б/р	Тнас	X3
4	Способ эксплуатации	б/р	Тэкс	X4
5	Глубина подвеса насоса	м	Нпод	X5
6	Частота качаний	кач/мин	n	X6
7	Длина хода	М	S	X7
8	Диаметр насоса	мм	Днас	X8
9	Штанги или НКТ	б/р	Тобр	X9
9.1	№ ступени	б/р	Нстп	X10
9.2	Диаметр ступени	мм	Дстп	X11

9.3	Длина ступени	м	Лстп	X12
10	Дебит жидкости	м ³ /сут	Qждк	X13
11	Дебит теоретический	м ³ /сут	Qтеор	X14
12	Доля жидкости (обводненность)	в долях	В	X15
13	Вязкость	сСт (мм ² /с)	ν	X16
14	Газовый фактор	м ³ /т	Г	X17
15	Дебит нефти	т/сут	Qнеф	X18
16	Дебит воды	м ³ /сут	Qвод	X19
17	Максимальный угол наклона ствола скважины	рад	α	X20
18	Глубина расположения максимального угла наклона	м	На	X21
19	Коэффициент подачи насоса	б/р	η	X22
20	Наработка	сут	T	T

В общем шаблоне Табл. 2 **Схематичное представление исходной таблицы** включены подтаблицы характеризующие каждую ступень колонны штанг и НКТ (пункты 9, 9.1, 9.2, 9.3).

Табл. 2 Схематичное представление исходной таблицы

Нскв	Тобр	Нстп	Дстп	Лстп
1	НКТ	1	73	1200
	НКТ	2	73	900
	...			
	штанги	1	22	800
	штанги	2	19	400
	...			
2	НКТ	1	73	1150
	...			

Перед началом регрессионного анализа, необходима унификации данных, для этого желательно исключить подтаблицу из базовой, предварительно её обработав, и добавить колонки в Табл. 2, характеризующие колонны НКТ и насосных штанг.

Колонны штанг и НКТ характеризуются усредненными диаметрами их ступеней

следующим образом:

$$d_{cp} = \frac{\sum_{i=1}^m d_i L_i}{\sum_{i=1}^m L_i}, \quad (1)$$

где d_i – диаметр i колонны, L_i – длина i колонны, m – количество ступеней. Используя длину и диаметр каждой ступени, получаем разнообразие значений и добавляем следующие столбцы в модель данных. Результаты вычислений добавлены в Табл. 1

Продолжение таблицы 1

21	Диаметр НКТ (усредненный)	мм	Днкт	X23
22	Диаметр насосных штанг (усредненный)	мм	Дштанг	X24

Оценка пригодности определенного параметра для анализа может быть произведена методом разбиения её значений на квартили (квартиль — граница на шкале измеряемого свойства, отделяющая 25% значений случайной величины от выборки. Различаются 3 квартиля: Q1 – первый на границе интервалов 0-25% и 25-50%, Q2 – второй на границе интервалов 25-50% и 50-75% называемый медианой и Q3 – третий на границе интервалов 50-75% и 75-100%). Обработка исходных данных этим методом показала, что технологические параметры имеют различные значения на границах и пригодны для анализа, за исключением значений усредненного диаметра НКТ (Табл. 3), которые во всех квартилях практически одинаковы, поэтому данный параметр исключается из дальнейшего рассмотрения [4].

Табл. 3 Разброс значений случайных величин среднего диаметра насосных труб и штанг

	min	1 квартиль	медиана	3 квартиль	max
НКТ	73	73	73	73	89

Дальнейший анализ проведен с использованием известных методов и статистического языка программирования «R», предназначенного для статистических обработок [5]. В результате анализа получена базовая регрессионная модель:

$$T = -167.1995 - 1.531624 X5 + 293.2946 X6 + 183.8179 X7 + 42.06031 X8 - \quad (2)$$

$$- 41.88825 X_{13} + 8.192688 X_{14} - 957.8543 X_{15} + 131.1325 X_{16} + 244.4943 X_{17} - 21.86063 X_{18} + 28.24314 X_{19} - 251.3793 X_{20} - 0.3154588 X_{21} - 363.4558 X_{22} - 379.91 X_{24}$$

Оценка уравнения (2) по известным статистическим критериям показывает, что коэффициент детерминации $R^2 = 0.59$. Коэффициент детерминации показывает, какая доля зависимой переменной обуславливается построенной функцией регрессии. Например, при полученном коэффициенте детерминации 0.59 регрессионная модель обуславливает 59% дисперсии, остальные же 41% считаются обусловленными факторами, не отраженными в модели. F-статистика (по Фишеру) = 5.009 с вероятностью нулевой гипотезы $6.88 \cdot 10^{-6}$. F-отношение это отношение рассчитанных с учётом степеней свободы объяснённой нашей моделью дисперсии к необъяснённой, нулевая гипотеза соответствует нулевому значению коэффициента детерминации, т.е. найденное уравнение регрессии статистически значимо. Стандартное отклонение невязки в полученной модели составляет 331 суток.

Для определения приемлемости такой ошибки при прогнозах наработки на отказ может быть использован метод известный под названием «ящик с усами», который позволяет графически представить разбивку на квартили значений наработки на отказ (Рис. 1). График «ящик с усами» (англ. box-and-whiskers plot) был впервые применен Джоном Тьюки, чтобы показать характер распределения набора данных [2]. График показывает распределение величины наработки на отказ, левая и правая граница прямоугольника («ящика») соответствует 1 и 3 квартилям соответственно, вертикальная жирная черта — медиану. Доверительный интервал показан пунктиром

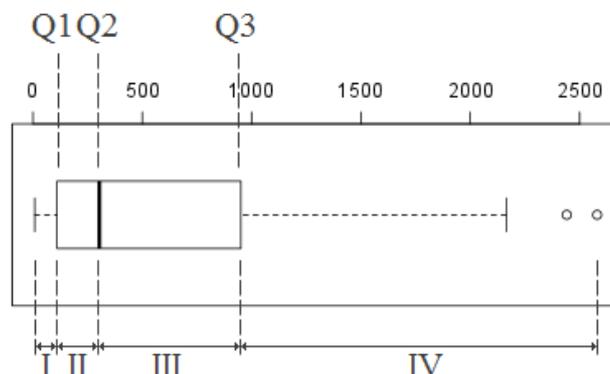


Рис. 1 Распределение наработки на отказ штанг («ящик с усами»)

Гистограмма, распределение и соответственное нормальное распределение остатков показаны на рисунке Рис. 2. Она характеризуется большим разбросом и неравномерностью (отклонением от нормального закона распределения). Как известно, распределение ошибок регрессионной модели должно соответствовать нормальному закону. Учитывая, что размер диапазона между 1 и 3 квартилем наработки штанговой колонны является примерно 900 суток, и что стандартное отклонение ошибки модели 331 суток, принимаем разброс ошибок широким.

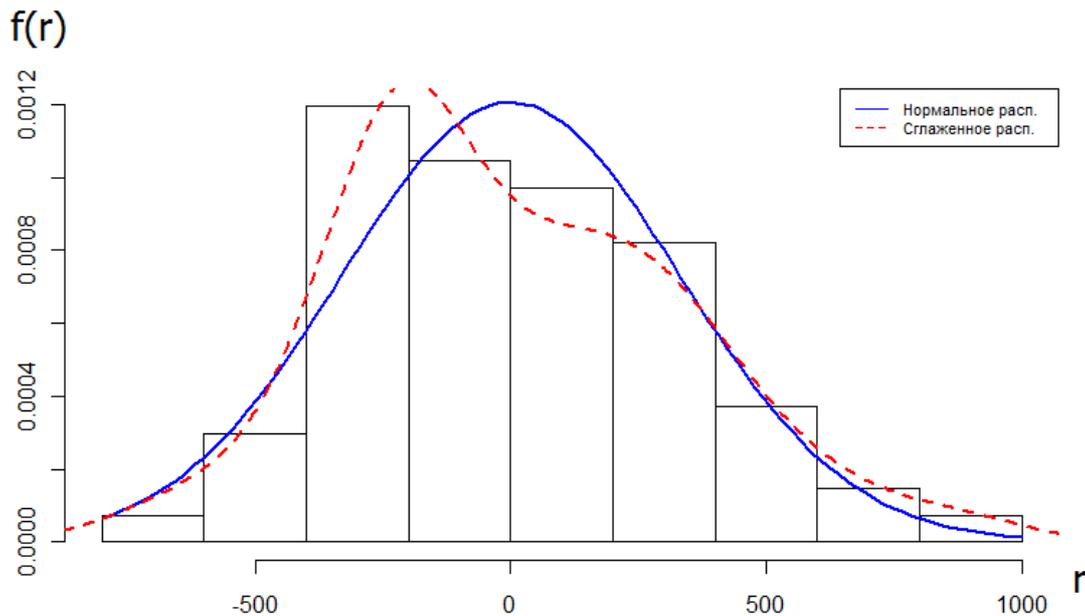


Рис. 2 Распределение остатков базовой регрессионной модели.

Проведенный анализ полученного уравнения (2) показал необходимость её совершенствования для получения модели, позволяющая повысить точность прогнозов.

Известно, если две или более переменные входящие в уравнение регрессии зависимы между собой, возникает неустойчивость параметров регрессии, осложняется анализ коэффициентов, и снижается качество получаемой модели. Данное состояние модели — означает её мультиколлинеарность. Модель обладающая таким состоянием, нуждается в сокращение размерности пространства входных данных удалением некоторых переменных. Одним из статистических методов выявления мультиколлинеарности является фактор роста дисперсии (VIF), для каждой независимой переменной:

$$\sqrt{\text{VIF}_i}; i = \overline{1, n} ,$$

где n – количество независимых переменных. Максимальное значение критерия принимают

наиболее зависимые переменные [1].

Приведем пример анализа мультиколлинеарности и её устранения. Максимальное значение $\sqrt{\text{VIF}}$ фактического дебита жидкости составляет 30,67 для линейной модели (2) (Рис. 3). Для оценки правомерности заключения по $\sqrt{\text{VIF}}$ построена проверочная модель:

$$X13 = -2.08 - 0.0009X5 - 0.07X6 - 0.06X7 + 0.005X8 + 0.003X14 + 0.35X15 - 0.03X16 - 0.05X17 + 1.12X18 + 0.85X19 - 0.24X20 - 0.0003X21 + 0.22X22 + 0.16X24 \quad (3)$$

Проверочная линейная модель (3) показывает, что имеется тесная линейная зависимость дебита жидкости от дебита нефти и воды (соответствующие p-value по критерию Стьюдента менее чем $2 \cdot 10^{-6}$), что подтверждает существующую взаимосвязь. Выявленная зависимость расположена в трехмерном пространстве на одной плоскости (Рис. 4). Остальные переменные влияют незначительно.

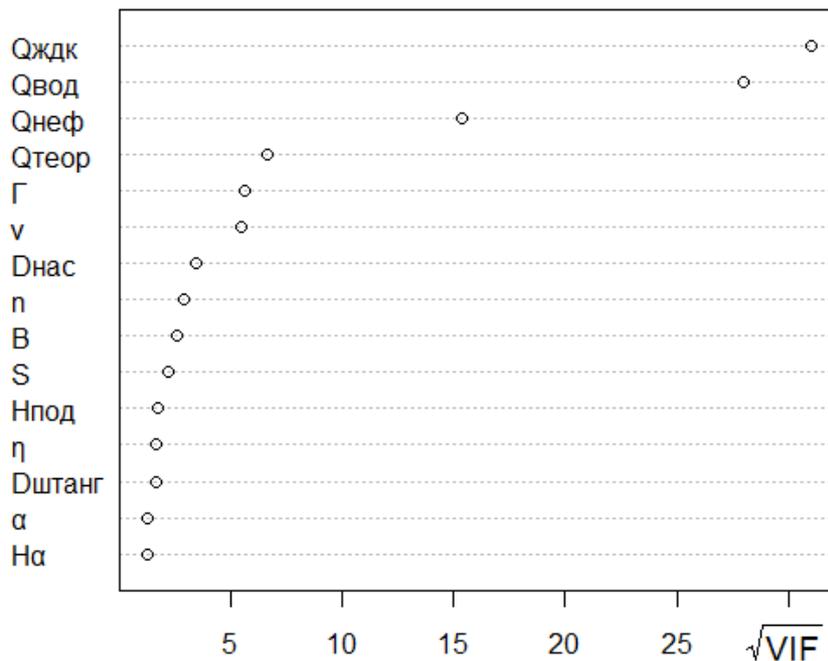


Рис. 3 Точечная диаграмма значений фактора роста дисперсии для входящих переменных.

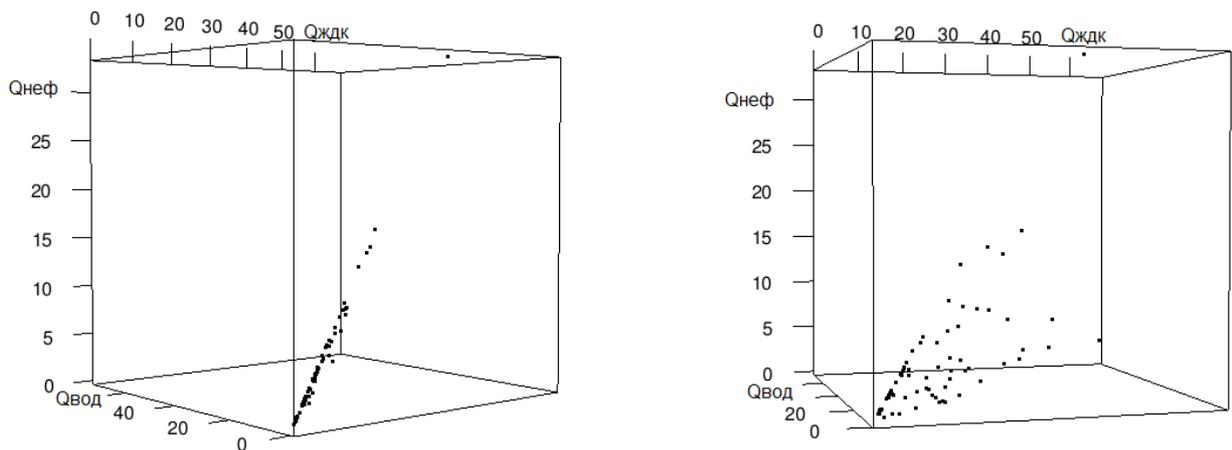


Рис. 4 Зависимость дебита жидкости от дебита нефти и воды.

Таким образом, доказано существование тесной зависимости между 3 переменными (дебит жидкости, нефти и воды) и устранение мультиколлинеарности достигается исключением одной из них, к примеру, переменной дебита жидкости как имеющей наибольший \sqrt{VIF} .

Вычисление \sqrt{VIF} после исключения дебита жидкости показывает, что максимальное значение \sqrt{VIF} уже для теоретического дебита составляет 5,78. Исследования проведенные, аналогично дебиту жидкости, построением уравнения регрессии, подтверждают наличие зависимости теоретического дебита от других переменных.

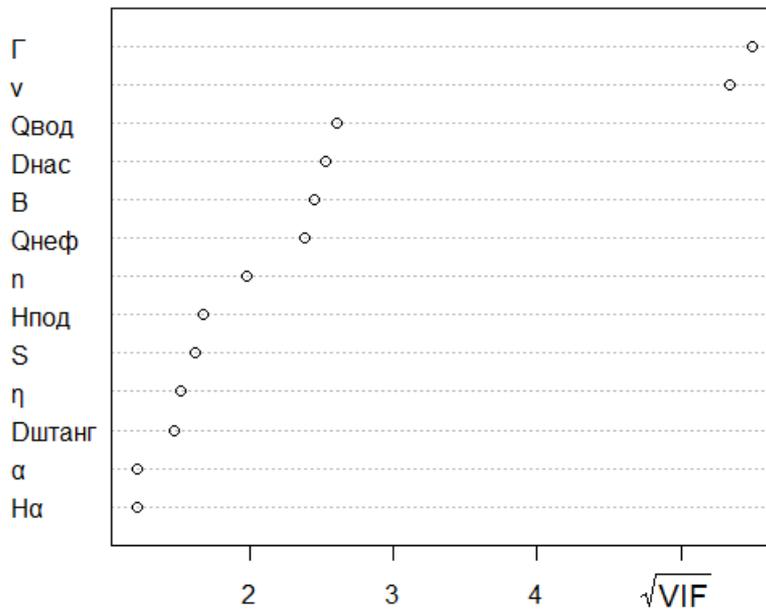


Рис. 5 Значения фактора роста дисперсии.

Из рисунка Рис. 5 видно, что после исключения из модели теоретического дебита, максимальное значение \sqrt{VIF} принимает газовый фактор равный 5,57, исследования которого, построением проверочной модели, показали, что значимый коэффициент получен лишь для одного параметра — вязкости добываемой жидкости (p-value менее $2 \cdot 10^{-16}$). Диаграмма разброса для газового фактора и вязкости опровергает наличие тесной линейной зависимости между этими параметрами, тем что скопление их происходит локально, только в 3 областях (Рис. 6). Следовательно, решение об устойчивой зависимости переменных и удалении одной из них – отклоняется. Опровержение взаимосвязи переменных со значением $\sqrt{VIF} = 5,57$, позволяет избежать оценки \sqrt{VIF} для остальных параметров, имеющих существенный отрыв по величине \sqrt{VIF} (Рис. 5)

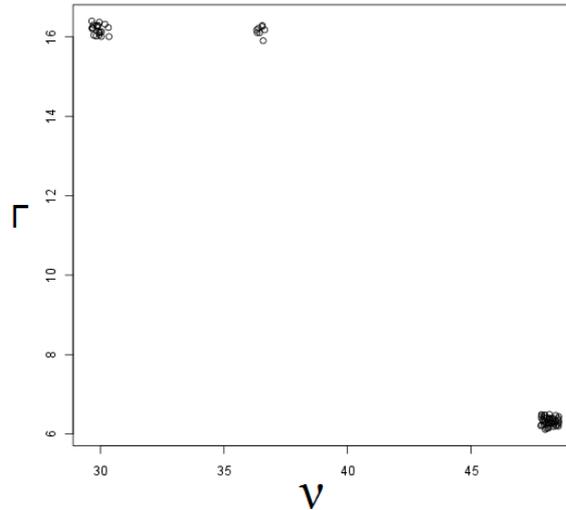


Рис. 6 Диаграмма разброса для вязкости и газового фактора

¹Линейная модель регрессии является линейной комбинацией входящих переменных и свободного члена. На основе входящих переменных можно вычислить новые нелинейные комбинации переменных, которые могут улучшить модель, и результирующее уравнение может принимать вид (4).

$$T = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \beta_{n+1} f_1(X_1, \dots, X_n) + \dots + \beta_{n+k} f_k(X_1, \dots, X_n), \quad (4)$$

$$T = \beta_0 + \sum_{i=1}^{n+k} \beta_i X_i \quad (5)$$

где $X_i, i \in \overline{1, n}$; n – число основных входящих переменных; $f_j, j \in \overline{1, k}$; k – число дополненных входящих переменных образованных функциональной зависимостью; $\beta_0, \dots, \beta_{n+k}$ – веса линейной модели.

Для получения нелинейных комбинаций, введем допущение, что каждая анализируемая скважина является прямолинейной с единственной точкой перегиба и состоит из вертикального и наклонного участка (Рис. 7). Допущенную форму ствола скважины характеризуют исходные данные: максимальный угол наклона и глубина расположения этого угла. Такая модель ствола скважины с точки зрения влияния её на работоспособность

¹ Убрали дебит воды и дебит нефти, без объяснений.

штанговой колонны может быть охарактеризована отношением $\frac{\alpha_{\max}}{l_{\max}}$, который представляет собой упрощенный вариант «обобщенного параметра интенсивности искривления ствола скважины» [3]. Подтверждением этой гипотезы может служить то, что при одном и том же значении максимального угла наклона в зависимости от глубины его расположения, степень влияния на работоспособность штанговой колонны будет различной. Другими словами чем меньше глубина расположения максимального угла по стволу скважины, тем больше его негативное влияние на работоспособность колонны и соответственно отношение $\frac{\alpha_{\max}}{l_{\max}}$ будет принимать большее значение. Это согласуется физикой процесса, чем меньше глубина расположения максимального угла, тем больше натяжение колонны, вызванная весом нижележащих штанг и жидкости на плунжере насоса, следовательно больше прижимающая Эйлера сила трения, по этой же причине выше напряжения от изгиба штанг, которые в значительной степени определяют работоспособность штанговой колонны.

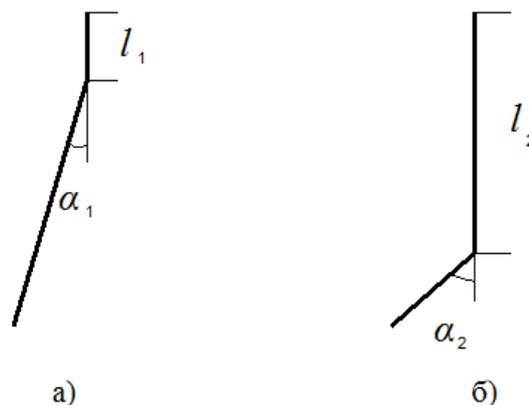


Рис. 7 Принятая форма скважины.

Таким образом, существует взаимосвязь величины максимального угла наклона ствола скважины и глубины его расположения. Для оценки предложенной гипотезы рассмотрим 2 эквивалентные скважины с одинаковыми условиями эксплуатации, типоразмерами колонн и глубиной подвески насоса (Рис. 7), при $l_2 > l_1$ и $\alpha_2 > \alpha_1$. Естественно, большая доля ожидаемых обрывов штанговой колонны приходится на место перегиба колонны. В варианте (а), перегиб приходится на верхнюю, наиболее нагруженную, часть колонны, смягчающим обстоятельством является малый угол наклона ствола, в варианте (б), перегиб имеет больший угол наклона, смягчающим обстоятельством является глубокое расположение этого перегиба, оставляя высоконагруженную верхнюю область колонны вертикальной. Логично предположить, что форма этих скважин одинаково влияет на величину наработки штанговой

колонны, и отношение максимального угла с глубиной его расположения, как нелинейная комбинация, может быть полезна для уточнения модели (X25).

Наряду с конфигурацией ствола скважины существенное влияние на напряжение в штанговой оказывает ускорение полированного штока. Формула (6) позволяет получить амплитуду пропорциональную Sn^2 , что является нелинейной комбинацией и которая может быть добавлена в регрессионную модель (X26).

$$x = \frac{S}{2} \cos(2\pi nt)$$

$$a = \frac{d^2(\frac{S}{2} \cos 2\pi nt)}{dt^2} = -2\pi^2 Sn^2 \cos 2\pi nt \quad (6)$$

На основании допущенной формы ствола скважины, рассчитаны составляющие веса штанг: вес вертикальной части штанговой колонны, вес штанг расположенных на наклонной части ствола и сила трения штанг о трубы в наклонной части (X27, X28, X29). Каждое слагаемое является нелинейной комбинацией и включается в модель. Аналогично рассчитывается вес жидкости на плунжер (X30), при допущении заполнения жидкости в насосной трубе до забоя скважины. Далее вычисляется коэффициент асимметрии цикла нагружения [источник].

$$r = \frac{P_{ум}^1}{P_{ум}^2 + P_{жс}}, \quad (7)$$

где $P_{ум}^1$ – вес штанг без силы трения, $P_{ум}^2$ – вес штанг с силой трения, $P_{жс}$ – вес столба жидкости в скважине, аналогично для включения в модель.

Исключив дублированные переменные в чистом виде, являющиеся по критерию Стьюдента, незначимыми: длина хода, число качаний в минуту, вес вертикальной составляющей колонны, вес наклонной составляющей колонны, максимальный угол искривления скважины, глубина расположения максимального угла, коэффициент асимметрии цикла и включив их комбинации получаем линейную модель (8):

$$T = - 3738.771 + 47.99561 X8 + 603.3145 X15 + 117.1691 X16 + 234.2125 X17 - 802.2394 X22 - 143.9662 X24 + 10.26169 (X5-X21) + 5.911904 X26 - 3134.857 X29 - 0.005792196 X30 - 622608 X25 \quad (8)$$

Диаграмма остатков приведена на Рис. 8, доверительные интервалы коэффициентов с вероятностью 0,95 показаны в Табл. 4.

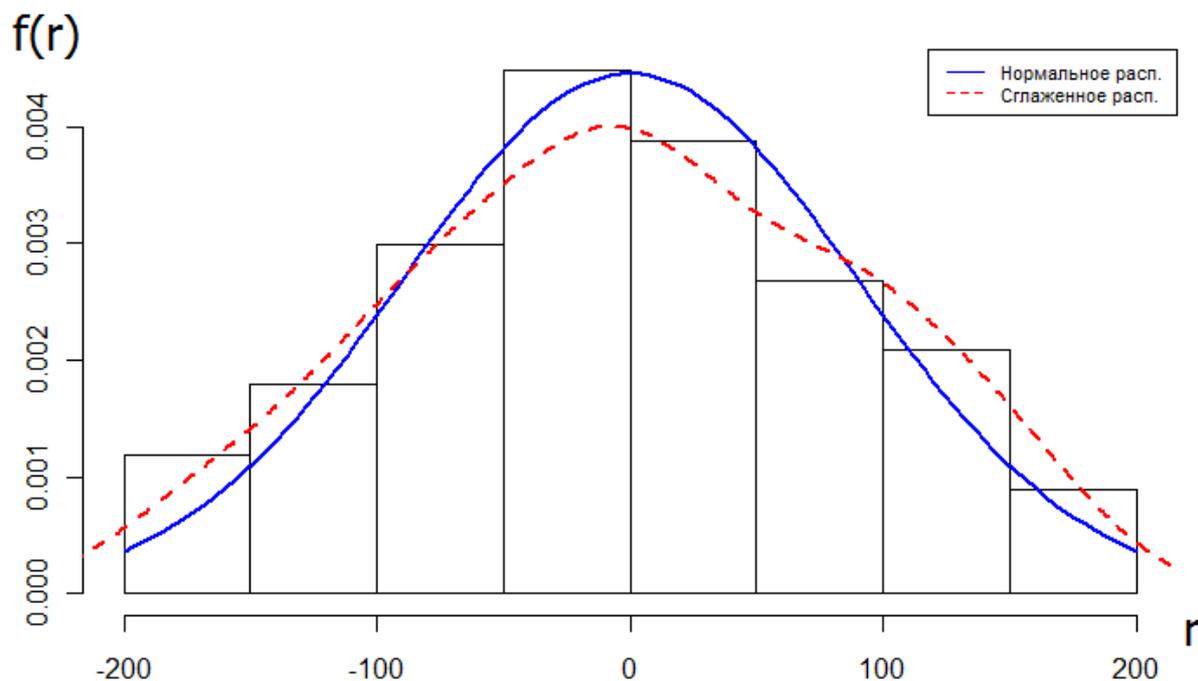


Рис. 8 Диаграмма остатков модели.

Табл. 4 95% - доверительные интервалы коэффициентов

Переменная	2.5%	97.5%
X8	43.97384	52.01738
X15	491.0044	715.6246
X16	102.1377	132.2004
X17	208.9188	259.5062
X22	-957.2574	-647.2214
X24	-185.3708	-102.5616
X5-X21	9.012028	11.51136
X26	4.647735	7.176073
X29	-3448.34	-2821.374
X30	-0.006333089	-0.005251303
X25	-717210.8	-528005.2

Анализ полученного уравнения регрессии показывает, что коэффициент детерминации $R^2 = 0.95$; скорректированный коэффициент детерминации $\text{adjusted-}R^2 = 0.95$; F-статистика (по Фишеру) = 119.2 с соответствующим уровнем вероятности $2.2 \cdot 10^{-16}$; Получена линейная модель с повышенным коэффициентом детерминации; сравнение скорректированных коэффициентов детерминации показывает более лучший показатель конечной модели по сравнению с начальной среднеквадратичным отклонением ошибки 89. Полученное уравнение регрессии может быть использовано для прогнозирования периода безотказной работы насосных штанг.

Литература

1. Hair JF, Anderson R, Tatham RL, Black WC: *Multivariate Data Analysis*. Prentice Hall: Upper Saddle River, N.J. 2006.
2. John W. Tukey.: *Exploratory Data Analysis*. Addison-Wesley. 1977.
3. Уразаков К.Р.: *Разработка научных основ и технологии добычи нефти из наклонно направленных скважин*. Диссертация на соискание ученой степени доктора технических наук в форме научного доклада. Уфа – 1994, 50 с.
4. *Вероятность и математическая статистика: Энциклопедия* / Под ред. Ю.В.Прохорова. — М.: Большая российская энциклопедия, 2003. — 912 с.
5. Интернет ресурс: <http://www.r-project.org/>