

Министерство образования Российской Федерации  
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет информатики

Кафедра теоретических основ информатики

УДК 681.03

ДОПУСТИТЬ К ЗАЩИТЕ В ГАК

Зав. кафедрой, д.т.н.

\_\_\_\_\_ Ю.Л. Костюк

«\_\_» \_\_\_\_\_ 2003 г.

Кузеванов Александр Викторович

**РАЗРАБОТКА И ПРОГРАММНАЯ РЕАЛИЗАЦИЯ АЛГОРИТМОВ  
ОЦЕНКИ ВЗАИМОДЕЙСТВИЯ СКВАЖИН НЕФТЯНЫХ  
МЕСТОРОЖДЕНИЙ**

Дипломная работа

Научный руководитель,  
менеджер ТомскНИПИнефть \_\_\_\_\_ В.Л. Сергеев  
ВНК, Д.т.н

Исполнитель,  
студент гр. 1481 \_\_\_\_\_ К.А. Кузеванов

Электронная версия дипломной работы помещена  
в электронную библиотеку. Файл

Администратор

Томск – 2003

## Реферат

Дипломная работа 42 с., 17 рис., 5 табл., 5 источников, 3 прил.

НЕПАРАМЕТРИЧЕСКАЯ ОЦЕНКА РЕГРЕССИИ, РАНГОВЫЙ КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ, ФУНКЦИОНАЛ КАЧЕСТВА, ДЕБИТ НЕФТЯНЫХ СКВАЖЕН, ПРИЕМИСОСТЬ НАГНЕТАТЕЛЬНЫХ СКВАЖИН, ПРОГНОЗ ДОБЫЧИ.

Объект исследования - скважины нефтяного месторождения, модели и алгоритмы взаимодействия скважин в условиях их нормальной эксплуатации (без остановки).

Цель работы - разработка и программная реализация алгоритмов взаимодействия нефтяных скважин и водо-нагнетательных скважин.

Методы исследований - теоретический и экспериментальный, на ЭВМ.

Результаты - разработан и реализован программный модуль для оценки взаимодействия нефтяных и водо-нагнетательных скважин.

# Содержание

Реферат .....	2
Содержание .....	3
Введение .....	5
1 Алгоритм оценки взаимодействия скважин на основе непараметрического метода аналогов .....	8
1.1 Оценка линейной регрессии .....	8
1.2 Алгоритмы на основе модифицированных оценок линейной регрессии .....	11
1.2.1 Модификации оценки выборочного среднего .....	11
1.2.2 Модификация оценки дисперсии .....	12
1.2.3 Модифицированная оценка линейной регрессии .....	13
1.3 Уточнение базовых оценок линейной регрессии .....	15
1.4 Алгоритм оценки взаимодействия скважин на основе ранговых коэффициентов корреляции .....	17
1.4.2 Алгоритм оценки взаимодействия скважин на основе рангового коэффициента корреляции Кэндела .....	18
1.5 Определение параметров алгоритмов построения оценок коэффициентов линейной регрессии .....	20
2 Аprobация алгоритмов взаимодействия скважин на модельных и реальных данных. ....	22
2.1 Построение оценок на модельных данных. ....	22
2.2 Аprobация алгоритмов на реальных данных .....	27
3 Программная реализация .....	34
3.1 Блок моделирования данных .....	34
3.1.1 Моделирования процессов взаимодействия скважин .....	34
3.1.2 Хранение реальных данных. ....	35

3.2 Построение оценок -----	36
3.3 Блок вывода результатов. -----	36
Заключение-----	38
Список использованных источников -----	39
Приложение А. КАРТА РАЗРАБОТКИ ОБЪЕКТА Ю <sub>22</sub> СЕВЕРО-ВАХСКОЙ ПЛОЩАДИ ПО СОСТОЯНИЮ НА 1.04.01 г-----	40
Приложение Б. Руководство программиста. -----	41
Приложение В. Руководство пользователя. -----	46

## Введение

В нефтегазодобычи одной из задач является исследования взаимодействия скважин между собой. Группа скважин на одном участке разработки представляет собой, по сути, сложную систему сообщающихся сосудов. Измерения различных характеристик скважин производится с ошибками и подвержены значительным случайным колебаниям, связанными с влиянием соседних скважин. Вообще, можно говорить, что взаимодействие скважин представляет собой некоторую функциональную стохастическую зависимость, которую и необходимо установить.

Одним из распространенных методов исследования взаимодействия скважин в практике нефтегазодобычи является гидропрослушивание [4]. Метод предусматривает создание гидродинамического импульса в скважине, именуемой возмущающей, посредством ее пуска, остановки или резкого изменения режима работы и регистрации этого импульса по изменению давления в другой скважине, именуемой реагирующей(наблюдательной), с последующей интерпретацией полученных данных. В качестве наблюдательных обычно используются несколько скважин, что дает возможность осветить гидропрослушиванием значительную по размеру зону пласта.

Однако использование метода гидропрослушивания для определения взаимодействия скважин в реальной промышленной обстановке связано с остановкой скважин и значительной потерей добычи нефти. В этой связи возникает задача оценки взаимодействия скважин в условиях их нормальной эксплуатации (без остановки).

В данной работе ставится задача в выяснении функциональной зависимости между дебитами добывающих и приемистостью нагнетательных скважин, а также между дебитами добывающих скважин динамически в процессе их функционирования.

Среднесуточный дебит добывающей скважины определяется как отношение суммарной добытой за месяц нефти(жидкости) к числу отработанных часов, умноженное на 24. Среднесуточная приемистость нагнетательных скважин определяется аналогичным образом как отношение суммарной закачанной за месяц воды в нагнетательную скважину к числу отработанных часов, умноженное на 24.

Дебиты скважин и участков объектов разработки представляют некоторые

нестационарные случайные процессы. Дебиты скважин измеряются с ошибками и подвержены значительным случайным колебаниям, связанным с влиянием соседних скважин, остановкой скважин, проведением различных мероприятий на скважинах окружения. Наблюдается и наложение данных факторов.

Определение закономерности традиционными статистическими методами[2] затруднительно. Например, если брать большой отрезок времени, например два года, закономерности размывается в связи с наложением многочисленных факторов. При отрезках времени год и менее закономерности проявляются, но наблюдается неустойчивость оценок, например, зависимость между дебитами добывающих и приемистостью нагнетательных скважин обратно пропорциональная, что противоречит физике процесса.

Должна быть организована специальная фильтрация данных. Например, брать оптимальные достаточно короткие временные промежутки для построения моделей зависимости между скважинами и осуществлять стабилизацию решений на этих промежутках, например, за счет использования дополнительной информации.

В данной работе рассматриваются три метода устойчивой оценки взаимодействия скважин. В основе первого метода используется модифицированная оценка линейной регрессии с учетом априорной информации [1-2]. В качестве априорных данных используются значения оценок параметров кусочно-линейной, взвешенной модели полученных за предыдущие моменты времени. В основе второго метода используются оценки рангового коэффициента корреляции К. Спирмена [3]. Третий метод связан с использованием рангового коэффициента корреляции Кендэла [3].

Исходная оценка взаимодействия скважин строится как комбинация оценок, полученных на основе приведенных выше методов.

Целью дипломной работы является:

Разработка и программная реализация алгоритмов взаимодействия скважин с использованием непараметрических методов обработки информации, статистик Спирмена и Кендэла.

Исследование точности алгоритмов взаимодействия скважин на модельных данных, полученных методом статистического моделирования.

Апробация алгоритмов на реальных данных с целью определения эффективности

работы нагнетательных скважин в условиях интенсификации добычи нефти (разделение эффектов от закачки жидкости и заглублении насосов).

# 1 Алгоритм оценки взаимодействия скважин на основе непараметрического метода аналогов

## 1.1 Оценка линейной регрессии

Пусть имеется добывающая скважина, где  $y(t)$  - дебит этой скважины, и есть  $m$  скважин окружения, где  $x_1(t), x_2(t), \dots, x_m(t)$  - дебиты скважин окружения. Основываясь на функциональной стохастической зависимости:

$$y(t) = f(x_1(t), x_2(t), \dots, x_m(t)), j = \overline{1, m} \quad (1)$$

между дебитом скважины  $y(t)$  и дебитами скважин окружения  $x_1(t), x_2(t), \dots, x_m(t)$ , рассмотрим задачу моделирования с использованием вероятностно-статистических и непараметрических методов на основе промысловых данных  $y(t_i), x_1(t_i), x_2(t_i), \dots, x_m(t_i), i = \overline{1, n}$ , полученных за период эксплуатации скважин. Данный подход позволяет моделировать функциональную зависимость между дебитами скважин в условиях априорной неопределенности о структуре модели, располагая лишь достаточно общими сведениями о функции взаимодействия скважин  $f(x_1(t), x_2(t), \dots, x_m(t))$  такими как ее ограниченность, однозначность, существование производных и т.п.

Известно, что оптимальным приближением к функции  $f(x_1(t), x_2(t), \dots, x_m(t))$  является условное математическое ожидание (регрессия)

$$r(X(t)) = \int_Y y F(y(t) / X(t)) dy \quad (2)$$

где  $F(y(t) / X(t))$  - условная плотность вероятности  $y(t)$  при заданных значениях вектора  $X(t) = x_1(t), x_2(t), \dots, x_m(t)$ .

Суть непараметрического подхода к моделированию функциональной зависимости (1) заключается в оценке условного математического ожидания с использованием



непараметрической оценки условной плотности вероятности. Доказано, что непараметрическая оценка при увеличении объема обучающей выборки  $n$  сходится к функции взаимодействия скважин  $f(X(t))$  (1.1)

$$r_n(y(t)/X(t)) \rightarrow f(X(t)), n \rightarrow \infty$$

при условии стационарности процесса изменения дебитов скважин во времени.

В том случае, когда восстанавливаемая нами функция регрессии  $r(X(t))$  имеет следующий вид:

$$r(X(t)) = a + b_1x_1 + b_2x_2 + \dots + b_mx_m \quad (3)$$

её называют функцией линейной регрессии [4].

Рассмотрим случай, когда в формуле (3), мы наблюдаем линейную зависимость двух переменных.

$$y(t_i) = a + bx(t_i), i = \overline{1, n} \quad (4)$$

где  $y(t_i), x(t_i)$  - значение дебитов скважин измеренные в момент времени  $t_i$ , в дальнейшем мы будем записывать эти обозначения в сокращенном варианте  $y_i, x_i$ , для упрощения записи формул.  $n$  - общее количество моментов времени, в которые производились наблюдения.

Для прогнозирования дебитов скважин нам необходимо восстановить зависимость (4) в точке  $k = n$  (т.е. в момент времени  $t_n$ ). Для этого нам необходимо произвести оценку коэффициентов  $a, b$ , обозначим их через  $a_k, b_k$ , соответственно. В связи с характеристиками функциональных зависимостей между дебитами скважин, изложенными во введении, мы будем использовать не весь доступный объем выборки наблюдений  $n$ , а некоторую предысторию процесса (4), объемом  $m$  наблюдений. Таким образом, мы будем строить оценку линейной регрессии, используя  $n-m$  последних наблюдений. Восстанавливаемая зависимость выглядит следующим образом:

$$y_k = a_k + b_k x_k, k = n \quad (5),$$

где  $y_k, x_k$  - значения дебетов скважин в момент времени  $t_k$ ,  $a_k, b_k$  - значения коэффициентов линейной регрессии, оцененных в точке  $k = n$ , с использованием предыстории процессов  $y_{k-i}, x_{k-i}, i = \overline{1, m}$ ,  $m$  - объем выборки предыстории процессов.

Итак, запишем оценку коэффициентов линейной регрессии [4], исходя из наших обозначений.

$$a_k = m_k(y) - \rho_k \frac{\sqrt{Dy_k(y)}}{\sqrt{Dx_k(x)}} m_k(x), \quad b_k = \rho_k \frac{\sqrt{Dy_k(y)}}{\sqrt{Dx_k(x)}} \quad (6)$$

где  $\rho_k$  - коэффициент корреляции линейной регрессии (5),  $m_k(y), m_k(x)$  - выборочные средние значения, а  $Dx_k(x), Dy_k(y)$  - дисперсия случайных величин  $x_k, y_k$ , соответственно.

Коэффициент линейной регрессии вычисляется по формуле:

$$\rho_k = \frac{\sum_{j=0}^{m-1} (y_{k-j} - m_k(y))(x_{k-j} - m_k(x))}{m \sqrt{D_k(x)} \sqrt{D_k(y)}}, k = n \quad (7)$$

Для оценки дисперсии величин  $y_k, x_k$  (см. [4, стр. 209]):

$$D_k(y) = \frac{\sum_{i=0}^{m-1} (y_{k-i} - m_k(y))^2}{m-1} \quad Dx_k(x) = \frac{\sum_{i=0}^{m-1} (x_{k-i} - m_k(x))^2}{m-1} \quad (8)$$

Оценка выборочного среднего:

$$m_k(x) = \frac{\sum_{i=0}^{m-1} x_{k-i}}{m} \quad m_k(y) = \frac{\sum_{i=0}^{m-1} y_{k-i}}{m} \quad (9)$$

Для получения более точных оценок, необходимо увеличивать объем выборки предысторий, но на больших объемах выборки часто наблюдается размывание, а при малых

отрезках времени, а особенно при наличии статистических выбросов в зависимости, наблюдается не устойчивость оценки, так например, при наличии выбросов у нас будет сильно увеличиваться оценка дисперсии, что приведет и к увеличению доверительных интервалов. Поэтому мы будем использовать модификации оценки линейной регрессии. В этой связи рассмотрим алгоритмы на основе модифицированных оценок линейной регрессии.

## **1.2 Алгоритмы на основе модифицированных оценок линейной регрессии**

Рассмотрим модификации оценки линейной регрессии, для получения устойчивых оценок функциональной зависимости взаимодействия скважин.

### **1.2.1 Модификации оценки выборочного среднего**

Для борьбы с выбросами, прежде всего можно заменить оценки выборочного среднего (9) на устойчивые оценки медианы, т.е. в формулах (6) – (8) вместо выборочного среднего мы будем использовать:

$$m_k(x) = med(x_{k-i}, i = \overline{0, m-1}, k = n) \quad (10)$$

$$m_k(y) = med(y_{k-i}, i = \overline{0, m-1}, k = n) \quad (11)$$

В этом случае мы не будем учитывать, при построении оценки, те моменты времени в которых произошли статистические выбросы, которые при использовании оценок вида (9) могли значительно исказить оценки выборочного среднего.

Второй способ изменения оценок (9) состоит в том, что мы не будем отбрасывать точки, в которых наблюдаются значительные отклонения от среднего значения на оцениваемом интервале. Потому, что, вообще говоря, мы не можем сказать из-за чего произошли отклонения, и как это повлияло на функциональную зависимость взаимодействия скважин. И вполне логичным тогда является шаг, при котором мы учитываем все точки, но при этом приписываем каждой точке некоторый вес:

$$m_k(x) = \frac{\sum_{i=0}^{m-1} K_k(x_{k-i})x_{k-i}}{\sum_{j=0}^{m-1} K_k(x_{k-j})} \quad m_k(y) = \frac{\sum_{i=0}^{m-1} K_k(y_{k-j})y_{k-j}}{\sum_{j=0}^{m-1} K_k(y_{k-j})} \quad (12)$$

где  $K_k(x), K_k(y)$  - весовые функции, рассчитываемые по формулам:

$$K_k(x) = \exp(-(\frac{m_k(x) - x_{k-j}}{h})^2 / 2), \quad K_k(y) = \exp(-(\frac{m_k(y) - y_{k-j}}{h})^2 / 2) \quad (13)$$

здесь значения  $m_k(x), m_k(y)$  - оценки средних значения случайных величин  $y_k, x_k$  вычисленных на интервале  $\overline{k, k-m+1}$  (размер интервала равен объему выборки предисторий  $m$ ), при  $k = n$ , вычисленных либо по формулам (9), либо по формулам (10) и (11). Полученные значения используются при построении оценок коэффициентов линейной регрессии (1.4).

## 1.2.2 Модификация оценки дисперсии

Следующим логическим шагом, конечно же, является модификация оценок дисперсии значения дебитов скважин  $y_k, x_k$  при построении оценок (6), где  $k = n$ . Здесь, по аналогии с оценками средних значений вместо оценок (8), алгоритма построения оценки линейной регрессии (см. параграф 1.1), мы построим взвешенные оценки дисперсии вида:

$$Dx_k(x) = \frac{\sum_{j=0}^{m-1} K_k(x_{k-j})(x_{k-j} - m_k(x))^2}{(m-1) \sum_{j=0}^{m-1} K_k(x_{k-j})} \quad (14)$$

$$Dy_k(y) = \frac{\sum_{j=0}^{m-1} K_k(y_{k-j})(y_{k-j} - m_k(y))^2}{(m-1) \sum_{j=0}^{m-1} K_k(y_{k-j})} \quad (15)$$

Здесь мы так же несколько изменим весовые функции, что бы они несли некоторую

информацию о дисперсии оцениваемых величин:

$$K_k(x) = \exp\left(-\frac{m_k(x) - x_{k-j}}{h\sqrt{D_x^*(x_{k-j})}}\right)^2 / 2, \quad K_k(y) = \exp\left(-\frac{m_k(y) - y_{k-j}}{h\sqrt{D_y^*(y_{k-j})}}\right)^2 / 2, \quad (h \approx 1) \quad (16)$$

где  $D_y^*, D_x^*$  - устойчивые оценки параметра масштаба:

$$D_y^*(y_{k-j}) = \text{med}(|y_{k-j} - m_k(y)|, j = \overline{0, m-1}, k = \overline{n, m})^2, \quad (17)$$

$$D_x^*(x_{k-j}) = \text{med}(|x_{k-j} - m_k(x)|, j = \overline{0, m-1}, k = \overline{n, m})^2.$$

### 1.2.3 Модифицированная оценка линейной регрессии

Запишем теперь алгоритм модифицированной линейной регрессии, в котором используем все полученные ранее в параграфах (1.2.1) и (1.2.2) изменённые оценки. Также добавим весовые функции в оценку коэффициента корреляции. Таким образом, окончательно наш алгоритм оценки линейной регрессии (4) будет выглядеть следующим образом:

$$y_k = a_k + b_k x_k, k = n \quad (18)$$

где  $y_k, x_k$  - значения дебетов скважин в момент времени  $t_k$ ,  $a_k, b_k$  - значения коэффициентов линейной регрессии, оцененных в точке  $k = n$ , с использованием предыстории процессов  $y_{k-i}, x_{k-i}, i = \overline{1, m}$ ,  $m$  - объем выборки предыстории процессов. Расчет коэффициентов (6)

$$a_k = m_k(y) - \rho_k \frac{\sqrt{Dy_k(y)}}{\sqrt{Dx_k(x)}} m_k(x), \quad b_k = \rho_k \frac{\sqrt{Dy_k(y)}}{\sqrt{Dx_k(x)}}$$

где оценка коэффициента корреляции:

$$\rho_k = \frac{\sum_{j=0}^{m-1} K_k(x)K_k(y)(y_{k-j} - m_k(y))(x_{k-j} - m_k(x))}{\sum_{j=0}^{m-1} K_k(x)K_k(y)\sqrt{D_k(x)}\sqrt{D_k(y)}}, k = n \quad (19)$$

оценка дисперсии (14) – (15):

$$Dx_k(x) = \frac{\sum_{j=0}^{m-1} K_k(x_{k-j})(x_{k-j} - m_k(x))^2}{(m-1)\sum_{j=0}^{m-1} K_k(x_{k-j})}$$

$$Dy_k(y) = \frac{\sum_{j=0}^{m-1} K_k(y_{k-j})(y_{k-j} - m_k(y))^2}{(m-1)\sum_{j=0}^{m-1} K_k(y_{k-j})}$$

значение весовых функций рассчитываем по формулам:

$$K_k(x) = \exp\left(-\left(\frac{m_k(x) - x_{k-j}}{h\sqrt{D_x^*(x_{k-j})}}\right)^2 / 2\right), K_k(y) = \exp\left(-\left(\frac{m_k(y) - y_{k-j}}{h\sqrt{D_y^*(y_{k-j})}}\right)^2 / 2\right), (h \approx 1)$$

$$D_y^*(y_{k-j}) = \text{med}(|y_{k-j} - m_k(y)|, j = \overline{0, m-1}, k = n)^2,$$

$$D_x^*(x_{k-j}) = \text{med}(|x_{k-j} - m_k(x)|, j = \overline{0, m-1}, k = n)^2.$$

Оценки среднего значения, на первом этапе вычисляем средние значения по формулам (10), (11):

$$m_k(x) = \text{med}(x_{k-i}, i = \overline{0, m-1}, k = n) \quad m_k(y) = \text{med}(y_{k-i}, i = \overline{0, m-1}, k = n)$$

Потом рассчитываем взвешенные оценки выборочного среднего (12) и подставляем в формулы (6), (14), (15) и (19):

$$m_k(x) = \frac{\sum_{i=0}^{m-1} K_k(x_{k-i})x_{k-i}}{\sum_{j=0}^{m-1} K_k(x_{k-j})} \quad m_k(y) = \frac{\sum_{i=0}^{m-1} K_k(y_{k-j})y_{k-j}}{\sum_{j=0}^{m-1} K_k(y_{k-j})}$$

где  $K_k(x), K_k(y)$  - весовые функции, рассчитываемые по формулам:

$$K_k(x) = \exp\left(-\left(\frac{m_k(x) - x_{k-j}}{h}\right)^2 / 2\right), \quad K_k(y) = \exp\left(-\left(\frac{m_k(y) - y_{k-j}}{h}\right)^2 / 2\right) \quad \text{при } k = n$$

### 1.3 Уточнение базовых оценок линейной регрессии

В этом параграфе я уточню необходимость использования индекса  $k$ . И так, нас интересуют, прежде всего, оценки коэффициентов регрессии в точке, где были произведены последние по времени измерения значений дебитов скважин, мы пытаемся определить какая функциональная зависимость наблюдается в текущий момент между скважинами, так как в конечном итоге нам необходимо построить прогноз взаимодействия скважин. При этом мы используем лишь небольшую выборку предысторий. Но ведь, более ранние наблюдения тоже несут в себе некоторую информацию. Теперь попытаемся учитывать то, какой была функциональная зависимость в точке  $n-1$ ,  $n-2$  и т. д. То есть построим оценки аналоги, которые будут учитывать априорную информацию о нашей зависимости и построим взвешенные оценки базовых коэффициентов  $a_k, b_k$ , где  $k = n$ .

Пусть мы оцениваем коэффициенты линейной регрессии вида (4). Построим оценки аналоги  $a_k, b_k$ , где  $k = \overline{n, m1}$ , в моменты времени  $t_k = k$ ,  $(n-m1)$  размер обучающей выборки (количество оценок “аналогов”). Оценки “аналоги” построены по формулам (5), где  $m$  – объем выборки предысторий, а  $k = \overline{n, m1}$ . Мы получим  $(n-m1)$  оценок коэффициентов линейной регрессии на шагах от  $m1$  до  $n$ . Теперь построим базовые оценки коэффициентов  $\overline{b_n}, \overline{a_n}$  линейной регрессии в момент времени  $t_k = n$  по следующей схеме:

$$\bar{b}_n = \frac{\sum_{i=0}^{m_1} K\left(\frac{b_k^* - b_{k-i}^*}{\sqrt{D_k^*(b)} * h}\right) * b_{k-i}^*}{\sum_{i=0}^{m_1} K\left(\frac{b_k^{**} - b_{k-i}^*}{\sqrt{D_k^*(b)} h}\right)}, \quad \bar{a}_n = \frac{\sum_{i=0}^{m_1} K\left(\frac{a_k^{**} - a_{k-i}^*}{\sqrt{D_k^*(a)} * h}\right) * a_{k-i}^*}{\sum_{i=0}^{m_1} K\left(\frac{a_k^{**} - a_{k-i}^*}{\sqrt{D_k^*(a)} * h}\right)}, \quad \text{где } k = \overline{n, m_1} \quad (20)$$

$$D_k^*(a) = \text{med}(|a_{k-i}^* - a_k^{**}|, i = \overline{1, m_1}), \quad D_k^*(b) = \text{med}(|b_{k-i}^* - b_k^{**}|, i = \overline{1, m_1}) \quad (21)$$

$$K(u) = \exp(-u^2)$$

где  $b_k^* = \text{med}(b_{k-i}, i = \overline{1, m_1})$  (процедура необходима в случае если в момент времени  $t_k = k$  наблюдается существенные различия оценок  $b_{k-i}, i = \overline{1, m_1}$  среди которых могут быть и отрицательные, например  $b_k \leq 0, b_{k-1} \approx 1$ ). Оценка  $b_k^*$  при анализе добывающей и нагнетательной скважин из физических соображений не должна быть отрицательной, в противном случае необходимо взять  $b_k^* = 0$ .



## 1.4 Алгоритм оценки взаимодействия скважин на основе ранговых коэффициентов корреляции

В параграфах (1.1) – (1.3) мы рассматривали непараметрические оценки выборочного коэффициента корреляции, который измеряет степень связи между случайными величинами. Ограничиваться только таким подходом мы не можем. В рамках непараметрического подхода, можно построить оценки коэффициентов корреляции используя понятие рангов, Рассмотрим две таких оценки: рангового коэффициента корреляции К. Спирмена и рангового коэффициента корреляции Кендела.

Предположим, что у нас есть  $n$  пар наблюдений  $(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)$ , которые составляют случайную выборку из некоторого двумерного непрерывного распределения. Мы составим вариационные ряды для  $x_i$  и отдельно для  $y_i$ . Если две переменные сильно зависимы или имеют высокую степень связи, мы в праве ожидать, что ранги двух элементов каждой пары примерно одинаковы. И наоборот, если элементы не зависимы между собой, то будет наблюдаться различие рангов

Таким образом, в методе описанном в параграфе (1.2.3) вместо взвешенной оценки корреляции (19), мы будем использовать полученные ниже оценки ранговых коэффициентов корреляции (25) и (30) (см. ниже).

### 1.4.1 Алгоритм оценки взаимодействия скважин на основе рангового коэффициента корреляции К. Спирмена.

Коэффициент ранговой корреляции К. Спирмена является простой модификацией коэффициента корреляции Пирсона [5], при которой величины  $x_i, y_i$  заменяются их рангами. Здесь, я не буду писать классическую формулу расчета коэффициента корреляции, а запишу формулы сразу с учетом того, что мы будем проводить уточнение базовых оценок с учетом оценок “аналогов”, как в главе 1.3. И так:

$$r_k = 1 - \frac{6 \sum_{j=0}^m (R_j(x) - R_j(y))}{m^3 - m}, \text{ где } |r_k| \leq 1 \quad (22)$$

$$R_j(x) = \text{ранг}(x_j, j = \overline{k, k-m+1}, k = \overline{n, m}), \quad (23)$$

$$R_j(y) = \text{ранг}(y_j, j = \overline{k, k-m+1}, k = \overline{n, m}). \quad (24)$$

По аналогии с параграфом (1.2.3) проводим уточнение базовых оценок  $r_k, k = \overline{n, m_1}$  в момент времени  $t_k = k$  с учетом оценок “аналогов”  $r_{k-i}^*, i = \overline{1, m_1-1}$

$$\bar{r}_n = \frac{\sum_{i=0}^{m_1} K\left(\frac{r_k^* - r_{k-i}}{\sqrt{D_k(r_k)} * h}\right) * r_{k-i}}{\sum_{i=0}^{m_1} K\left(\frac{r_k^* - r_{k-i}}{\sqrt{D_k(r_k)} * h}\right)}, \quad (25)$$

где

$$r_k^* = \text{med}(r_{k-i}, i = \overline{0, m_1}), \quad (26)$$

а

$$D_k(r_k) = \text{med}(|r_{k-i}^* - r_k^{**}|, i = \overline{0, m_1}), \quad K(u) = \exp(-u^2) \quad (27)$$

#### 1.4.2 Алгоритм оценки взаимодействия скважин на основе рангового коэффициента корреляции Кэндела

Коэффициент ранговой корреляции Кэндела, является другой мерой корреляции, которая строиться следующим способом:

$$\tau_k^* = \frac{\sum_{j=1}^{m-1} \sum_{i=j+1}^m c(R_j(y_{k-j}) - R_i(y_{k-i}))}{0,5m(m-1)} \quad (28)$$

где функция

$$c(R_j(y), R_i(y)) = \begin{cases} 1 & \text{при } R_j(y_{k-j}) < R_i(y_{k-j}) \\ -1 & \text{при } R_j(y_{k-j}) > R_i(y_{k-j}) \end{cases} \quad (29)$$

где  $R_j(y_{k-j})$  ранги последовательности  $y_{k-j}, j = \overline{1, m-1}, k = \overline{n, m1}$  при условии, что соответствующие значения  $x_{k-j}$  упорядочены в порядке возрастания.

По аналогии с (25) проводим уточнение базовых оценок  $\tau_n$  в момент времени  $t_k = k$  с учетом оценок “аналогов”  $\tau_{k-i}, i = \overline{n, m1}$

$$\bar{\tau}_n = \frac{\sum_{i=1}^{m1} K\left(\frac{\tau_k^* - \tau_{k-i}^*}{\sqrt{D_k(r_k)} * h}\right) * \tau_{k-i}}{\sum_{i=1}^{m1} K\left(\frac{\tau_k^* - \tau_{k-i}^*}{\sqrt{D_k(\tau_k)} * h}\right)} \quad (30)$$

где

$$\tau_k^* = \text{med}(\tau_{k-i}, i = \overline{1, m_1}) \quad (31)$$

$$D_k(\tau_k) = \text{med}(|\tau_{k-i} - \tau_k^*|, i = \overline{1, m_1}), \quad K(u) = \exp(-u^2) \quad (32)$$

## 1.5 Определение параметров алгоритмов построения оценок коэффициентов линейной регрессии

В рассмотренных ранее алгоритмах не доопределенными параметрами являются:

1. Объем выборки  $m$  предыстории  $y_{k-i}, x_{k-i}, i = \overline{1, m}$ , используемый для построения каждой из пары оценок “аналогов”  $a_k, b_k$ , где  $k = \overline{n, m_1}$
2. Число оценок “аналогов”  $a_k, b_k, k = \overline{n, m_1}$  - характеризуемое числом  $m_1$ .
3. Параметр сглаживания  $h$  в комбинированных оценках параметров линейной регрессии определяющий значение весовой функции.

Рассмотрим решение задачи определения управляющего параметра  $h$  в алгоритмах оценки взаимодействия скважин в предположении, что объем выборки предыстории  $m$  и число оценок “аналогов”, т.е.  $m_1$  используемый для построения базовых оценок параметров линейной регрессии, определены. Тогда, оценку  $h^*$  параметра  $h$  определяем как решение экстремальной задачи

$$h^* = \arg \min_h J(h, m, m_1) \quad (33)$$

где функционал качества  $J(h)$  вычисляем по формуле:

$$J(h, m, m_1) = \sum_{k=n}^{n-c} (y_k - y_k^*)^2 \quad (34)$$

$y_k$  – значение дебита скважины в момент времени  $t = t_k$ , а  $y_k^* = \bar{a}_k + \bar{b}_k x_k$  прогнозные значения дебита добывающей скважины,  $\bar{a}_k, \bar{b}_k$  – уточненные значения базовых оценки параметров линейной регрессии (4) в момент времени  $t = t_k$ ,  $c$  – объем контрольной выборки ( $c \approx 3 - 5$ ). Значения  $m$  и  $m_1$  в функционале (34) подбираются эмпирически в пределах  $m \approx 10 - 12, m_1 \approx 6 - 10$ . При разных значениях данных параметров производится определение управляющего параметра  $h$  по алгоритму (33).

В качестве процедуры поиска параметра  $h^*$  желательно использовать методы без

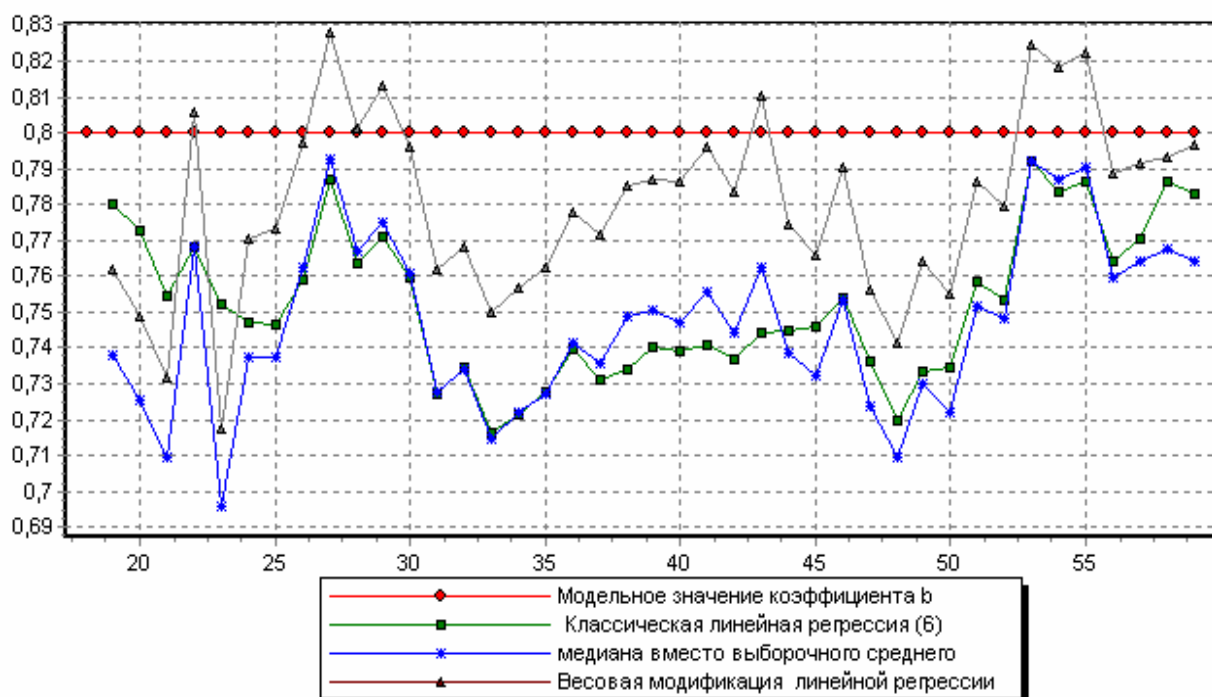
вычисления производных, дихотомии, золотого сечения и т.п., поэтому было решено искать его с помощью прямого поиска.

В качестве начального приближения  $h$  при использовании весовой функции  $K(u) = \exp(-u^2 / 2)$  рекомендуется использовать значения  $h \approx 0,5 - 2$ .

## 2 Апробация алгоритмов взаимодействия скважин на модельных и реальных данных.

### 2.1 Построение оценок на модельных данных.

Рассмотрим примеры работы алгоритмов оценки коэффициентов линейной регрессии на модельных данных. На рисунке 1 продемонстрированы результаты вычисления коэффициента  $b_k$  (4) с использованием алгоритмов: классической оценки регрессии (6), весовой модификации, формула (19) и медианной модификации выборочного среднего (10). Оценка производилась на модели стационарного процесса (см. рис. 3).



**Рисунок 1 – Оценка параметра  $b_k$ . Алгоритмы: классической оценки регрессии (6), весовой модификации, формула (19) и медианной модификации выборочного среднего (10).**

Как мы видим, алгоритм (19) дал более точную оценку коэффициента, а алгоритм, использующий оценку (10), имеет наибольшую ошибку из трех (см. таблицу 1). На рисунке 2 показаны результаты работы алгоритмов ранговой оценки корреляции Спирмена(22) и

Кендэла (28).

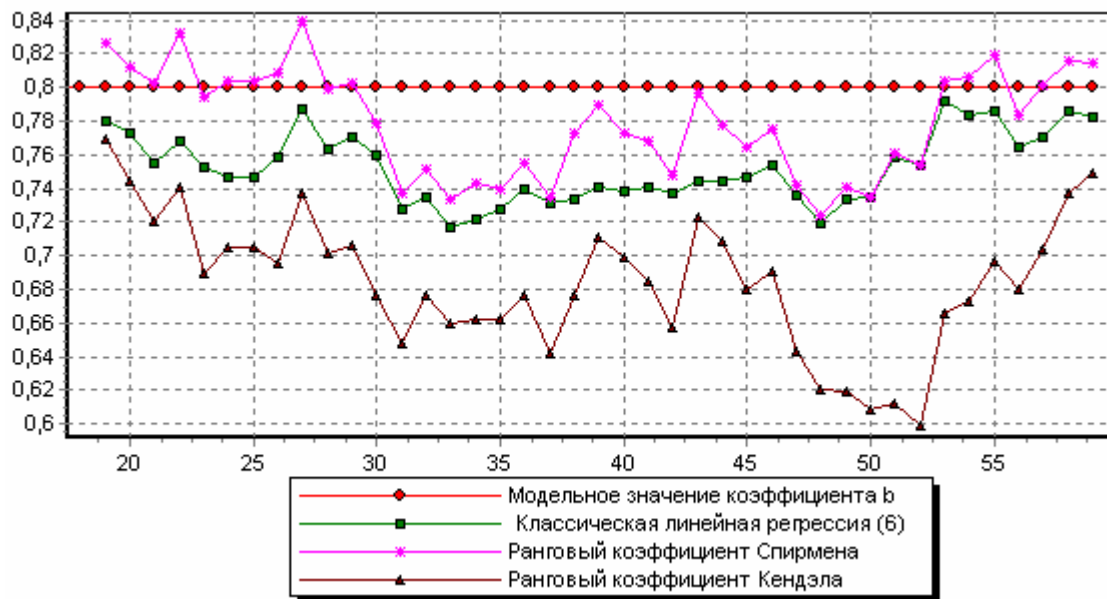


Рисунок 2 - Оценка параметра  $b_k$ . Алгоритмы: классической оценки регрессии (6), ранговой оценки корреляции Спирмена(22) и Кендэла (28).

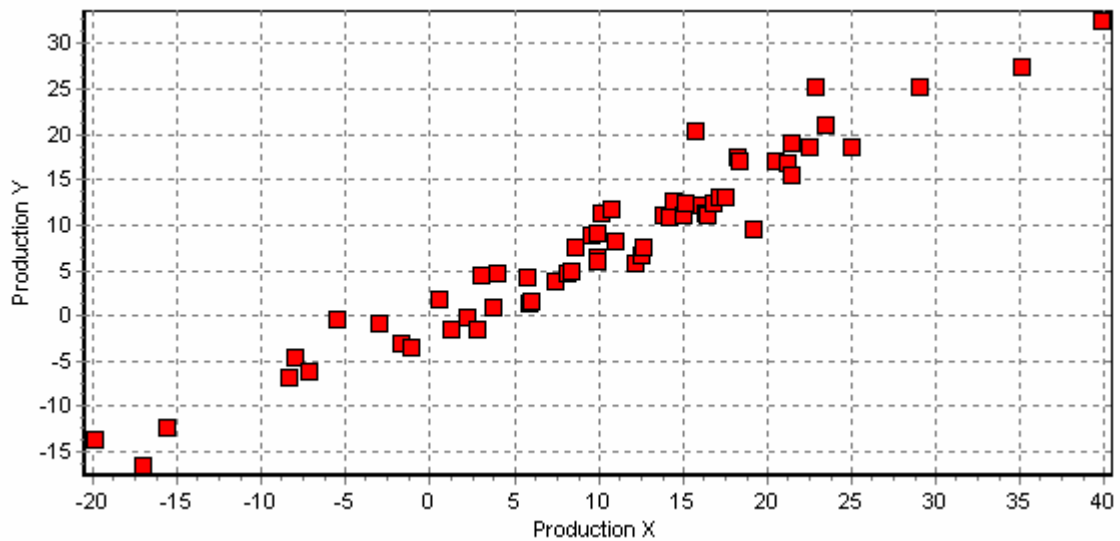


Рисунок 3 – Моделирование статического процесса.

**Таблица 1 Оценка точности построения коэффициента  $b_k$ . Классический алгоритм (6), весовой модификации, формула (19) и медианной модификации выборочного среднего (10) и ранговые коэффициенты Спирмена(22) и Кендэла(28). Оценка построена как**

$$\text{среднеквадратичное отклонение: } \frac{\sum_{k=m1}^n (b_k - b_k^*)^2}{(n - m1 + 1)}$$

	Классический алгоритм	Медианная оценка	Весовая модификация	Ранговый коэффициент Спирмена	Ранговый коэффициент Кендэла
Стационарный процесс	0,0027	0,0034	0,00101	0,001	0,015
Стационарный процесс с выбросами и зашумлением	0,14	0,14	0,0332	0,046	0,074
Нестационарный процесс	0,031	0,028	0,02362	0,028	0,055

Точность оценок заметно падает, если данные зашумлены и существуют точки выбросов. Мы смоделируем процесс с сильным шумом и точками выбросов (см. рис. 6). Общая картина результатов работы алгоритмов в этом случае схожа с рассмотренными ранее оценками, но здесь наблюдается значительное падение точности классического алгоритма оценки линейной регрессии (см. рис. 4, 5). В то время как при построении оценки коэффициента по формуле (20), точки выбросов оказали меньшее влияние на оценку из-за приписывания и маленького веса, что и отразилось на точности оценки (см. таб. 1).



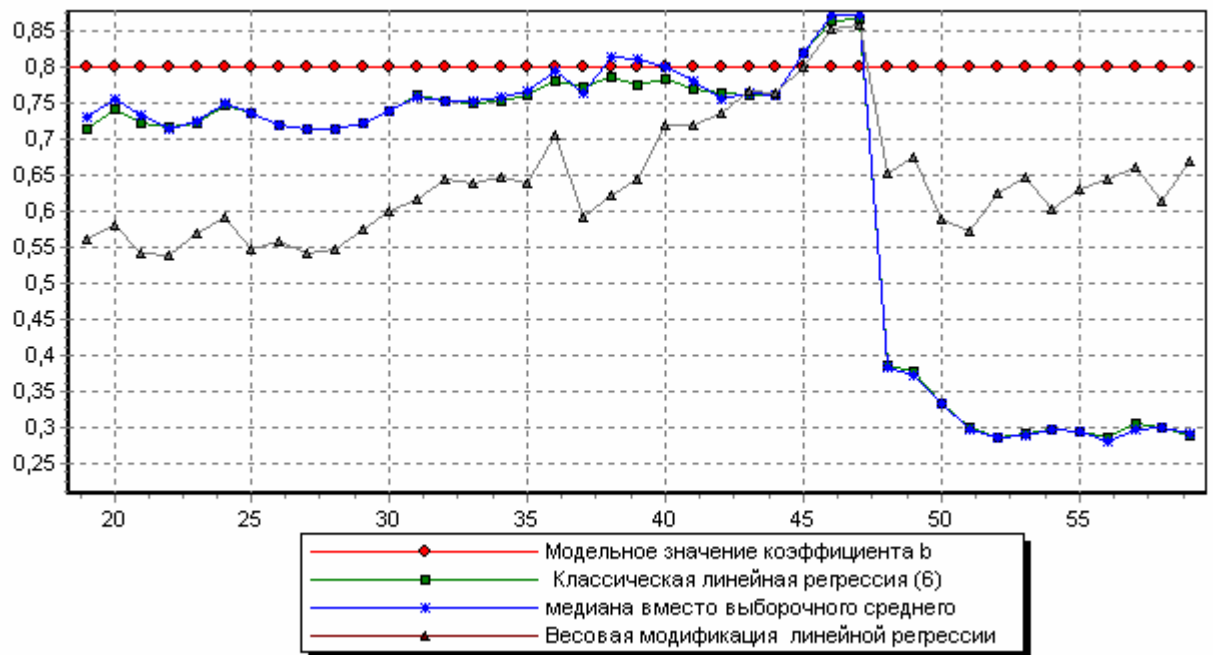


Рисунок 3 – Оценка параметра  $b_k$ . Стационарный процесс с выбросами.

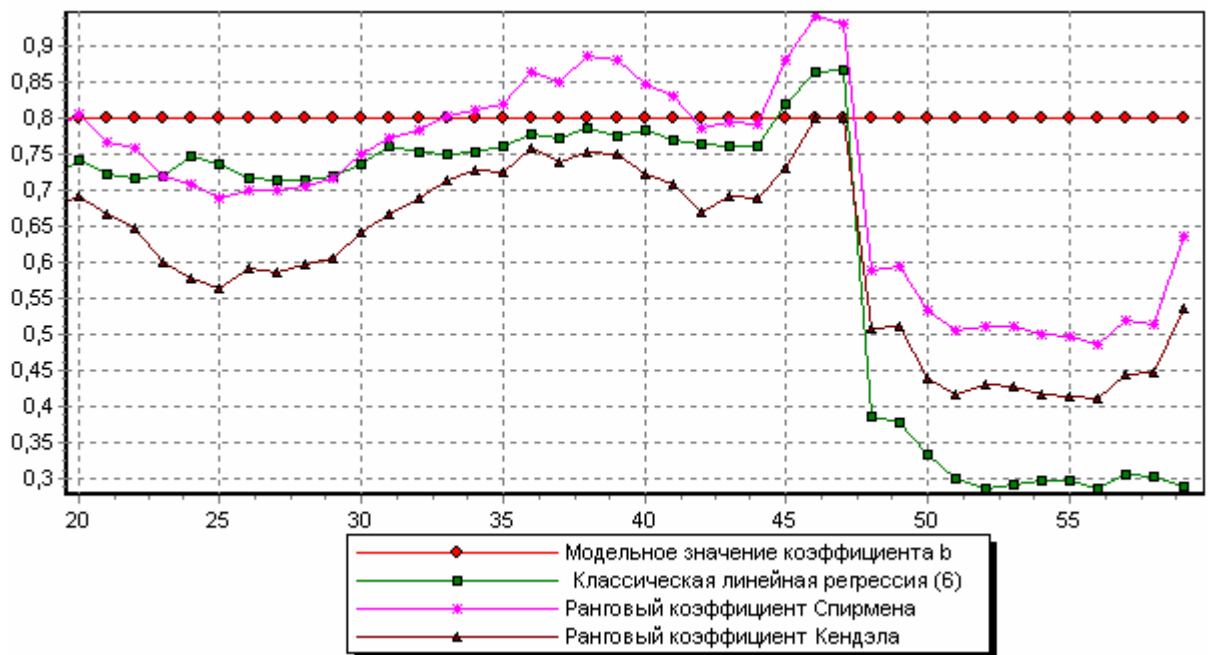
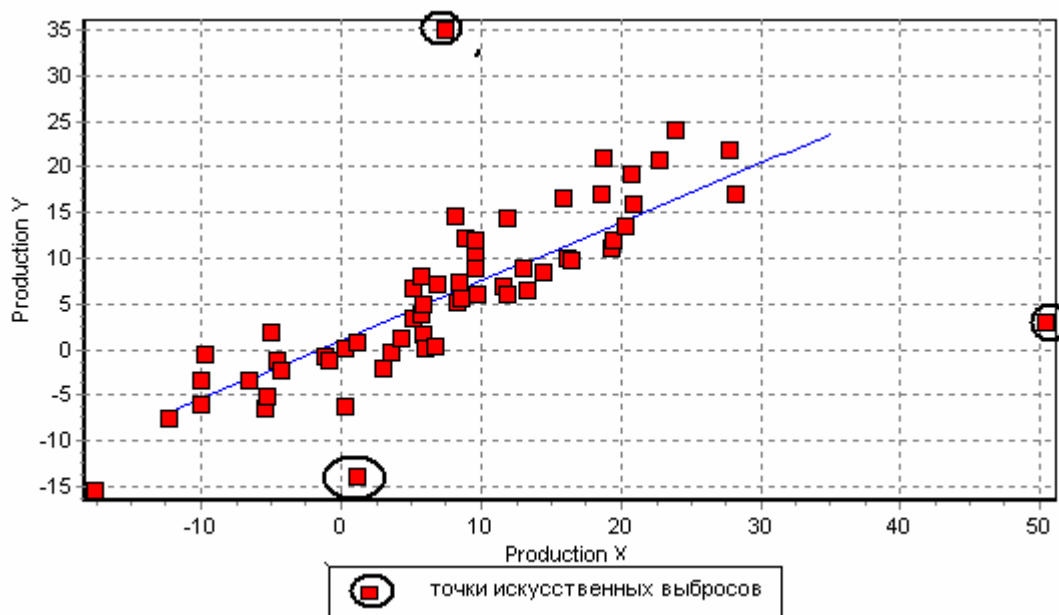
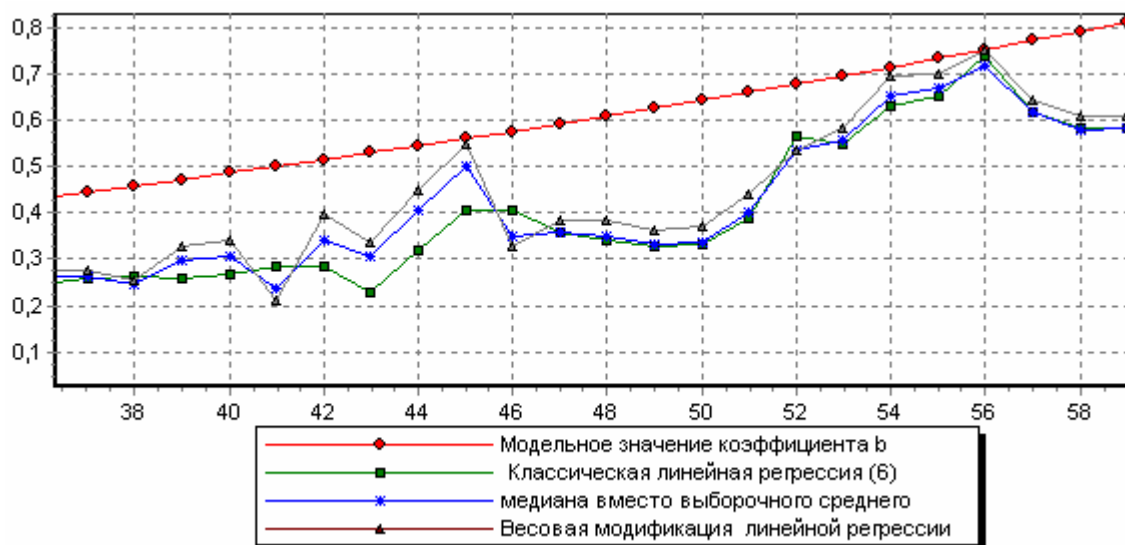


Рисунок 4 – Оценка параметра  $b_k$ . Стационарный процесс с выбросами, Ранговые коэффициенты.



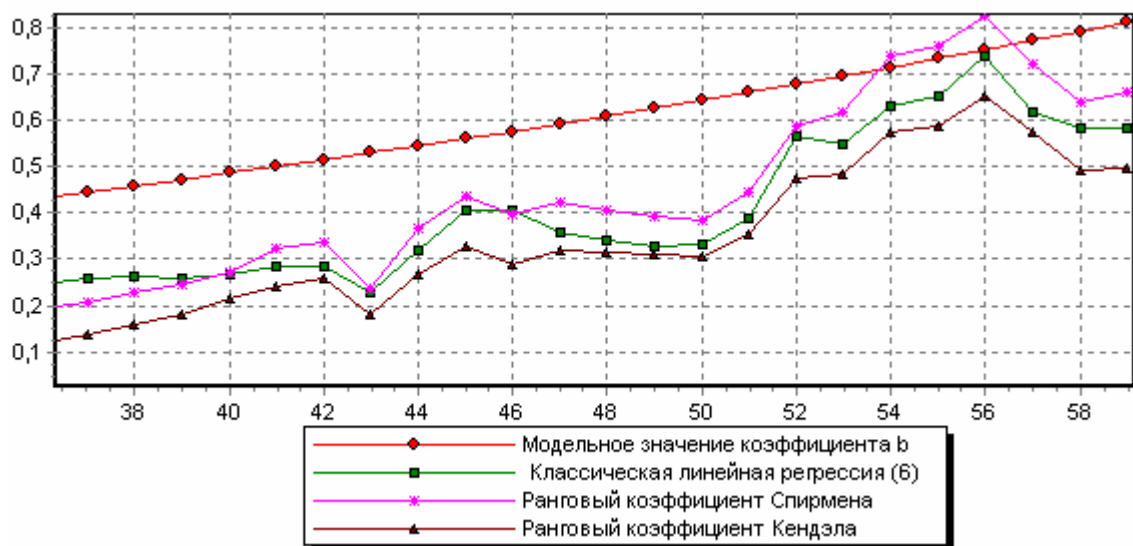
**Рисунок 6 – Модельные данные. Стационарный процесс с выбросами и зашумлением.**

Что бы закончить рассмотрение модельных данных, посмотрим, как себя ведут оценки при нестационарном процессе (см. рис. 7-8).



**Рисунок 7 - Оценка параметра  $b_k$ . Алгоритмы: классической оценки регрессии (6), весовой модификации, формула (19) и медианной модификации выборочного среднего (10). На стационарный процесс.**

Здесь стоит отметить, что медианная модификация (10) линейной регрессии дает более точную оценку чем классический вариант.



**Рисунок 8 – Не стационарный процесс. Оценка параметра  $b_k$ . Ранговые коэффициенты корреляции**

## 2.2 Апробация алгоритмов на реальных данных

В качестве примера, иллюстрирующего работу алгоритмов прогноза уровня добычи и оценки технологической эффективности ГТМ, рассмотрена задача определения дополнительной добычи нефти от интенсификации заглублением насосов с учетом изменения уровня закачки и взаимодействия скважин. Исходными данными являлись месячные значения добычи нефти, жидкости и закачки жидкости по скважинам участков 2.2 и 3.2 пласта Ю<sub>2</sub><sup>2</sup> Северо-Вахской площади (приложение А).

Результаты работы алгоритмов приведены в таблицах 2, 3 и рисунках 9, 10 для участка 2.2. В таблицах 4, 5 и рисунках 11, 12 соответственно приводятся аналогичные результаты для участка 3.2. Используемые для расчетов исходные данные приведены на рисунках 13 и 14 для групп скважин участков 2.2 и 3.2 пласта Ю<sub>2</sub><sup>2</sup> соответственно.

На рисунке 9 линией 1 изображена реальная добыча нефти за семь, последующих с начала проведения интенсификации месяцев, для скважин 1021,1022 участка 2.2. Линия 3 - базовый прогноз добычи нефти, полученный при задании закачки в нагнетательных скважинах и значений отборов жидкости в скважинах окружения на уровнях соответствующих их значениям до начала проведения интенсификации [3]. Линия 2- прогноз добычи нефти с использованием непараметрической модели (19). Следует отметить,

что одновременно с заглублением насосов значительно увеличена закачка в нагнетательные скважины участка 2.2, что соответствует увеличению прогнозных значений добычи нефти. Вычисление двух прогнозных уровней добычи нефти позволяет рассчитать не только общий эффект от заглубления насосов и увеличения закачки, но и разделить эффекты. Соответствующие данные приведены в таблице 2

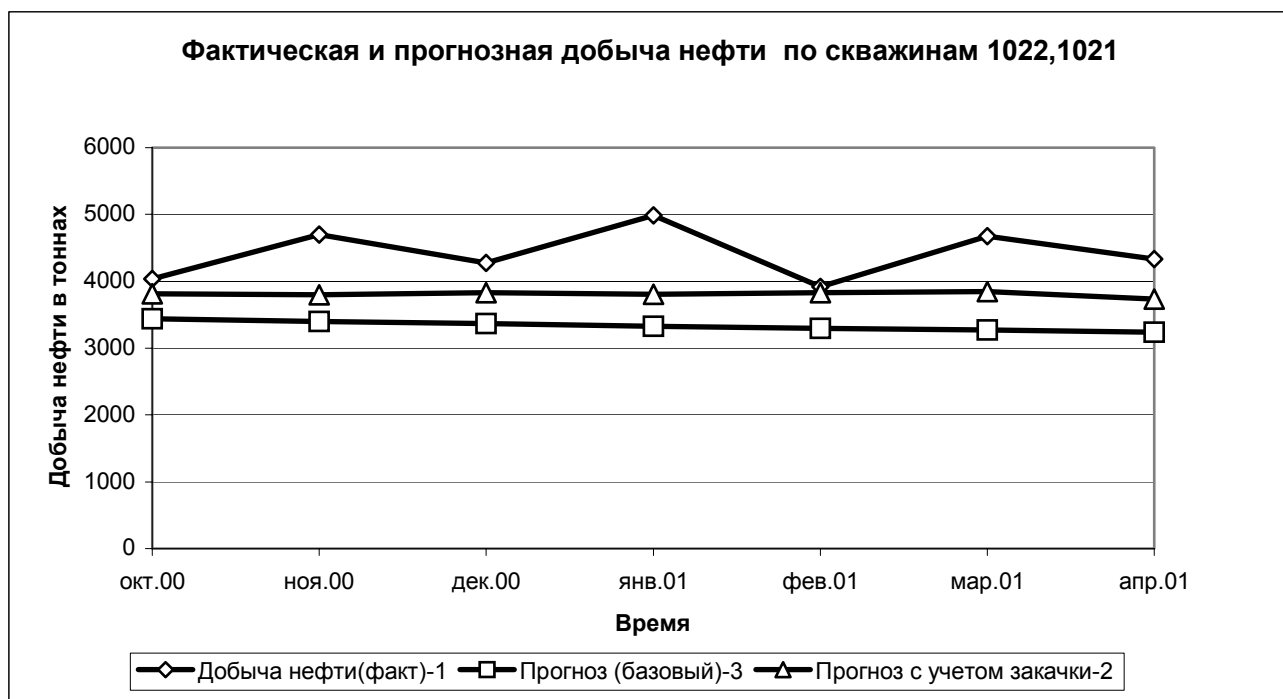
На рисунке 10 приводятся расчетные значения прогнозной добычи нефти в скважинах окружения участка 2.2 в двух вариантах. Линия прогноз 2 соответствует расчетным значениям добычи на скважинах окружения с учетом увеличения закачки в нагнетательных скважинах ячейки 2.2 и заглубления насосов на скважинах 1021,1022. Линия прогноз 1 соответствует расчетным значениям добычи на скважинах окружения с учетом увеличения закачки. Увеличение добычи нефти в данном случае объясняется интерференцией скважин которую в данном случае нетрудно определить. Конкретные значения снижения добычи нефти за счет интерференции скважин приводятся в последнем столбце таблицы 3

На рисунке 11 линией 1 изображена реальная добыча нефти за шесть, последующих с начала проведения интенсификации месяцев, для добывающих скважин участка 2.3. 2045, 1204, 1205 участка 2.2. Линия 3 - базовый прогноз добычи нефти, до начало проведение интенсификации. Линия 2- прогноз добычи нефти с использованием (19). Более подробные данные приводятся в таблице 4. Здесь следует отметить, что с учетом данных по закачке в прогнозируемом периоде прогнозный уровень добычи нефти снижается по сравнению с базовым прогнозным уровнем.

На рисунке 12 приводятся данные по добыче нефти в скважинах 2094 и 2095 . Линия прогноз 2 соответствует расчетным значениям добычи на скважинах 2094 и 2095 с учетом закачки в нагнетательных скважинах ячейки 3.2 и заглубления насосов на скважинах 1204,1205,2045 на уровне их значений. Линия прогноз 1 соответствует прогнозным значениям добычи с учетом соответствующих прогнозируемому периоду значений закачки при значениях добычи на скважинах 1204,1205 и 2045, соответствующих базовому уровню добычи.

**Таблица 2 Результаты расчета дополнительной добытой нефти при одновременном проведении геолого-технических мероприятий по заглублению насосов и увеличении закачки на скважинах 1022, 1021 участка 2.2. пласта Ю<sub>2</sub><sup>2</sup> Северо-Вахской площади**

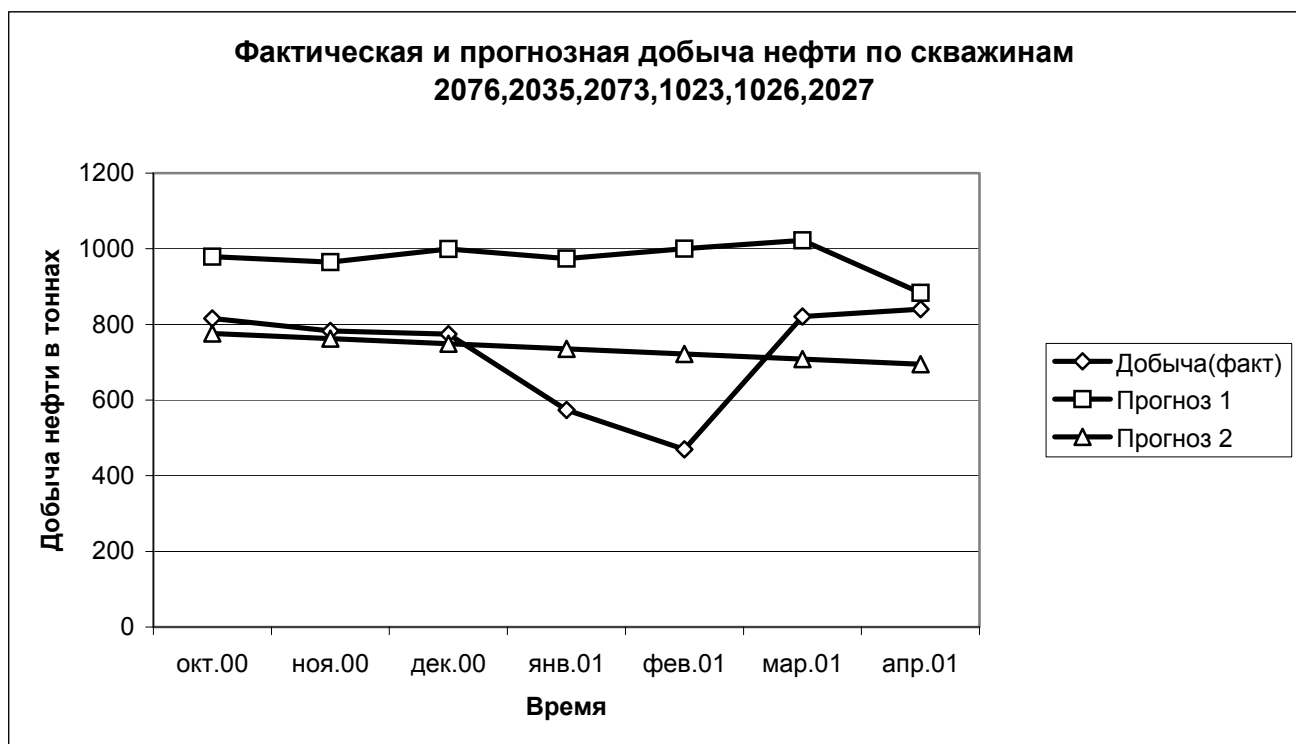
	Добыча нефти (в тоннах)	Базовая прогнозная добыча, без учета закачки	Прогноз с учетом увеличения закачки	Общая дополнительная добыча	Дополнительная добыча за счет увеличения закачки	Дополнительная добыча за счет заглубления
10.2000	4035	3440	3813	595	373	222
11.2000	4700	3440	3801	1300	401	899
12.2000	4273	3363	3829	910	466	449
01.2001	4989	3330	3809	1658	479	1179
02.2001	3918	3299	3831	619	532	87
03.2001	4679	3270	3849	1409	579	830
04.2001	4331	3243	3733	1080	491	598
Всего за 7 месяцев	30924	23344	26667	7580	3321	4259



**Рисунок 9 фактическая и прогнозная добыча нефти.**

**Таблица 3 - Результаты расчета по прогнозу добычи скважин окружения  
2076,2035,2073, 1023,1026,2027 участка 2.2 пласта Ю<sub>2</sub><sup>2</sup> Северо-Вахской площади**

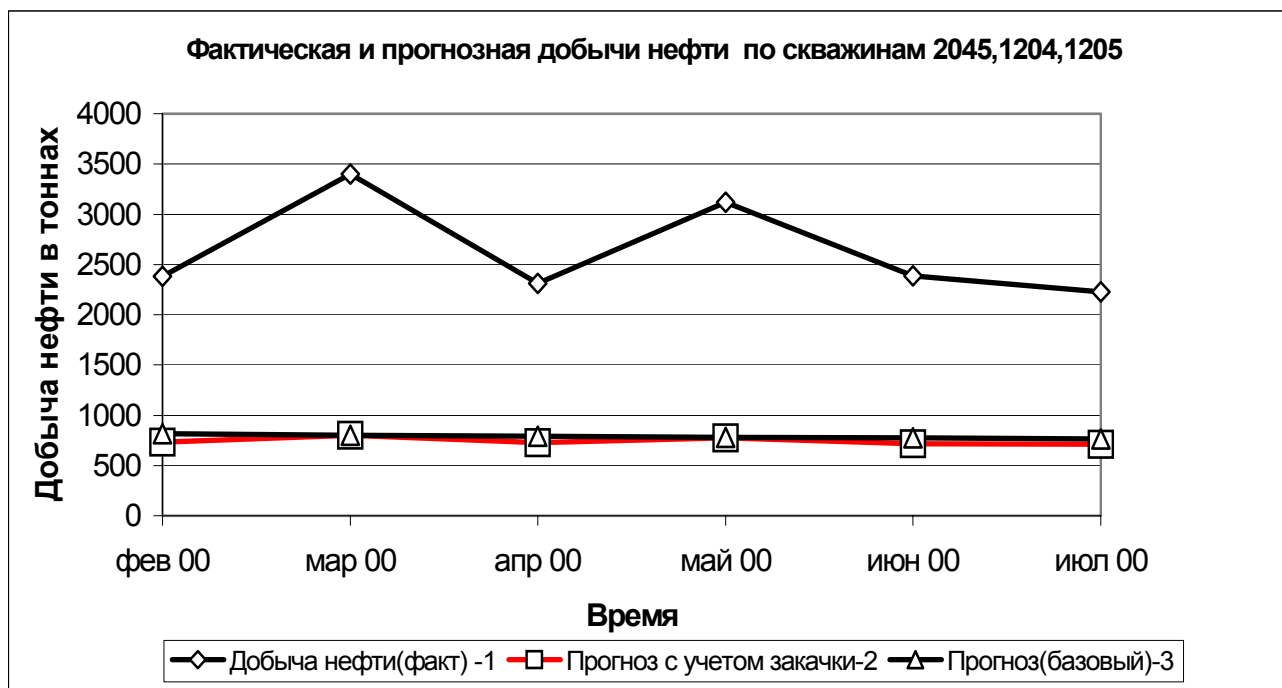
Расчетный период(месяц ,год)	Добыча нефти(в тоннах)	Прогноз добычи нефти с учетом закачки и заглубления насосов	Прогноз добыча с учетом закачки	Снижение добычи нефти за счет интерференции скважин
10.2000	816	776	979	163
11.2000	783	762	964	181
12.2000	774	749	999	225
01.2001	574	735	974	400
02.2001	470	721	1000,5	530,5
03.2001	821	708	1022,5	201,5
04.2001	840	695	883	43
Всего за 7 месяцев	5078	5147	6823	1745



**Рисунок 10**

**Таблица 4 Результаты расчета дополнительной добытой нефти при проведении мероприятия по интенсификации добычи(заглублении насосов) на скважинах 2045, 1204,1205 участка 3.2 пласта Ю<sub>2</sub><sup>2</sup> Северо-Вахской площади**

Расчетный Период (месяц,год)	Добыча Нефти (в тоннах)	Базовая прогнозная добыча, без учета закачки	Прогноз с учетом закачки	Общая дополнительная добыча	Дополнительная добыча за счет заглубления насосов	Дополнительная добыча за счет закачки
02.2000	2381	815	730	1566	1651	-85
03.2000	3403	802	800	2601	2602,5	-1.5
04.2000	2312	793	725	1519	1586	-67
05.2000	3121	780	776	2341	2345	-4
06.2000	2385	776	718	1609	1667	-58
07.2000	2228	763	711	1465	1517	-52
Всего за 6 месяцев	15810	4729	4461	11101	11369	-268

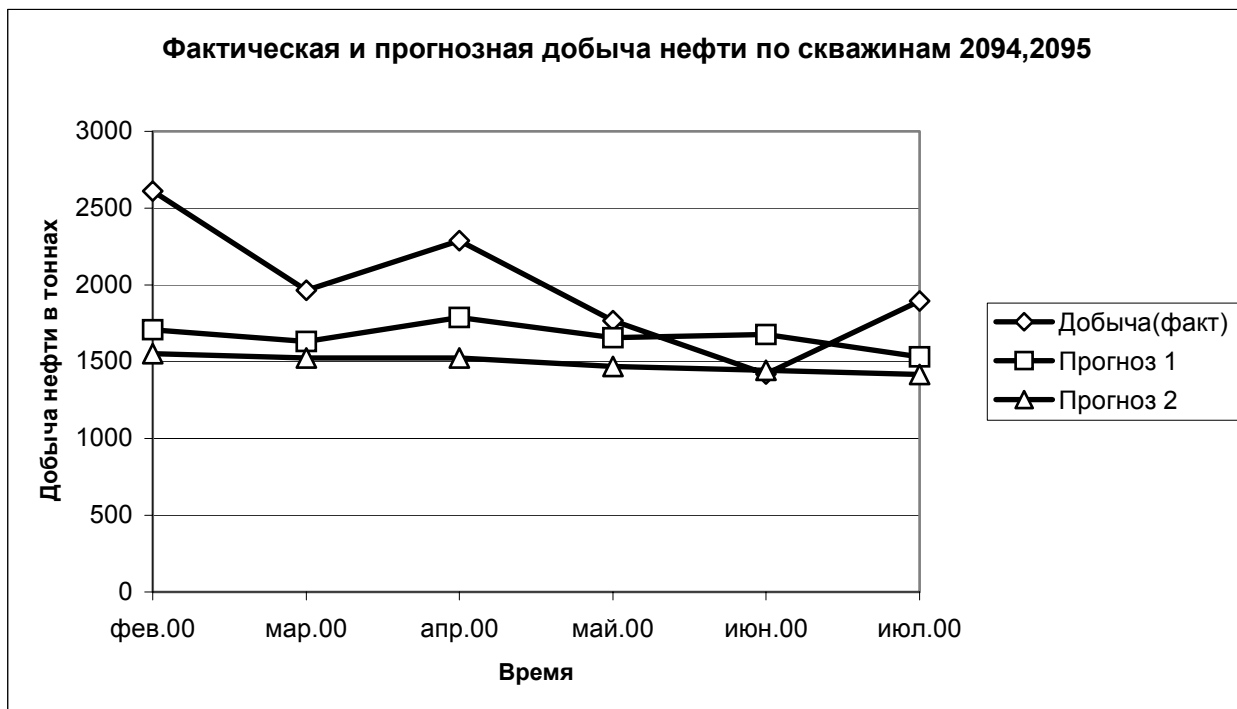


**Рисунок 11**

**Таблица 5 Результаты расчета прогноза добычи на скважинах окружения 2094,2095 при заглублении насосов на скважинах 1204,1205,2045 участка 3.2 пласта Ю<sub>2</sub><sup>2</sup>.**

Расчетный период(месяц, год)	Добыча нефти(в тоннах)	Прогноз добычи нефти с учетом закачки и заглубления насосов	Прогноз добычи нефти с учетом закачки	Снижение добычи за счет интерференции скважин
02.2000	2110	1551	1707	156
03.2000	1968	1524	1613	109
04.2000	2089	1524	1788	264
05.2000	1768	1470	1656	186
06.2000	1421	1443	1678	233
07.2000	1896	1416	1532,6	117
Всего за 6 месяцев	11951	8928	9993	1064





**Рисунок 12.**

Разность между двумя прогнозными значениями добычи означает в данном случае снижение добычи в скважинах окружения за счет интерференции скважин 2045,1204,1205 со скважинами 2094 и 2095.

Метод построения комбинированных оценок, по сравнению с методами, основанными на характеристиках вытеснения, дает возможность провести более детальный анализ сторонних факторов и учесть эффект по каждому из факторов в отдельности. На примерах оценки технологической эффективности интенсификации добычи нефти заглублением насосов, проводимой на скважинах пласта Ю<sub>2</sub><sup>2</sup> получены эффекты от увеличения закачки воды и эффекты от интерференции скважин.

Одним из важных результатов является возможность разделения эффектов, связанных с изменением уровня закачки и интерференции скважин

## 3 Программная реализация

Для решения поставленных задач был разработан программный модуль с использованием среды разработки Delphi.

В программном модуле можно выделить три основных блока:

1. Блок моделирования данных
2. Построения оценок.
3. Блок анализа и вывода результатов.

Рассмотрим работу каждого блока.

### 3.1 Блок моделирования данных

#### 3.1.1 Моделирования процессов взаимодействия скважин

Одной из поставленных задач, является возможность проверки работоспособности реализованных в программе алгоритмов оценки на модельных данных. Для этого в программе реализовано две схемы моделирования данных: стационарного и не стационарного процессов.

Первая схема – схема стационарного процесса, т. е. когда функциональная зависимость не меняется во времени. По сути, нам надо построить модели двух процессов: добычи нефти из добывающей скважины и закачка воды в нагнетательную скважину. Но, исходя из того, что нам необходимо чтобы эти два процесса были взаимосвязаны друг с другом, мы будем моделировать процессы по следующей схеме:

$$y_t = b * y_{t-1} + c\zeta \quad (35)$$

$$x_t = y_{t-1}$$

где  $c, b$  - константы, а  $\zeta$  - некоторая случайная величина, помеха. Коэффициент корреляции данных в этом случае регулируется параметром  $b$ . Тогда  $y_t$  - дебит

добывающей скважины,  $x_t$  - приемистость нагнетательной скважины, в момент времени  $t$ .

Вторая схема – нестационарный процесс, здесь используется та же схема (35), но чтобы процесс был нестационарным константы  $b$  мы используем выражение следующего типа:

$$b_t = b * t^2 \quad (36)$$

где  $b$  - начальное значение коэффициента. Таким образом, мы обеспечим достаточно сильную зависимость между  $y_t$  и  $x_t$  на коротких интервалах моделирования (причем на конечном этапе моделирования зависимость будет сильнее всего), но по мере удаления от конца интервала моделирования зависимость будет очень сильно размываться.

Так же при моделировании процессов, для учета наложения на процессы различных случайных факторов (на пример ошибки измерения), на процессы накладываются шум:

$$y_t = y_t + z\lambda$$

$$x_t = x_t + z\varepsilon$$

где  $\varepsilon, \lambda$  - некоторые случайные величины, а уровень шума регулируется параметром  $z$ .

Точки выбросов задаются вручную.

### 3.1.2 Хранение реальных данных.

Для хранения реальной промысловой информацией месторождения используется СУБД InterBase.

После того как пользователь указал месторождение с которым будет работать программа, производится построение графической схемы месторождения (см. рис. В.).

### 3.2 Построение оценок

В данном программном модуле производится построение следующих оценок:

1. Построение классической оценки линейной регрессии (6)
2. Линейная регрессия с медианой вместо выборочного среднего (10),(11)
3. Взвешенная линейная регрессия (18)
4. Ранговый коэффициент Спирмена (22)
5. Ранговый коэффициент Кэндела (28)
6. Производится уточнение базовых оценок по схеме (20)
7. Комбинированная оценка, как взвешенной среднее оценок 4 -6.

Для взвешенных оценок производится определения параметра сглаживания (33). Определения параметра производится с помощью прямого поиска. Интервалы поиска задает пользователь, но по умолчанию установлены интервалы поиска, которые определены эмпирическим способом в результате анализа результатов построения оценок на модельных данных.

### 3.3 Блок вывода результатов.

Все результаты полученные на предыдущем шаге представляются в графическом и табличном виде.

Для построенных оценок производится оценка точности по формуле:

$$J(h, m, m_1) = \sum_{k=n}^{n-c} (y_k - y_k^*)^2 \quad (37)$$

где  $y_k$  - истинные значения дебита, а  $y_k^*$  - оценка дебита..

Когда нам требуется построить прогноз добычи нефти на реальных данных нам требуется указать каким методом построить прогноз (1-7). Выбор метода производится по следующей схеме.

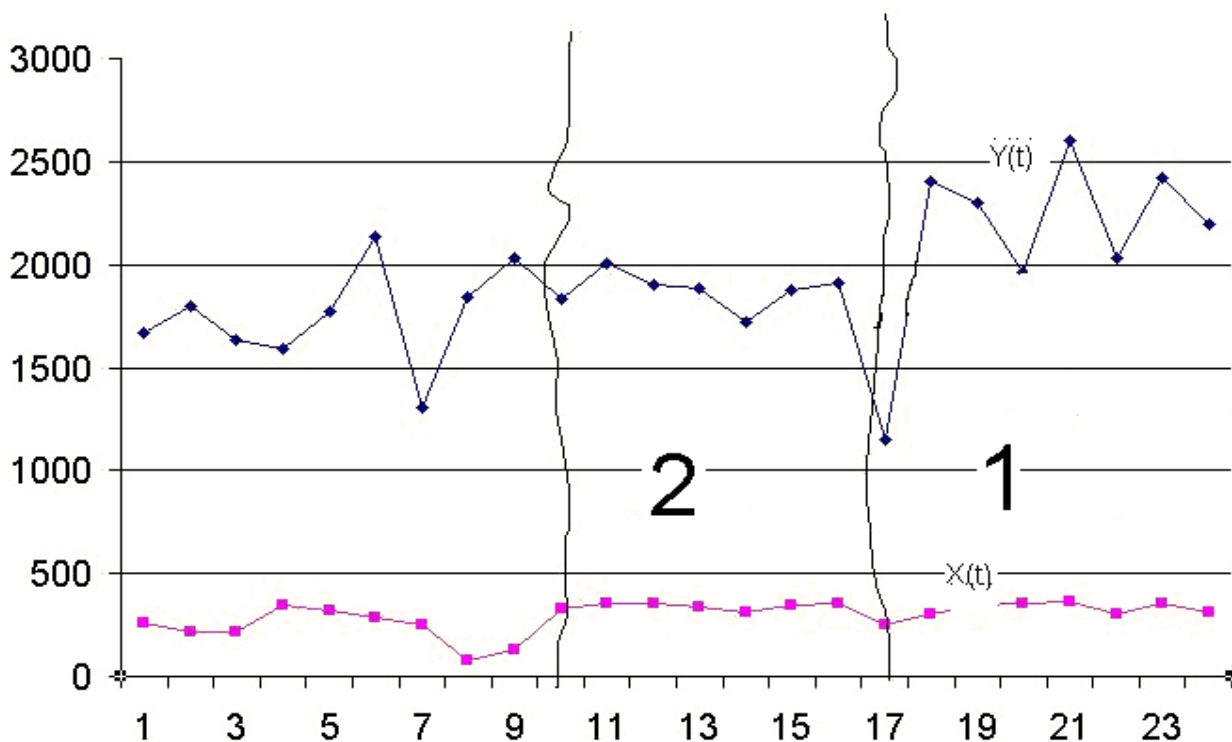


Рисунок 13

Для построения прогноза, мы будем строить оценку зависимости  $Y(t)$  от  $X(t)$  в области, на рисунке 13 обозначенной 1. На области 2 построим оценки 1-7. По полученным оценкам построим прогноз  $Y(t)$  от  $X(t)$  на интервале 1. Следующим шагом на этом же интервале построим для каждой оценки функционал (37). Для построения итогового прогноза мы выбираем тот алгоритм, который имеет минимальное из полученных значение функционала. В конечном итоге пользователь указывает каким методом он желает построить прогноз добычи нефти и на каком интервале..

Все табличные данные можно сохранить в электронном виде в формате Excel.

## Заключение

В данной работе разработан и реализован программный модуль для решения задачи оценки взаимодействия скважин нефтяного месторождения в условиях их нормального функционирования (без остановки скважин), наложения различных многочисленных факторов.

Проведена апробация алгоритмов и программ на модельных (полученных методом статистического моделирования) и реальных данных по закачке жидкости в нагнетательные (возмущающие) скважины, дебитов реагирующих нефтяных скважин.

Показано, что обучение, организованное на коротких промежутках времени предыстории процесса разработки, позволяет выявить закономерности между дебитами и приемистости скважин. Эффективными являются алгоритмы, основанные на комбинации непараметрических оценок регрессии, ранговых коэффициентов корреляции Кендэла и Спирмена.

Разработанный программный модуль позволяет выбирать (из списка скважин) возмущающие и реагирующие скважины создавать базу данных по дебитам и приемистости скважин и выбирать соответствующие алгоритмы обработки данных.

Разработанный модуль апробирован на задаче определения эффективности работы нагнетательных скважин в условиях интенсификации добычи нефти заглублением насосов. Установление зависимости между нагнетательными и добывающими скважинами по истории разработки позволило определить объем добычи нефти, полученной от увеличения закачки, разделить эффекты увеличения добычи нефти от закачки и заглубления насосов. Определены объемы добычи нефти отдельно по двум одновременно проведенным мероприятиям.

## Список использованных источников

1. Сергеев В.Л. Идентификация систем с учетом априорной информации.- Томск: Издательство научно-технической литературы, 1999.-146 с.
2. Демиденко Е.З. Линейная и нелинейная регрессии. - М.:Финансы и статистика,1981. – 300 с.
3. Дэвис Дж.С. Статистический анализ данных в геологии.-М; ”Недра”, 1990.- 320 с.
4. Уилкс. Математическая статистика.- М: Финансы и статистика,1986.- 545с.
5. Э. Ллойд. Справочник по прикладной статистике.-М: Финансы и статистика, 1990, т.2.-525с.





## Приложение Б. Руководство программиста.

Программный модуль написан с помощью среды разработки Borland Delphi 7. Программа состоит из 6 модулей:

ParentUnit – отвечает за создание формы MDIParent и запуск остальных модулей.

ParamUnit – отвечает за инициализацию и моделирование данных. А так же запуск оценок и обновление графических данных в результате работы.

DataUnit – выводит на форму “Данные” все полученные в результате построения оценок промежуточные данные.

GraphUnit – организует работу с графиками функций.

MathUnit – здесь описаны все алгоритмы построения оценок.

InterfUnit – работа с реальными данными.

Глобальные структуры данных:

ProductionX, ProductionY :array of extended; - хранение модельных данных, приемистости нагнетательной и дебита добывающей скважин соответственно.

Observation:array of TObservation; - хранение реальных данных, загружаемых из таблицы Excel.

SizeN: integer – Объем выборки дебитов скважин.

PreHystory: integer – Объем выборки предыстории, используемых для построения оценок.

LeftBorder: integer; - количество оценок аналогов.

ParamH:extended; - параметр сглаживания h

StepHParam: - шаг изменения параметра h при процедуре его уточнения для веовых функций.

Nose – уровень шума накладываемого на модельные данные (в процентах).

Переменные хранения коэффициентов линейной регрессии и коэффициента корреляции:

Классическая линейная регрессия:

ClassicLinRegretionBk:array of extended;

ClassicLinRegretionAk:array of extended;

ClassicLinRegretionCor:array of extended;

Весовая модификация выборочного среднего:

ModifyMiddleLinRegretionBk:array of extended;

ModifyMiddleLinRegretionAk:array of extended;

ModifyMiddleLinRegretionCor:array of extended;

Медианная модификация среднего значения:

MedModifyLinRegretionBk:array of extended

MedModifyLinRegretionAk:array of extended;

MedModifyLinRegretionCor:array of extended;

Весовая модификация дисперсии

ModDispersionLRBk:array of extended;

ModDispersionLRAk:array of extended;

ModDispersionLRCor:array of extended;

Модифицированная оценка линейной регрессии:

AllDispBk:array of extended;

AllMedDispAk:array of extended;

AllMedDispCor:array of extended;

Коэффициент корреляции Спирмена:

ClassicSpirmanRk:array of extended;

ClassicSpirmanAk:array of extended;

ClassicSpirmanCor:array of extended;

Коэффициент корреляции Кендэла:

ClassicKendelRk:array of extended;

ClassicKendelAk:array of extended;

ClassicKendelCor:array of extended;

Основные функции программного модуля:

procedure ClassicLinearRegression(x,y:array of extended;var Bk,Ak,Cor:array of extended; n,m1,m:integer); - построение классической оценки линейной регрессии: x,y – приемистость и дебиты скважин. Bk, Ak – коэффициенты линейной регрессии.

procedure ModifLinearRegression(x,y:array of extended;var Bk,Ak,Cor:array of extended; n,m1,m:integer;h:extended); - построение оценки с взвешенным выборочным средним

procedure MedModifLinearRegression(x,y:array of extended;var Bk,Ak,Cor:array of extended; n,m1,m:integer); - построение оценки линейной регрессии, с медианой вместо взвешенного среднего.

procedure ClassicSpirmanEvaluation(x,y:array of extended;var Rk,Ak,Cor:array of extended; n,m1,m:integer); - Ранговый коэффициент корреляции Спирмена.

procedure ClassicKendelEvaluation(x,y:array of extended;var Rk,Ak,Cor:array of extended; n,m1,m:integer); - Ранговый коэффициент корреляции Кендэла.

procedure ModifDispersionLinearRegression(x,y:array of extended;var Bk,Ak,Cor:array of extended; n,m1,m:integer;h:extended); Взвешенная оценка дисперсии.

procedure AllModifLinearRegression(x,y:array of extended;var Bk,Ak,Cor:array of extended; n,m1,m:integer;h,h1:extended); Модифицированная оценка линейной регрессии.

procedure AveragingEvaluation(B:array of extended;var eBsr:extended; n,m1:integer;h:extended); Уточнение базовых оценок. Значение помещается в переменную eBsr/

function EvaluationHparam(var h:extended;h1:extended;X,Y:array of extended;BeginStep,Prec:extended;n,m1,m:integer;Mode:byte):integer; - процедура вычисления параметра h для взвешенных оценок.

Функции работы с графиками.

procedure DrawSeries(y:array of extended;Point:TColor); overload; - вывести график функции y (здесь и далее требуется указать на какую форму вывести, с помощью параметров MDIChild форм.)

procedure DrawSeriesXY(x,y:array of extended;Point:TColor); overload;- вывести график функции y, где значения по оси ЩЧ находятся в x.

procedure DrawSeries(y:array of extended; Point:TColor; Mode:byte; Righ,Left:integer);overload; - вывести график функции y в серия Mode (для форм, где отображается несколько графиков)

procedure AddSeries(Color:TColor;Title:string;Vis:boolean); overload; Добавить график на форму.

Функции интерфейса пользователя:

procedure GeneretedButtonClick(Sender: TObject); - запустить моделирование.

procedure RunWorkButtonClick(Sender: TObject); - запустить построение оценок.

function CheckParam(mode:byte):boolean; - процедура инициализации. Сбор введенной пользователем информации и проверка её на допустимость.

procedure StartGrafics ; - процедура инициализации графиков.

procedure RefreshData(Mode:byte); процедура, производящая обновление данных, обновление графиков.

procedure ClearEvl; - процедура окончания работ. Закрытие Excel приложения,

освобождение памяти сохранения данных в файле если нужно.

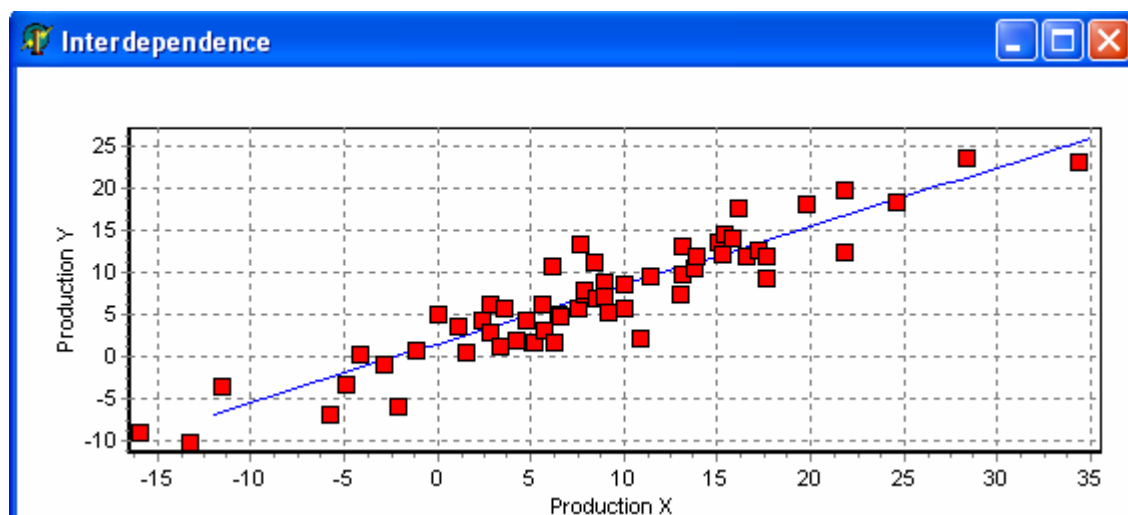
## Приложение В. Руководство пользователя.

Программа состоит из исполняемого файла: Iteraction.exe.

После запуска программы вы увидите главную форму программы и инициализации вычислений (см. рис. В.1). Все управление происходит через главное меню. В разделе графики, вы можете посмотреть при работе зависимость между величинами X и Y (см. рис. В.2), графики оценки параметров линейной регрессии. А так же вывести на экран форму данных (рис. В.3). В разделе “Данные” происходит выбор режима работы: Модельные данные или реальные данные. Загрузка реальных данных происходит из Excel таблицы.



Рисунок В.1 – Главная форма.

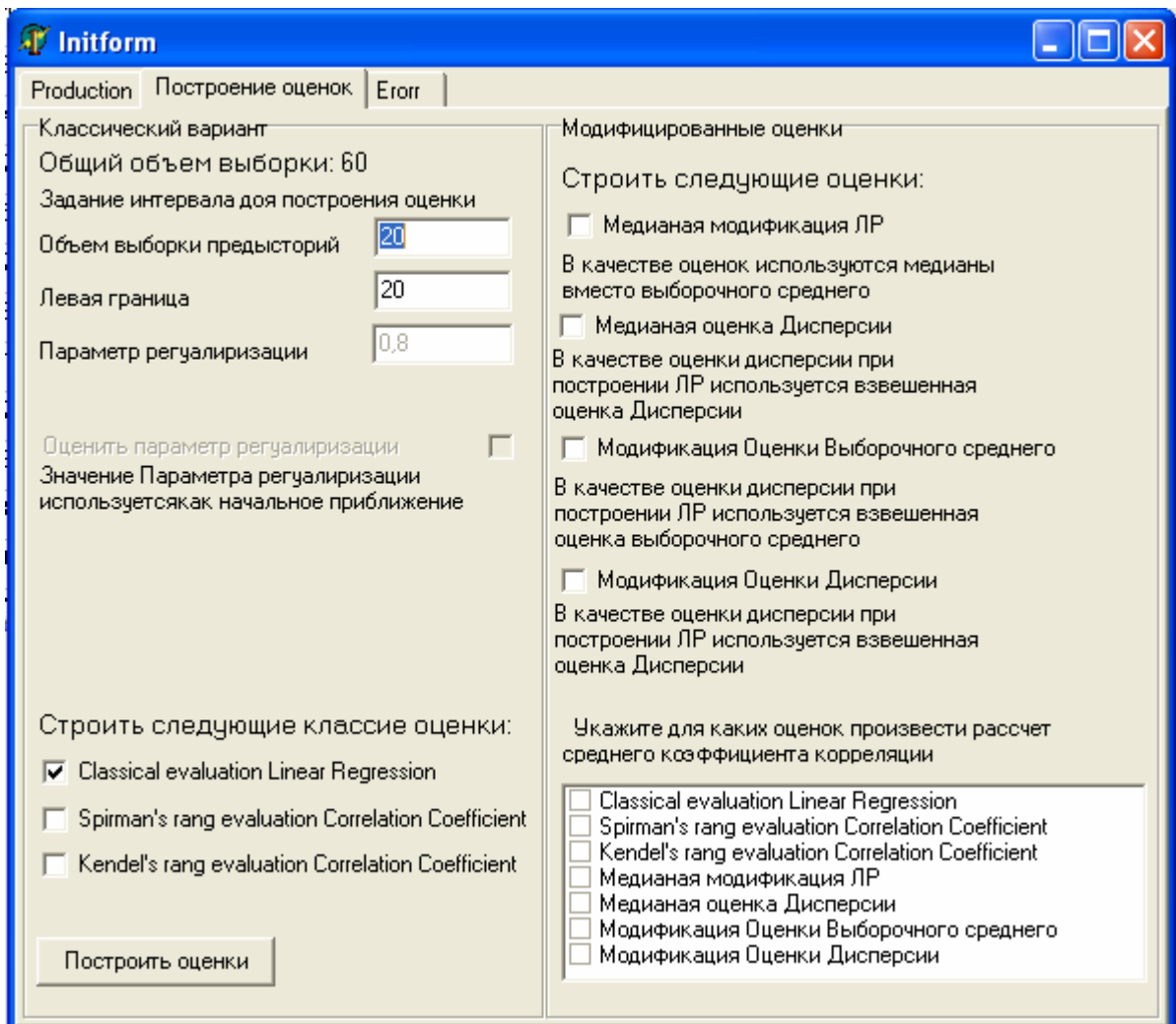


**Рисунок В.2 – форма с графиком зависимости X и Y. Прямая - прямая регрессии.  
 Параметры оценены по классическому варианту оценки линейной регрессии.**

№	Выборка X	Выборка Y	Correlation Coefficient
1	5,61725040256935	6,22031035016159	0,8
2	17,6713729670112	9,40610919928692	0,8
3	-4,12080816487735	0,250075564518885	0,8
4	-5,75651560411203	-6,9001260207679	0,8
5	15,3559945782488	14,6925674028662	0,8
6	16,5325274159363	12,0009033524187	0,8
7	21,8624089842977	19,9251973977801	0,8
8	7,5489261103557	5,79830077825406	0,8
9	17,5945663441565	12,0048971877608	0,8
10	7,85018572191985	7,38842773396091	0,8
11	-1,11058521398516	0,661457398214441	0,8
12	13,767855178826	10,5330568027433	0,8
13	13,0003795911034	7,44680839229833	0,8
14	-15,9334625164966	-8,97582317234228	0,8
15	-2,13614870330377	-5,85190046200116	0,8
16	-11,5584167962861	-3,61495593472128	0,8

**Рисунок В.3 – Форма данных.**

Итак рассмотрим теперь основную форму: инициализации вычислений (см. рис. В.4).



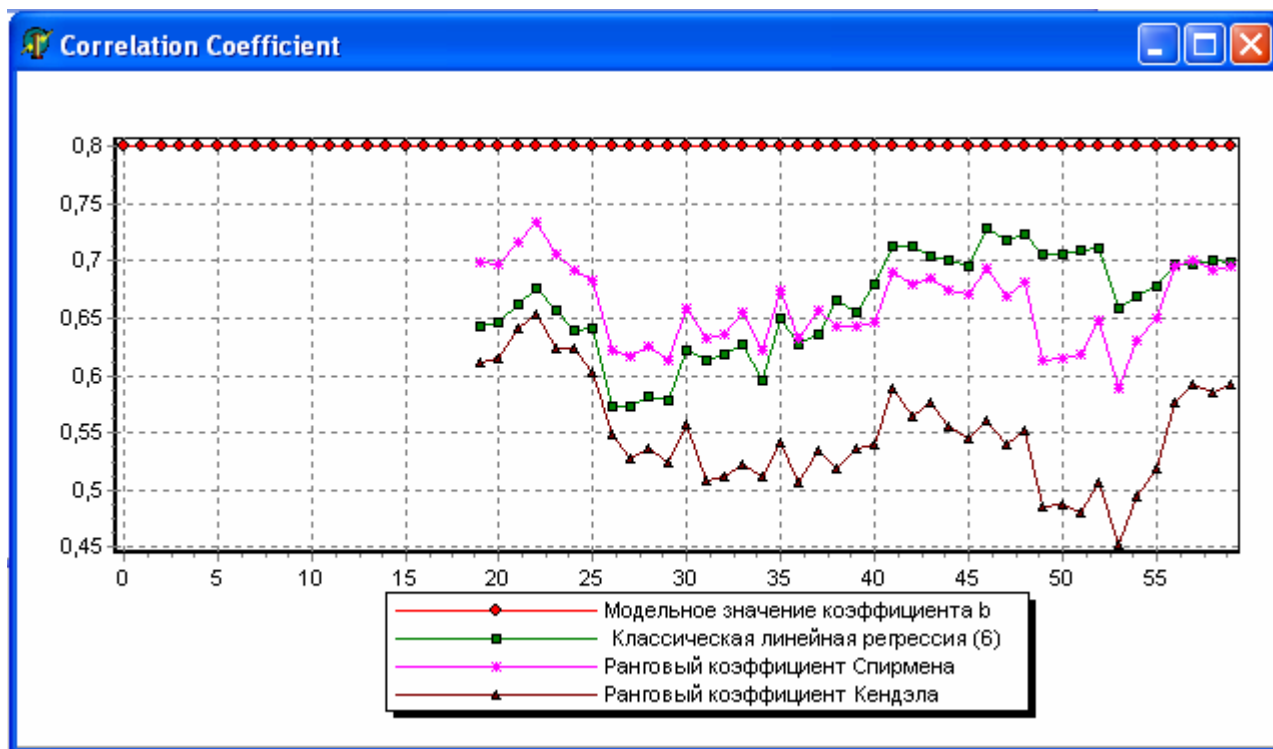
**Рисунок В.4 – Форма инициализации вычислений.**

При работе с модельными данными здесь вы задаете объем выборки наблюдений. Тип моделируемого процесса: стационарный, не стационарный. Задается параметр  $b[k]$  (35), для не стационарного процесса это начальное значение. Так же вы можете указать уровень шума, накладываемый на данные. Выбросы добавляются уже после построения. Для этого в форме данных (см. рис. В.3) выберете закладку исходные данные. За тем с помощью двойного щелчка на ячейки, значение которой вы хотите изменить, вы сможете редактировать значение. После нажатия клавиши перевода каретки. Изменённое значение занесется в модельные данные. Эти изменения вы сможете тут же увидеть на форме с графиком зависимости (см. рис. В.2)

В закладке “Построение оценок” задаются интервалы построение оценок, объем выборки предыстории, значение параметра сглаживания, только для оценок с весовыми



функциями. Здесь же происходит выбор: построение каких оценок произвести, а так же, требуется ли произвести уточнение базовых оценок. После этого необходимо запустить вычисления, с помощью кнопки “построить оценки”. Результаты вычислений выведутся в окно формы данных (см. рис. В.3). Оценки параметров линейной регрессии можно посмотреть и в графическом виде, они выведутся на форму “Коэффициенты корреляции” (см. рис. В.5).



**Рисунок В.5 – График оценки коэффициента  $b[k]$  линейной регрессии.**

Все графики в программе вы можете вы можете скролировать с помощью правой клавиши мыши.

В меню файл вы можете указать, в какой файл сохранить результаты работы. Все вычисленные значения сохраняются в электронной таблице Excel.