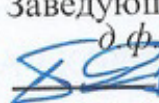


МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение
высшего образования
«ТЮМЕНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
ИНСТИТУТ МАТЕМАТИКИ И КОМПЬЮТЕРНЫХ НАУК
Кафедра фундаментальной математики и механики

РЕКОМЕНДОВАНО К ЗАЩИТЕ В ГЭК
И ПРОВЕРЕНО НА ОБЪЕМ
ЗАИМСТВОВАНИЯ

Заведующий кафедрой
д.ф.-м.н., доцент

Гатосов А.В.
2018 г.

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ
АВТОМАТИЗИРОВАННЫЙ РАСЧЕТ ВЗАИМОВЛИЯНИЯ СКВАЖИН

01.04.01 Математика

Магистерская программа «Математическое моделирование»

Выполнил работу
Студент 2 курса
очной формы обучения



Лещенко
Антон
Александрович

Руководитель работы
к.ф.-м.н.,
доцент



Бутакова
Нина
Николаевна

Рецензент
директор департамента
ООО «ТИНГ»



Комягин
Анатолий
Игоревич

Тюмень, 2018

ВВЕДЕНИЕ	2
ГЛАВА 1. НАУЧНО-ИССЛЕДОВАТЕЛЬСКАЯ РАБОТА.....	4
1.1 Математическая постановка задачи.....	4
1.2 Анализ математических методов решения задач корреляции в различных областях наук	5
1.2.1 Метеорология	5
1.2.2 Экономика.....	13
1.2.3 Физика Солнца	20
ГЛАВА 2. ОПИСАНИЕ МЕТОДОВ РЕШЕНИЯ ЗАДАЧ КОРРЕЛЯЦИИ.....	31
2.1 Спектральный анализ	31
2.2 Вейвлет-анализ	34
2.3 Фрактальный анализ	37
ГЛАВА 3. МЕТОД ОПРЕДЕЛЕНИЯ УРОВНЯ ВЗАИМОВЛИЯНИЯ НАГНЕТАТЕЛЬНЫХ И ДОБЫВАЮЩИХ СКВАЖИН С ПРИМЕНЕНИЕМ ВЕЙВЛЕТ-АНАЛИЗА	40
ГЛАВА 4. ТЕСТИРОВАНИЕ АЛГОРИТМА НА ИДЕАЛЬНОМ ПРИМЕРЕ ДАННЫХ	46
ГЛАВА 5. ТЕСТИРОВАНИЕ АЛГОРИТМА НА РЕАЛЬНОМ ПРИМЕРЕ ДАННЫХ	52
ЗАКЛЮЧЕНИЕ	56
СПИСОК ЛИТЕРАТУРЫ.....	57

ВВЕДЕНИЕ

Цель диссертационной работы состоит в создании автоматического алгоритма, способного оценивать степень влияния нагнетательной скважины на добывающую скважину.

Для достижения данной цели был поставлен ряд **задач**:

1. Анализ математических методов решения задач корреляции в различных областях наук.
2. Разработка алгоритма для расчета взаимовлияния скважин на основе промысловых данных.
3. Тестирование разработанного алгоритма на идеальном примере данных.
4. Тестирование разработанного алгоритма на реальном примере данных.

Оптимальное управление режимами эксплуатации скважин является важной и актуальной задачей современной нефтепромысловой практики: на многих месторождениях процесс заводнения пластов характеризуется низкой эффективностью, наблюдается высокая перекомпенсация отборов жидкости.

Для определения качественных характеристик взаимодействия нагнетательных и добывающих скважин в системе поддержания пластового давления ставится задача оценки взаимовлияния скважин – исторического отклика показателей работы добывающих скважин (дебита жидкости) на изменение режимов работы нагнетательных скважин (приемистости) («шахматки»).

В настоящее время в ПК Атлас «Управление Заводнением» реализована методика, позволяющая оценивать корреляцию показателей скважин в системе «пара скважин». Такой подход имеет недостаток, а именно – трудо-и-время-затратность, пользователю необходимо самостоятельно оценивать тренды, сопоставлять события для того чтобы определить влияние нагнетательной скважины на добывающую. В связи с чем, целью данной работы и является создание алгоритма, который способен оценить взаимовлияние скважин автоматически, с минимальным участием

человека, кроме того алгоритм позволит исключить влияние человеческого фактора, а именно, субъективную оценку.

ГЛАВА 1. НАУЧНО-ИССЛЕДОВАТЕЛЬСКАЯ РАБОТА

1.1 Математическая постановка задачи

Исходными данными для нашей задачи о взаимовлиянии скважин являются показатели приемистости нагнетательных скважин и дебита жидкости добывающих скважин, которые представляют собой временные ряды $\{X_t\}$ - последовательность значений переменной, соответствующей возрастающему порядку моментов времени t :

$$\{X_t\}, \text{ где } t = 1, 2, \dots, T$$

Временной ряд – упорядоченная совокупность измерений во времени определенной величины какого-либо процесса. Временные ряды делятся на непрерывные, дискретные, равномерные и неравномерные. Вид ряда определяется способом записи его результатов. Одним из общих признаков всех временных рядов является наличие в них резких изменений, то есть трендов, которые являются периодическими колебаниями разной длительности, то есть процесс может быть, как и с различными резкими переменами, так и может характеризоваться устойчивостью (длительный временной интервал с малыми колебаниями), а также незначительные перемены, например составляющие шума.

Основная задача анализа временных рядов состоит в интерпретации исследуемых (наблюдаемых) данных, которая осуществляется с помощью использования различных математических моделей. Данными моделями, чаще всего, являются функции, которые достаточно детально изучены.

Модель временного ряда состоит из двух видов данных качества сигнала, а именно, «правильных» и «неправильных» изменений, это означает, что модель является ничем иным, как «сигнал + шум». В таком виде модели «сигнал» записывается с помощью определенного набора функций, а «шум» интерпретируется с помощью иного, особого вида функций, который носит название белого шума, то есть резкие и частые

изменения в динамике исследуемого ряда. В случаях, когда невозможно определить границу между данными чистого сигнала и шумом, в качестве моделей для анализа временных рядов могут использоваться различные случайные процессы.

1.2 Анализ математических методов решения задач корреляции в различных областях наук

Изучение различных областей наук показывает, что существует сходство, а часто и общность механизмов построения, функционирования и развития различных систем. Следствием этого является применимость одних и тех же математических методов в различных областях.

В работе представлено несколько примеров реальных задач корреляции параметров с их последующим решением.

Коэффициент корреляции - нормированная величина, которая изменяется в пределах от -1 до $+1$. Если значение коэффициента корреляции не равно нулю и больше, то это позволяет сделать вывод о том, что существует связь между показателями, то есть говорит о том, что данные показатели зависимы друг от друга (например, объем продаж на рынке торговли вырос, а, следовательно, увеличилась выручка от данной продажи). Если коэффициент корреляции меньше нуля, то связи между показателями нет, они являются независимыми друг от друга (например, спрос на товар увеличился после того, как произошло планомерное понижение цены).

1.2.1 Метеорология

Нахождение взаимосвязи между метеорологическими данными, позволяет анализировать влияние одних показателей на другие: временные ряды изменения температур, давления, скорости ветра, выпадения осадков, влажности и др. В ходе анализа строится модель долгосрочного прогнозирования, с помощью которой предсказывается атмосферная изменчивость, циркуляции воздушных масс и др.

К примеру, один из известных английских ученых, а именно, Гильберт Уокер работая в Индии рассматривал проблему о появлении муссонов и возможности их предсказания. Из-за муссонов нередко возникали катастрофические ливни, которые наносили огромный ущерб. Изучая, показания с барометров, ему удалось найти зависимость между полученными им данными со станций в западной и восточной части Тихого океана, тем самым он обратил внимание на возможность предсказания погодных аномалий.[2]

Задача.

Изучение корреляции метеорологических параметров между парами метеорологических станций Саудовской Аравии, расположенных в разных стратегических местах. [19]

Данные: поверхностное давление, температура воздуха, относительная влажность, скорость ветра, количество осадков по девяти станциям (рис. 1-5).

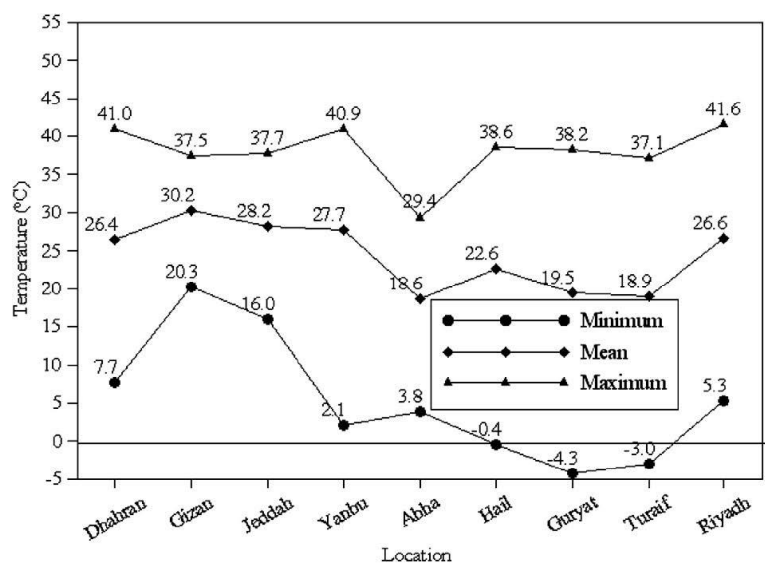


Рис. 1. Представление временных рядов минимальных, средних и максимальных температур зафиксированных в различных городах Саудовской Аравии.

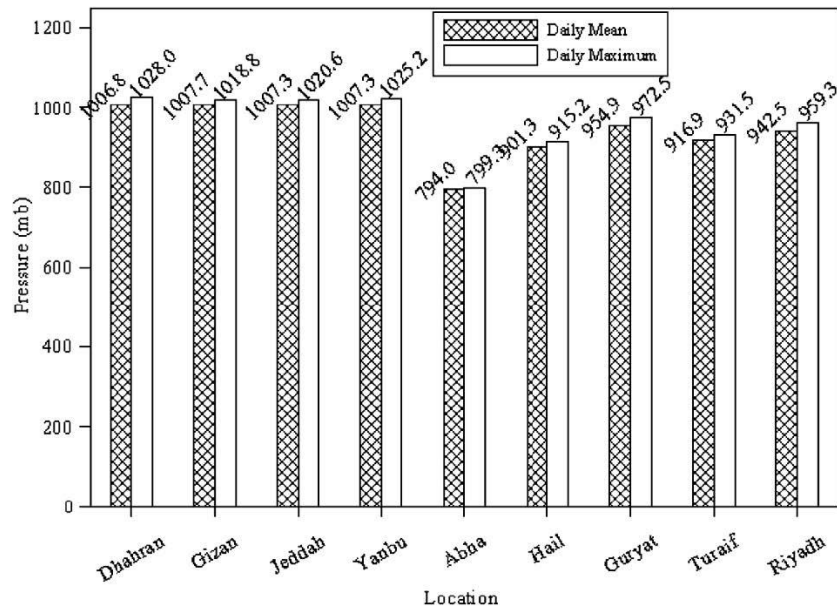


Рис. 2. Представление временных рядов средних и максимальных давлений в городах Саудовской Аравии

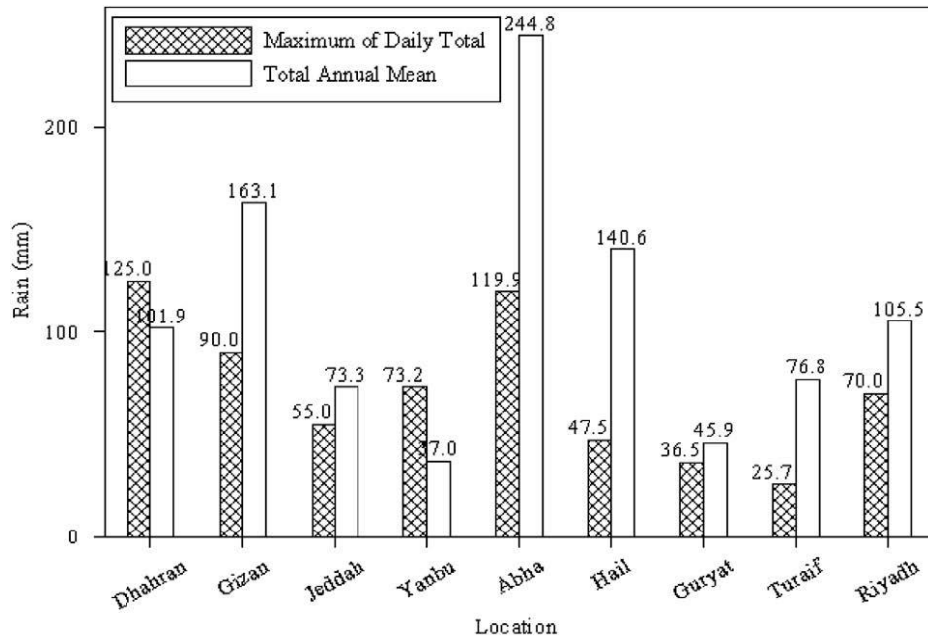


Рис. 3. Представление временного ряда атмосферных осадков зафиксированных в различных городах Саудовской Аравии

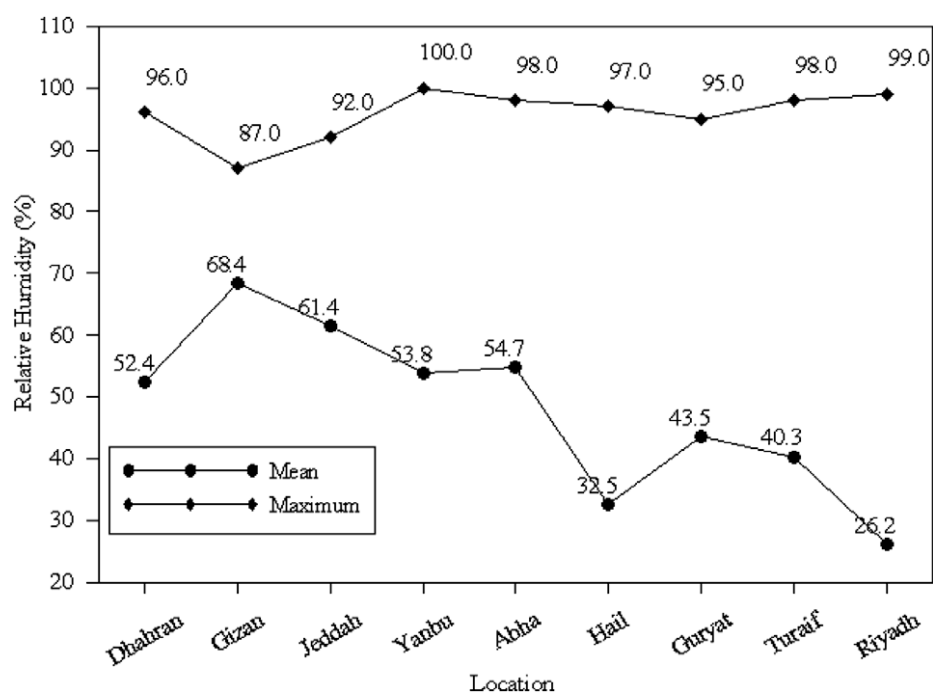


Рис. 4. Представление временных рядов средней и максимальной влажности в городах Саудовской Аравии

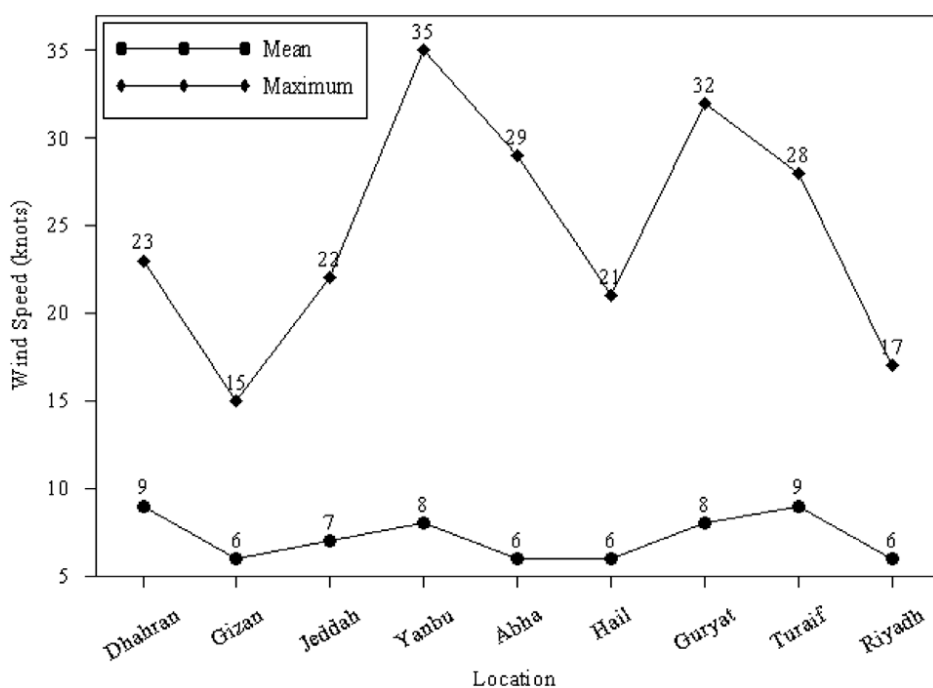


Рис. 5. Представление временных рядов средней и максимальной скорости ветра в городах Саудовской Аравии

Математические методы: вейвлет-анализ.

Для анализа корреляции метеорологических данных в Саудовской Аравии использовался кросскорреляционный коэффициент r :

$$r = \frac{\sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{X})^2 \sum_{i=1}^n (y_i - \bar{Y})^2}}, \quad (1)$$

где x_i, y_i – значения временного ряда, i – количество замеров, \bar{X}, \bar{Y} – среднее значение выборок. Данный коэффициент r также называется линейным коэффициентом корреляции Пирсона. Данный коэффициент определяет насколько сильна связь между двумя коррелируемыми показателями (различными процессами, временными рядами, сигналами) в случае наличия между ними линейной зависимости. На практике коэффициент применим, когда распределение исследуемого процесса близко к нормальному распределению.

Если этот коэффициент положительный, то он показывает, что если одна величина возрастает, то другая в этот момент имеет тенденцию к повышению, тогда как отрицательное значение указывает на то, что увеличение одной переменной ведет к снижению тенденции другой. Исследования Эдисона и Добеши показывают, что вейвлеты помогают обнаруживать масштаб и структуру тенденций.

Вейвлеты характеризуются как с помощью параметра растяжения ($a > 0$), так и с помощью параметра сдвига b ($-\infty < b < \infty$). Одномерная вейвлет-функция соответствует двумерной в природе:

$$\psi_{a,b}(x) = \psi\left(\frac{x-b}{a}\right), \quad (2)$$

Для решения задачи используется непрерывное вейвлет-преобразование:

$$W(a,b) = \int_{-\infty}^{\infty} f(x)\psi^*\left(\frac{|x-b|}{a}\right)dx, \quad (3)$$

где $f(x)$ – одномерная функция, а ψ^* – (* означает функцию комплексного сопряжения) анализирующий вейвлет, который еще называют базисным (материнским).

Базисным вейвлетом для данной задачи выбирается МНАТ-вейвлет вида:

$$\psi\left(\frac{|x-b|}{a}\right) = \frac{1}{\sqrt{2\pi}a} \left(2 - \frac{|x-b|^2}{a^2}\right) \exp\left(-\frac{|x-b|^2}{2a^2}\right), \quad (4)$$

Вейвлет-спектр определен в следующем виде:

$$M(a) = \frac{1}{a} \int_{-\infty}^{\infty} W(a,b)^2 db, \quad (5)$$

Кросскорреляционный вейвлет коэффициент, или его еще называют коэффициентом взаимокорреляции, можно записать как:

$$r_W = \frac{\int_{-\infty}^{\infty} W_1(a,b) W_2(a,b) db}{\sqrt{M_1(a)M_2(a)}}, \quad (6)$$

Отношение между корреляционным коэффициентом и корреляционным вейвлет коэффициентом можно представить в виде:

$$r = \frac{\int_{-\infty}^{\infty} r_W(a) \sqrt{M_1(a)M_2(a)} a^{-1} da}{\sqrt{\int_{-\infty}^{\infty} M_1(a) a^{-1} da \int_{-\infty}^{\infty} M_2(a) a^{-1} da}}, \quad (7)$$

В данной задаче понятие вейвлета используется для расчёта коэффициента кросс-корреляции между парами прибрежных и внутренних станций-замеров (табл. 1-3), например в городах Dhahran (восточное побережье), Abha (горный район на юго-западе), Jeddah (западное побережье) и Gizan (юго-западное побережье).

Таблица 1. Кросскорреляционные коэффициенты, рассчитанные для береговых станций

Пары мест замеров	Кросскорреляционный коэффициент				
	Поверхностное давление	Температура воздуха	Относительная влажность	Скорость ветра	Осадки (дожди)
Dhahran-Gizan	0,9400	0,9218	0,3472	0,0748	0,0789
Dhahran-Jeddah	0,8875	0,8729	0,2132	0,1720	-0,0046
Dhahran-Yanbu	0,8833	0,8787	0,1295	0,1103	0,0037
Gizan-Jeddah	0,9524	0,8706	0,0483	0,1096	-0,0001
Gizan-	0,9366	0,8744	0,0483	0,1096	0,0022

Yanbu					
Jeddah- Yanbu	0,9552	0,9430	0,3658	0,3853	0,2123

Таблица 2. Кросскорреляционные коэффициенты, рассчитанные для внутренних станций

Пары мест замеров	Кросскорреляционный коэффициент				
	Поверхностное давление	Температура воздуха	Относительная влажность	Скорость ветра	Осадки (дожди)
Abha–Guryat	0.5994	0.8729	0.4677	0.0062	0.0240
Abha–Hail	0.5155	0.8767	0.4055	0.0640	-0.0010
Abha–Riyadh	0.6602	0.9155	0.5608	0.1602	0.1450
Abha–Turaif	0.5028	0.8590	0.4450	0.0687	-0.0058
Guryat–Hail	0.5750	0.8954	0.5125	0.1448	0.0019
Guryat–Riyadh	0.8069	0.9122	0.5700	0.1471	0.0611
Guryat–Turaif	0.7137	0.9217	0.7344	0.2502	0.1199
Hail–Riyadh	0.6831	0.9088	0.5766	0.0986	0.0269
Hail–Turaif	0.5010	0.8648	0.5844	0.1140	-0.0059
Riyadh–Turaif	0.6676	0.8835	0.6180	0.0851	0.0130

Таблица 3. Кросскорреляционные коэффициенты, рассчитанные для береговых и внутренних станций

Пары мест замеров	Кросскорреляционный коэффициент				
	Поверхностное давление	Температура воздуха	Относительная влажность	Скорость ветра	Осадки (дожди)
Abha–Dhahran	0.6373	0.9047	0.4328	0.0554	0.0381
Abha–Gizan	0.6699	0.9039	0.3319	-0.0435	0.0985
Abha–Jeddah	0.6397	0.8316	-0.0746	0.2316	0.0102
Abha–Yanbu	0.6637	0.8562	0.0404	0.0905	-0.0025
Gurayat–Dhahran	0.8083	0.9172	0.5485	0.0509	0.0264
Gurayat–Gizan	0.8132	0.8923	0.4537	0.2094	0.0054
Gurayat–Jedda	0.7938	0.8990	0.1685	0.2427	0.0370
Gurayat–Yanbu	0.7599	0.9075	0.1195	0.3325	0.0576
Hail–Dhahran	0.7101	0.9178	0.5359	0.0373	0.0117
Hail–Gizan	0.6914	0.9000	0.4516	0.0914	-0.0153
Hail–Jeddah	0.6583	0.8372	0.0873	0.0884	-0.0104

Пары мест замеров	Кросскорреляционный коэффициент				
	Поверхностное давление	Температура воздуха	Относительная влажность	Скорость ветра	Осадки (дожди)
Hail–Yanbu	0.6779	0.8504	0.0168	0.0396	-0.0081
Riyadh–Dhahran	0.9609	0.9767	0.6376	0.3037	0.1765
Riyadh–Gizan	0.9078	0.9154	0.4931	0.0893	0.0600
Riyadh–Jeddah	0.8475	0.8660	0.0624	0.2374	-0.021
Riyadh–Yanbu	0.8337	0.8753	0.0279	0.1168	-0.022
Turaif–Dhahran	0.6647	0.8816	0.5647	-0.0059	0.0171
Turaif–Gizan	0.6616	0.8624	0.4690	0.0965	-0.0099
Turaif–Jeddah	0.6572	0.8419	0.1505	0.0495	0.0497
Turaif–Yanbu	0.6447	0.8592	0.1154	0.0373	0.0248

Результат: В ходе расчета определены максимальные значения коэффициентов корреляции для всех параметров береговых станций среди всех пар. Было найдено так же то, что данные по временным рядам температур сильно коррелируют для всех пар станций, а от общего числа осадков наоборот слабо коррелируют. Временные ряды давления показали хорошую корреляцию по сравнению с относительной влажностью и скоростью ветра. В случае прибрежных станций максимальные коэффициенты соответствуют Jeddah–Yanbu: давление 0.955, температура 0.943, скорость ветра 0.385 и общие осадки 0.212 соответственно. В случае внутренних станций был обнаружен максимальный коэффициент корреляции 0,807 для временных рядов давления пары Guryat–Riyadh, для температуры, относительной влажности и скорости ветра станций Guryat–Turaif 0,922, 0,734, 0,250.

Вывод: Полученные коэффициенты корреляции можно использовать для прогноза появления аномально жарких и очень сухих ветров самум и хамсин, которые вызывают песчаные бури летом и появление сильного зимнего ветра шемаль, который приносит похолодания.

1.2.2 Экономика

Нахождение взаимосвязи между экономическими показателями важно для исследования, оно позволяет анализировать влияние одних показателей на другие: увеличение объема продаж, выручка, цена товара, спрос на товар и др. В ходе анализа строится модель долгосрочного прогнозирования, с помощью которой предсказывается спрос на товар, выручка, расход.

Задача.

Исследование корреляции динамики цены на нефть-сырец и отопительную нефть (рис. 6). [1]

Данные: цена на нефть-сырец, цена на отопительную нефть.

Математические методы: фрактальный анализ.

Случайный процесс, например, такой как изменение валютных рядов, вводит понятие фрактала (объекта, который включает в себя самоподобные структуры) в гауссовский или пуассоновский случайные процессы. Большой объем экспериментальных данных имеет фрактальную структуру, анализ такого вида данных можно произвести посредством проведения фрактального анализа.[13]

Основная численная характеристика для фрактальных структур это размерность D :

$$D = \lim_{\delta \rightarrow 0} \frac{\ln N(\delta)}{\ln(1/\delta)}, \quad (1)$$

«где $N(\delta)$ – минимальное количество шаров радиусов δ , покрывающих это множество. Основанием для введения этого определения является асимптотика для $N(\delta)$, которая в случае множеств определяется выражением»[10]:

$$\left(\frac{1}{\delta}\right)^D = N(\delta). \quad (2)$$

Для того чтобы получить надежное значение для оценки D , посредством использования различных существующих алгоритмов, необходимо обладать большим объемом исходных данных (примерно

$10^3 - 10^4$ значений), поэтому результаты, которые получаются в ходе выполнения расчетов, дают возможность оценить только общие свойства исследуемого ряда на больших масштабах, внутри которых можно наблюдать то, что временной ряд может менять характер своего поведения несколько раз. Чтобы наиболее детально рассмотреть свойства временного ряда требуется, локально, определить размерность D , то есть рассчитать ее в каждой точке ряда. В работе [1] был разработан специальный алгоритм, с помощью которого вычислялись фрактальные характеристики исследуемого ряда и исследованы хаотические временные ряды с помощью индекса фрактальности и др.

Для временных рядов используется индекс Херста H , который связан с фрактальной размерностью соотношением $D = 2 - H$ и является показателем персистентности, способности сохранять определенную тенденцию, временного ряда. Различают три разных режима, которые могут существовать на рынке: $H = 0,5$ поведение цен описывается моделью случайного блуждания, при $H > 0,5$ цены находятся в состоянии тренда, а при $H < 0,5$ цены находятся в состоянии снижения, или частых колебаний в достаточно узком диапазоне цен.



Рис. 6. Совместное представление нормированных значений цен на нефть-сырец (BC - Brent Crude, темная линия) и отопительная (НО - Heating Oil, светлая линия)

В анализе всех фрактальных рядов на первом этапе необходимо рассчитать показатель Херста, который определяется использованием метода нормированного размаха. Расчет данного показателя позволяет определить наличие фрактальных структур в динамике исследуемого ряда, то есть выявить природу происхождения сигнала, которая должна быть фрактальной, чтобы можно было в дальнейшем применять для анализа различные фрактальные методы, если она не является фрактальной, то методы фрактального анализа не подходят для исследования временного ряда. Ниже представлены основные выражения для расчета данного показателя.

Математическое ожидание ряда x за время Δ :

$$X = \frac{1}{\Delta} \sum_{i=1}^{\Delta} X_i, \quad (3)$$

$X(t)$ – накопленное отклонение значения сигнала от среднего значения $(x)_{\Delta}$:

$$X(t, \Delta) = \sum_{u=1}^t \{x(t) - (x)_{\Delta}\}, \quad (4)$$

Размах отклонений будем искать в виде:

$$R(\Delta) = \max X(t, \Delta) - \min X(t, \Delta), \quad (5)$$

С помощью извлечения квадратного корня из дисперсии можно получить в результате стандартное среднеквадратическое отклонение:

$$S(\Delta) = \sqrt{\left(\frac{1}{\Delta} \sum_{t=1}^{\Delta} \{x(t) - (x)_{\Delta}\}^2\right)}. \quad (6)$$

Размах носит название нормированного, если его можно разделить на стандартное среднеквадратическое отклонение, что может позволить применить его к разным системам.

После выполнения всех расчетов, с помощью формул (3)-(6), на основе которых были получены данные, делаются выводы: оценивается среднее значение показателя Херста, определяются пределы отклонения динамики, и

некоторые другие показатели. В данной задаче значение показателя Херста равно 1.13891, это говорит о том, что ряд, представленный на рис. 7, имеет фрактальную природу, это позволяет сделать вывод, что к этой задаче применим фрактальный анализ.

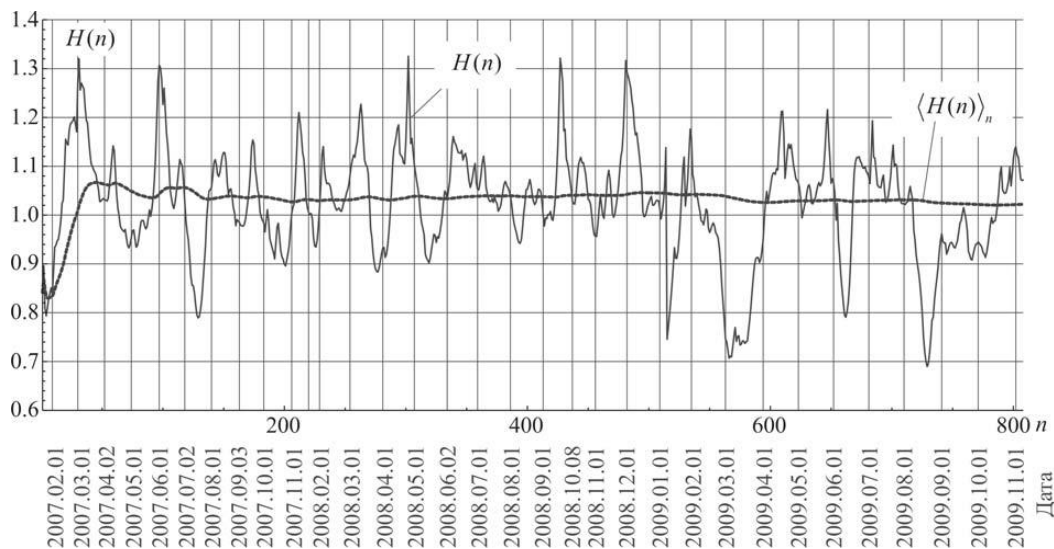


Рис.7. Зависимость динамического значения H от дня наблюдения n , рассчитанная на основании предыдущих 90 значений. $\langle H(n) \rangle_n$ – среднее значение $H(n)$ за период наблюдения n

Далее по методу Грассберга-Прокаччия вычисляется корреляционный интеграл, по которому определяется корреляционная размерность. После чего основываясь на методе фрактального анализа «ложных ближайших соседей» строится график. С помощью использования данного метода можно выявить на основе анализа предыдущих значений цены нефти, устойчивое событие именуемое трендом, который связан как с повышением, так и с понижением данной цены в ближайшем будущем. Причем предполагаемые положительные и отрицательные тренды имеют различные признаки. «Сам метод «ложных ближайших соседей» состоит из следующих шагов:

1. Пусть размерность $m = 1$. Находим для каждой точки $\bar{x}(i)$ временного ряда ближайшего «соседа» $\bar{x}(j)$ в m -мерном пространстве.
2. Вычисляется расстояние между точками $\|\bar{x}(i) - \bar{x}(j)\|$.

3. Находится расстояние между данными точками на следующем шаге $\|\bar{x}(i+1) - \bar{x}(j+1)\|$ и определяется согласно формуле

$$R_i = \frac{\|\bar{x}(i+1) - \bar{x}(j+1)\|}{\|\bar{x}(i) - \bar{x}(j)\|} \quad (7)$$

4. Если $R_i > R_t$, где R_t – это подходящий порог, то точка $\bar{x}(j)$ является ложным ближайшим соседом по отношению к точке $\bar{x}(i)$. В результате производится подсчет количества ложных ближайших соседей P для каждой точки $\bar{x}(i)$.

5. Вычисляется соотношение P/N , где N – общая длина ряда, алгоритм повторяется для $m = m + 1$.

6. Алгоритм продолжается до тех пор, пока частное P/N не станет близким к нулю.

Рекомендуемое значение порога $R_t = 2$ ». [1]

С помощью этого алгоритма, авторы статьи разработали программу, которая позволила рассчитать количество ложных ближайших соседей для разных временных рядов. Результат выполнения данной программы для ряда, который был исследован в данной статье, представлен на рис. 8 в виде графика с зависимостью количества ложных ближайших соседей P , отнесенной к длине ряда N , от размерности пространства вложения m .

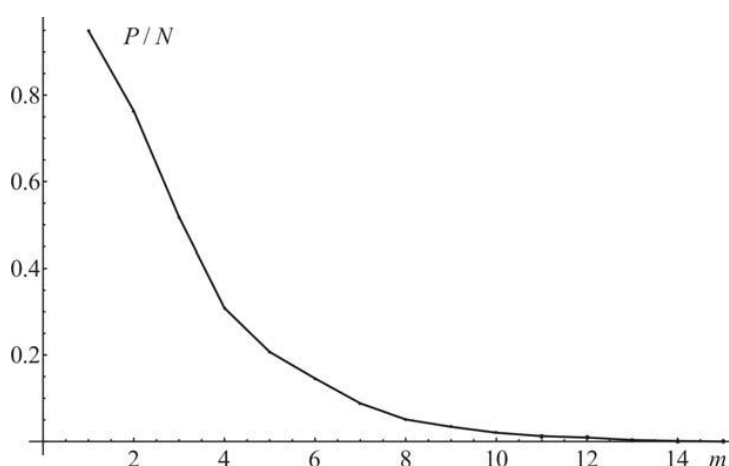


Рис. 8. Зависимость числа «ложных ближайших соседей» P , отнесенного к общей длине ряда N , от размерности пространства вложения m для временного ряда нормированных значений цены на нефть-сырец

Для нахождения закономерности авторы решили ограничить длину ряда, сделать ее не превышающей 120 отсчетов. [1] Связав с предположением, что количество игроков рынка покупки нефти за большой временной период значительно меняется. Опираясь на другие исследования, О.И. Антипов модифицировал метод ложных ближайших соседей: соотношение числа соседей P к общей длине ряда N при переходе из двумерного фазового пространства ($m = 2$) в трехмерное ($m = 3$) определяет “скорость” сворачивания аттрактора (компактное подмножество фазового пространства динамической системы) в пространство вложения. Чем меньше значение, тем быстрее он сворачивается. Это может свидетельствовать о том, что на динамику исследуемого ряда влияет меньшее число факторов (т.е. игроков рынка энергоносителей). Однако, справедливо и обратное. Далее проводилось вычисление отношения P/N по измененному алгоритму. На входе задавалось так называемое окно наблюдений w , которое являлось ничем иным, как количество отсчетов временного ряда, предшествующих дню, для которого вычислялось значение P/N . Из вышесказанного следует, что длина ряда (в случае данного примера это ряд на рис.б), для которого было вычислено соотношение, составила w отсчетов, то есть $N = w$ соответственно. Далее, если отступить от начала исследуемого ряда на длину окна, то можно для каждого последующего дня вычислить соотношение P/w , до конца наблюдаемого периода, это делается для того, чтобы отследить динамику изменения исходного временного ряда.

В процессе исследования общего временного ряда - длина окна w изменялась в пределах от 60 до 120 дней, что соответствует по времени примерно 2-4 месяцам. Наиболее результативным и оптимальным, с точки зрения полученных итогов, оказалось значение окна w равное 90. На выходе получались слишком короткие ряды, поэтому также менялось и значение порога R_t в пределе от 2 до 10. Наиболее оптимальный результат получался при $R_t \in (5; 6)$. Следует также отметить, что при изменении значений w и R_t в указанных выше пределах общая картина результатов изменялась не очень

сильно. На рис.9 показаны результаты вычисления отношения P/w для цены на нефть-сырец, когда значение окна w равнялось 90 и значению порога составляло $R_t = 2e$. На графике жирной линией изображено среднее значение величины P/w для каждого дня ряда за весь исследуемый период, данное значение определяемо по формуле:

$$\langle P/w \rangle_n = \frac{1}{n} \sum_{i=1}^n (P/w)_i. \quad (8)$$

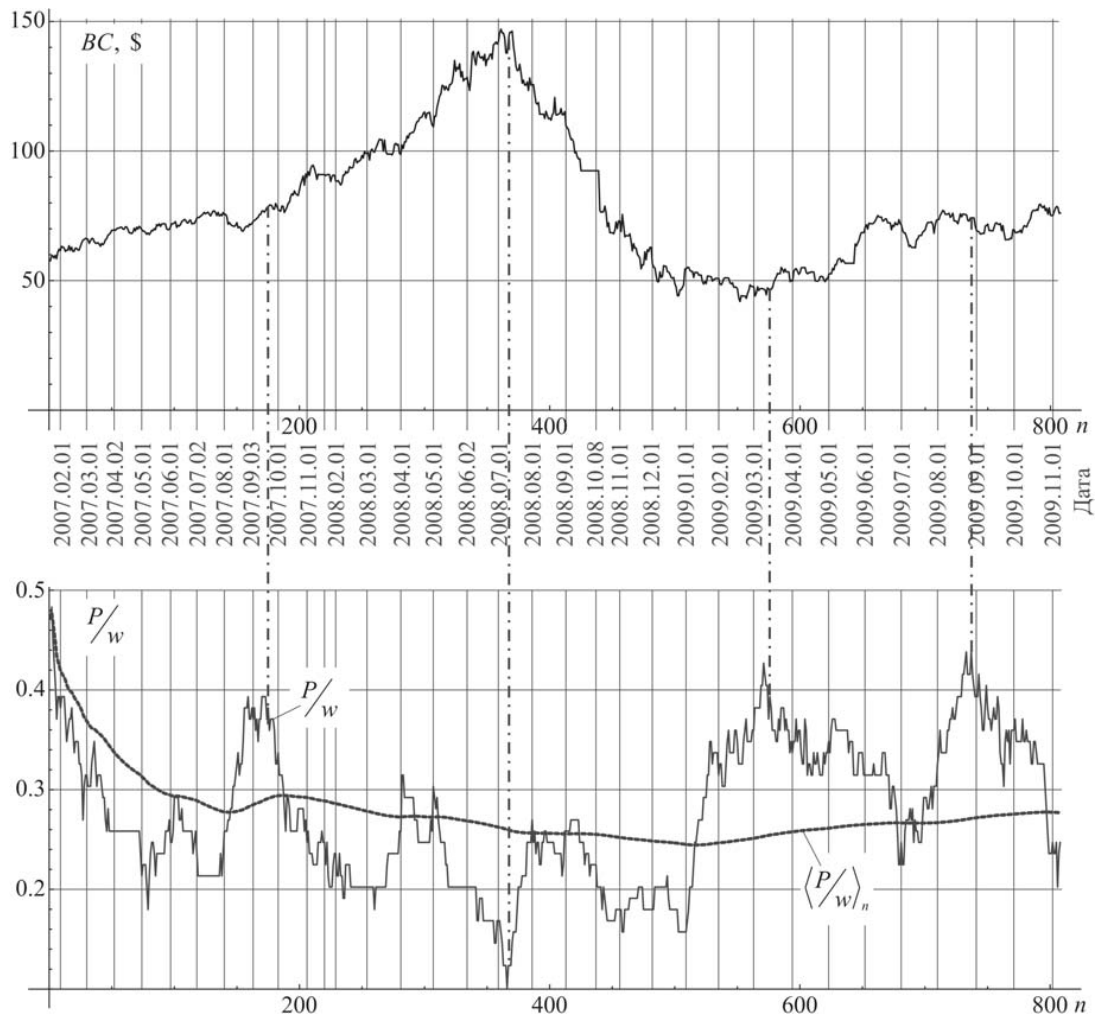


Рис. 9. Сопоставление экстремумов вычисленной зависимости количества ложных ближайших соседей P , отнесенной к длине окна w с соответствующими значениями цены нефти-сырца (начало ряда размером окна в 90 дней не показано)

Обратив внимание на рис. 9 можно увидеть важную закономерность: перед каждым устойчивым по времени событием наблюдается четко выраженный экстремум зависимости P/w от n . Перед снижением цены на

соответствующий вид энергоносителя это минимум зависимости P/w от n , а перед ростом цены наблюдается максимум.

Результат: Из закономерностей, которые были получены в ходе ранее выполненных расчетов, можно сделать вывод о том, что как только достигается точка экстремума минимального значения отношения (резкое отклонение величины P/w от средней, которая на графике обозначена как $\langle P/w \rangle_n$) наблюдается длительное событие, связанное со снижением цены на энергоноситель. Кроме того, следует также добавить, что проходя область резкого отклонения величины от среднего значения, наблюдаемого в сторону роста или увеличения, можно сделать вывод о том, что в ближайшее время будет происходить длительное событие, связанное с увеличением цены на соответствующий вид энергоносителя. Данное явление объясняется следующим предположением: уменьшение числа покупателей энергоносителя говорит и о скором снижении цены на соответствующий товар. То есть, если учесть данное предположение, то можно выдвинуть гипотезу, которая будет верной, о том, что рост числа покупателей (игроков или участников, что то же самое) данного рынка приведет, возможно, к плавному увеличению цены, иногда данный рост числа участников может вызвать и резкий рост цены на товар.

Вывод: Полученные значения можно использовать для прогноза роста или спада цены, а также для определения устойчивости трендов, то есть насколько долго может длиться снижение или рост цены на нефть.

1.2.3 Физика Солнца

Данные полученные из различных наблюдений за солнечными пятнами с одной стороны указывают на наличие некой определенной периодичности их возникновения, а с другой стороны показывают существенную хаотичность данного процесса. Изменяются как периоды солнечных циклов (периодические изменения в солнечной активности), так и амплитуды количеств пятен. Наблюдения могут позволить выявить не только видимые,

но и некоторые скрытые периодичности в разных исследуемых солнечных явлениях, поэтому данные наблюдения за солнечными пятнами много раз подвергались анализу с помощью различных методов, в том числе и с помощью метода спектрального (Фурье) анализа.

Задача.

Выявление закономерностей в динамике числа солнечных пятен.
Прогноз изменения ритмов количества солнечных пятен.

Данные: Количество значений численности солнечных пятен за различные временные промежутки.

Математические методы: спектральный (Фурье) анализ.

В качестве исходных данных для расчета использовались ежедневные и среднегодовые данные количества солнечных пятен, взятые с сервера SIDC (Sunspot Index Data Center, daily и yearly sunspot number). Месячные данные SIDC не были использованы, потому что применяемые при получении их процедуры сглаживания сильно искажали информацию о спектральном составе исходного исследуемого сигнала. С помощью ежедневных данных, используя специальную фильтрацию, предложенную автором данной статьи, соответствующей частоте дискретизации с периодом в 1 месяц, были получены месячные данные, которые в дальнейшем были использованы для расчетов спектров и прогноза. В первых трех вариантах расчетов в качестве исходных данных были использованы, полученные с помощью фильтрации, месячные (monthly) данные, а в четвертом варианте - среднегодовые (yearly). Далее будут представлены результаты выполненных расчетов прогноза и спектров для каждого из четырех различных вариантов.

Основной задачей первого варианта было тестирование алгоритма прогноза. В качестве исходных данных были использованы 411 отсчетов месячных данных количества солнечных пятен, начиная с августа 1943 года. По полученным данным анализировался спектр, далее по нему проводилось

продление данных на 20 лет вперед, то есть за пределы окна наблюдений в 411 месяцев. Результаты данных расчетов представлены на рис.10 ниже. Черная линия показывает исходные данные, синяя и красная линии, это те данные, которые были восстановлены по полученным оценкам спектра в различных интервалах частот.

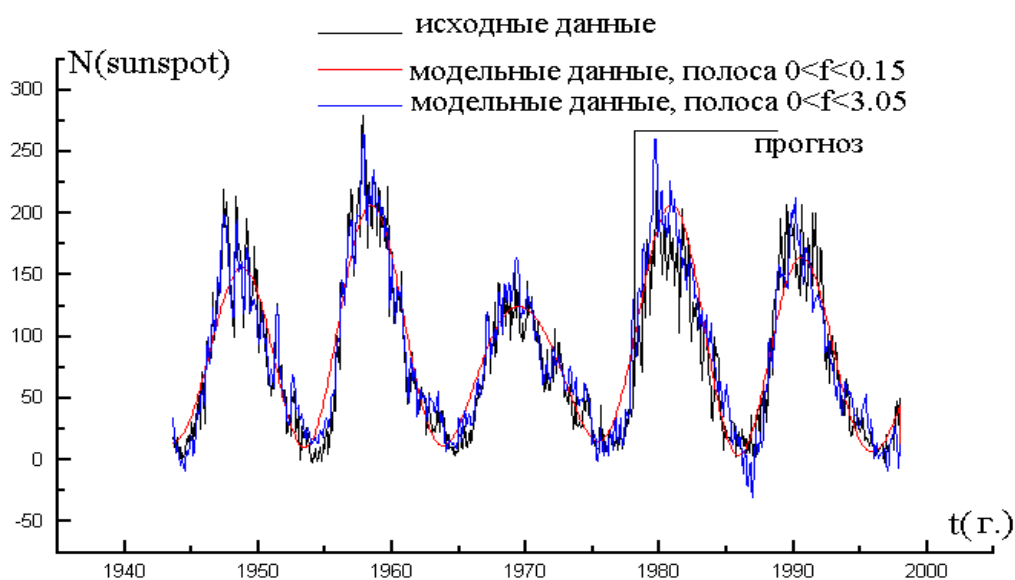


Рис. 10. График зависимости количества солнечных пятен от времени

Второй вариант расчетов почти полностью совпадает с первым, он отличается только выбором положения окна наблюдений: 411 отсчетов данных выбирались в интервале с января 1964 года по ноябрь 1998 года включительно. По этим данным был оценен спектр для второго варианта, далее был восстановлен, после выполнения преобразований, сигнал, как в пределах исследуемого временного окна, так и на 20 лет вперед за его пределы, как это было в первом варианте расчета. Результаты выполненных расчетов изображены на рис.11, где, так же как и в предыдущем рисунке, черная линия является исходным набором данных, а красная и синяя - это модельные данные, которые были получены после оценки спектра.

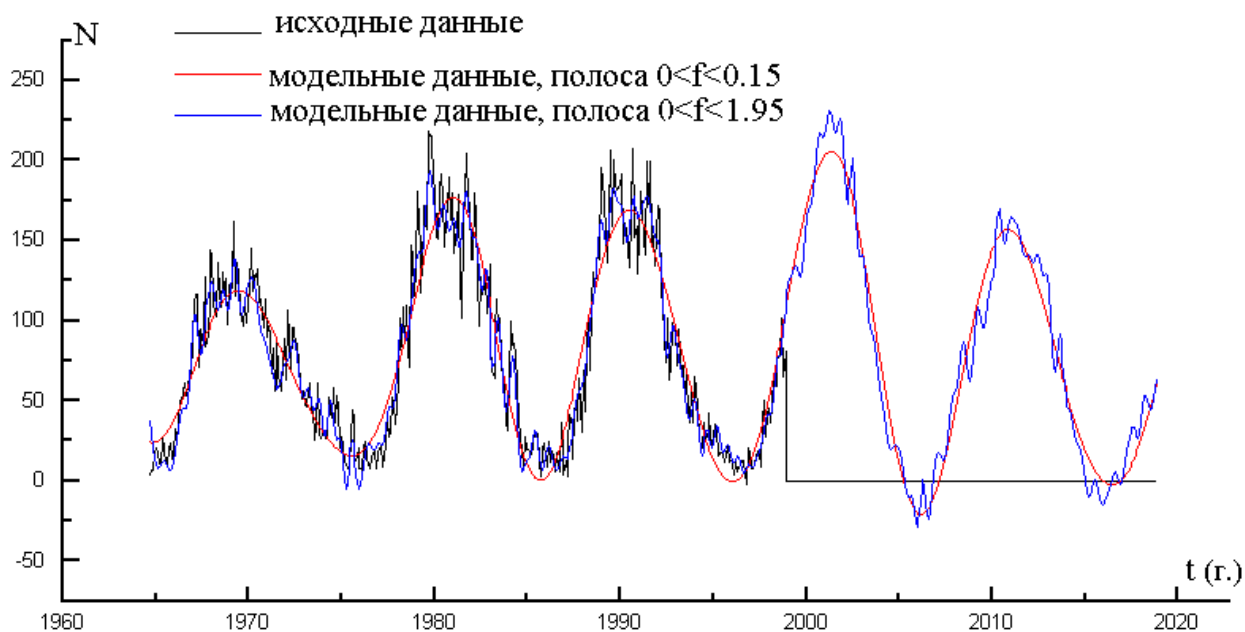


Рис. 11. График зависимости количества солнечных пятен от времени

Отличие третьего варианта от двух предыдущих состоит в длительности исследуемого временного периода исходных данных. Данный период включал в себя 684 отсчета, которые содержали месячные показатели количества солнечных пятен - начиная с января 1942 года по декабрь 1998 года включительно. Этот период прогнозирования охватывает 20 лет. График результатов расчёта показан на рис. 12. На этом же рис. 12 зафиксирован прогноз NASA, помимо исходных (чёрная кривая) и преобразованных (синяя и красная кривая).

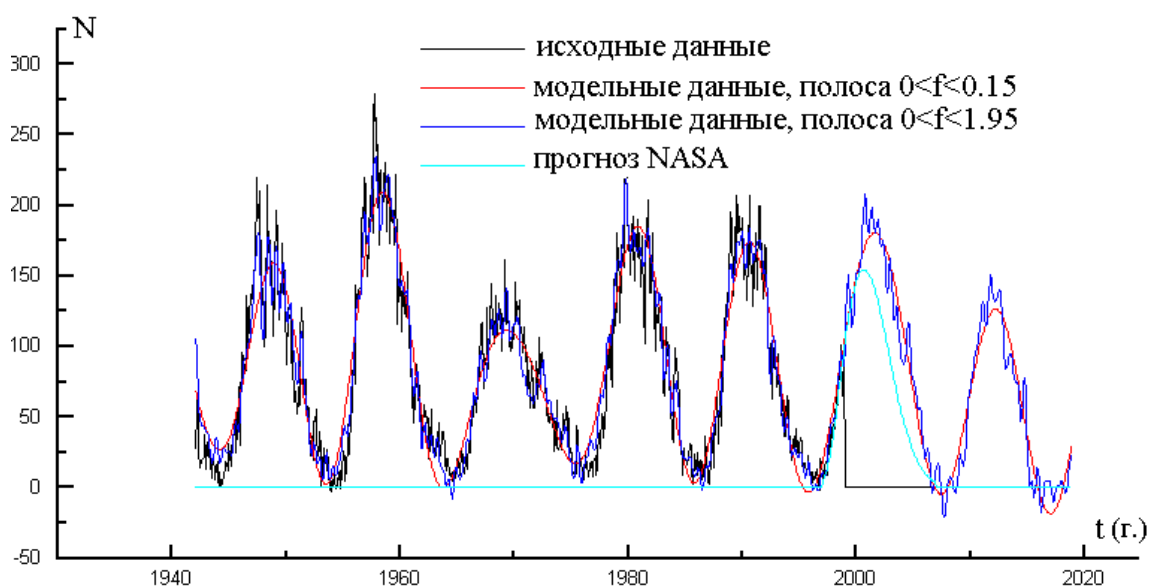


Рис. 12. График зависимости количества солнечных пятен от времени

В следующем четвертом варианте с 1700 года по 1999 год среднегодовые показатели количества солнечных пятен были использованы, как исходные данные. Согласно результатам прогноза по месячным данным за 1999 год была определена численность солнечных пятен. Предварительные расчёты свидетельствовали о том, что часть исходных данных из 299 отсчётов необходимо восполнить как минимум одним отсчётом для того, чтобы осуществился так называемый "захват" последующего 23 цикла. График данных результатов представлен на рис. 13. Обозначения сохраняются, как и на вышепредставленных графиках: чёрной линией обозначены исходные данные, а синей и красной показаны преобразованные.

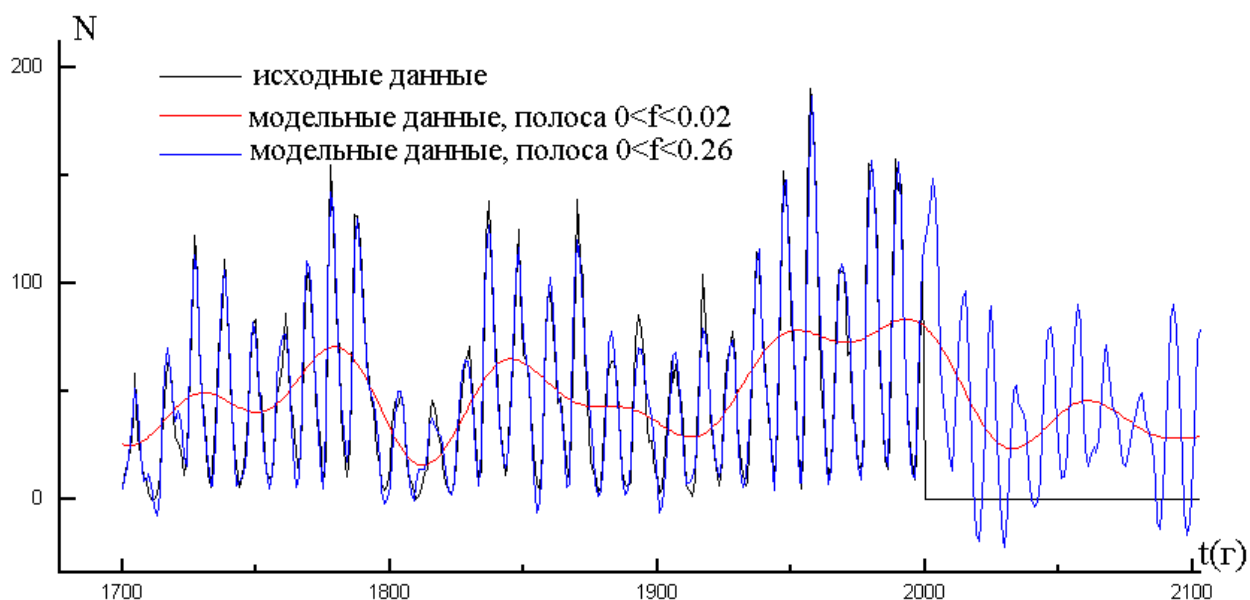


Рис. 13. График зависимости количества солнечных пятен от времени

Нетрудно заметить, что на изображенных выше графиках, преобразованные (синие линии) с достаточной точностью совпадают с исходными данными (черные линии), как в рамках окна наблюдений, так и вне его на пролонгируемом участке. Стандартное среднеквадратическое отклонение в пределах данного окна наблюдений составляло приблизительно 4%. Так же следует заметить, что происходит превышение максимума сглаженных данных (красная линия) для первого прогнозируемого цикла примерно на 10%. Обнаружены различия только на уровне

мелкомасштабных возмущений временной продолжительностью, которая составляет примерно два месяца. Такое различие подтверждается полосой частот, в которой происходит оценка спектра, а ее ширина определяется частотой дискретизации виртуальных отсчетов времени t_k . Количество неопределенных множителей Лагранжа, а также виртуальных отсчетов времени принималось за 300 в первых трех вариантах. Дискретизация виртуальных отсчетов времени $T = (t_M - t_1)/(M - 1)$ в первых двух вариантах принималась равной $\frac{59}{30}$ месяца и $\frac{923}{300}$ месяца в третьем варианте. Количество неопределенных множителей Лагранжа составляло 290 в четвертом варианте, а дискретизация виртуальных отсчетов времени $T = \frac{549}{289}$ года.

В первых трех вариантах положение максимумов и минимумов в циклах рационально устанавливать по красной линии, которая определяет преобразованные показатели сигнала. Эти сигналы были восстановлены по спектру в полосе частот $[0-0.15 \text{ г}^{-1}]$. Далее в таблице 4 рассмотрены ситуации возникновения максимумов t_{\max} и соответствующие им значения N_{\max} для вышеописанных вариантов расчетов.

Таблица 4. Максимальные числа солнечных пятен в различные солнечные циклы

Варианты	I		II		III		IV	
	t_{\max}	N_{\max}	t_{\max}	N_{\max}	t_{\max}	N_{\max}	t_{\max}	N_{\max}
18	11.1948	168	-	-	11.1948	159	1948	148
19	05.1958	218	-	-	06.1958	209	1958	186
20	03.1969	122	05.1969	118	05.1969	111	1969	109
21	11.1980	216 прг	12.1980	177	10.1980	185	1980	157
22	06.1990	177 прг	06.1990	169	08.1990	174	1990	156
23	-	-	04.2001	206 прг	08.2001	180 прг	2003	148 прг
24	-	-	11.2011	157 прг	03.2012	126 прг	2015	96 прг

Анализируя данную таблицу необходимо также учитывать и то, что среднегодовые всегда меньше месячных данных в максимумах циклов.

Рассмотрим показатель первого цикла N_{max} в первом варианте (помечено "прг"), то заметим, что выбранный нами показатель примерно на 16% больше двух последующих. Такой эффект схож с явлением Гиббса в ходе аппроксимации функции частными суммами рядов Фурье. Здесь смена максимумов и минимумов количества солнечных пятен был определен результатом интерференции многих спектральных линий. Так что линии, которые соответствуют показателям частот, превосходящим 0.15 г^{-1} , определяют едва различимые изменения в сравнении с достаточной степенью гладкого цикла в период длительностью 11 лет. Количество же спектральных линий, разрешенных, установить интервалом наблюдений

Далее рассмотрим рис.14. На нем показаны амплитудные спектры солнечных пятен для первых трех вариантов. Интервал наблюдений для первых двух вариантов составлял 411 месяцев. Количество спектральных линий в полосе частот $[0-0.15 \text{ г}^{-1}]$ для данных вариантов было равным 3 и 4, соответственно. Интервал наблюдений для третьего варианта составил 684 месяца, и количество линий возросло до 7.

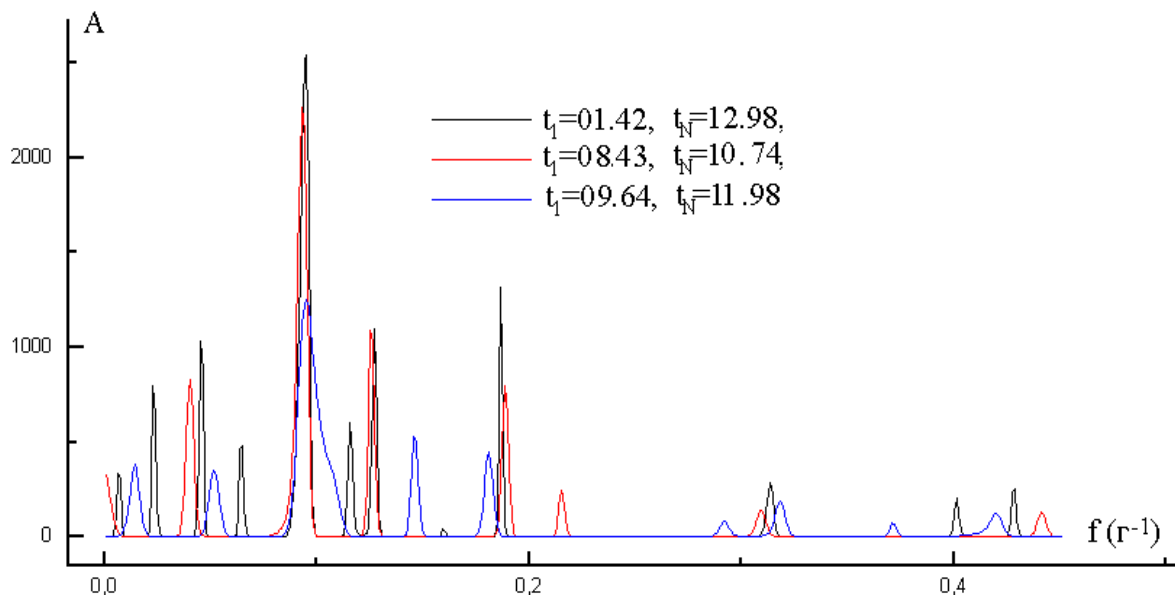


Рис. 14. Амплитудные спектры солнечных пятен для первых трех вариантов

Если допустить то, что явление Гиббса в первых 2-х вариантах наблюдалось в одинаковой степени, то значение $N_{max} = 206$ для 23 цикла варианта II необходимо ограничить до 170-175. Обратив внимание на третий

вариант, в силу большей длительности интервала наблюдений, явление Гиббса должно быть слабее (количество основных линий взаимовлияния выросло с трех до 7), так что значение $N_{\max} = 180$ необходимо понизить лишь на величину приблизительно равную 10, полученный результат согласуется с прогнозом 23 цикла второго варианта.

Также заметим, что от увеличения промежутка наблюдения следует смещение положения максимума, а также и ослабление явления Гиббса, что является очевидным, так как происходят изменения числа, положения и форм профилей линий. А это в свою очередь изменяет интерференционную картину. Происходит изменение амплитуды и положение на временной оси циклов. Максимальная длина интервала наблюдений для солнечных пятен на данный момент занимает 299 лет.

Возможна та ситуация, при которой положение интерференционных максимумов, найденное по спектральным оценкам для среднегодовых значений количества солнечных пятен, будет различным с соответствующими параметрами, найденными по месячным данным.

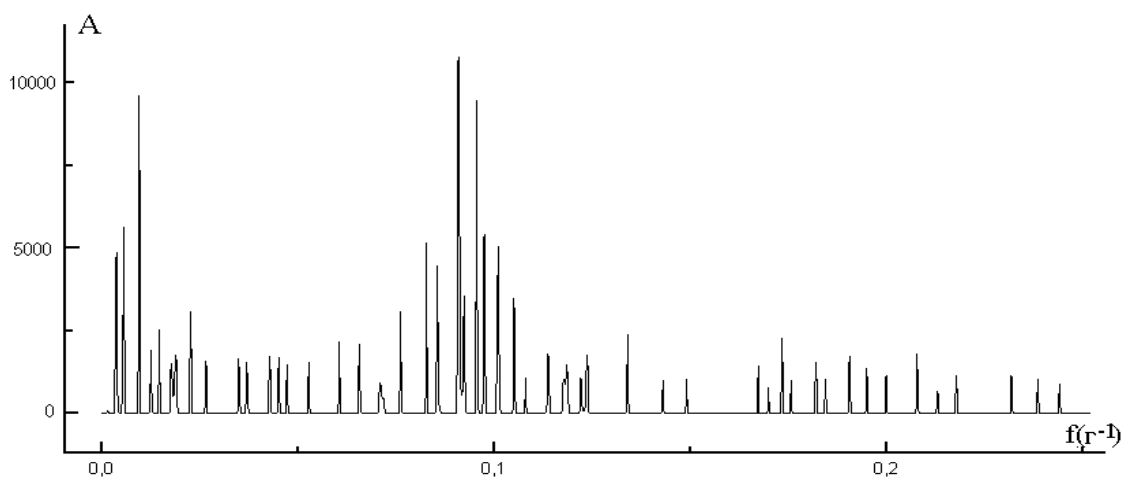


Рис. 15. Амплитудный спектр среднегодовых данных

Увеличение продолжительности периода наблюдений повлекло за собой обогащение данного амплитудного спектра, который базировался на среднегодовые данные (Рис.15), а количество линий с 7 изменилось до 34 в полосе частот $[0-0.15 \text{ г}^{-1}]$. А толщина линий значительно уменьшилась,

целесообразно амплитуда увеличилась. Довольно ясной стала линия шириной 0.0004 г^{-1} , которая соответствует периоду колебаний 106 лет. Данная спектральная картина распадается на мультиплетные полосы. «Мультиплет – это совокупность частиц системы, обладающей сходными свойствами». Мультиплет из 9 линий соответствует 11-летнему солнечному циклу.

Немало интересным является случай с низкочастотными составляющими амплитудного спектра с периодом от 50 лет и больше с полосой частот $[0-0.02 \text{ г}^{-1}]$. Нетрудно понять по рис. 13, что на данный момент мы располагаемся в начале участка, восстановленного по данной полосе временного хода количества солнечных пятен (красная линия). Аналогичная картина была в пределах 1800 года, в это время численность солнечных пятен в максимумах рядом стоящих циклов уменьшилась со 132 до 47, а период между данными максимумами составлял 17 лет, в то время как между минимумами 14 лет. Соответствующий прогнозу по месячным данным в третьем варианте максимум 23 цикла обязан отстоять на одиннадцать лет от максимума предыдущего цикла. В свою очередь прогноз по среднегодовым данным данный период увеличивает до 13 лет, на что прогнозируемая длительность 23 цикла станет 14 лет (от минимума и до минимума). Резкий скачок в связи с уменьшением амплитуды цикла произойдет только в 24 цикле, а среднегодовое количество солнечных пятен понизится до 93.

Из сказанного ранее, истинность прогноза можно определить на основании вида структурных компонентов, применяющихся при аппроксимации сигнала. На рис. 16 можно увидеть структурные компоненты первых трех вариантов, и заметить, как четко определяется структура одиннадцатилетнего цикла. Но данная структура «зашумлена» едва различимыми вариациями, которые в свою очередь провоцируют замывание цикличности с ростом периода. Можно увидеть, что происходит совпадение

структурных компонентов наших трех вариантов на начальном участке в течение примерно 25 лет.

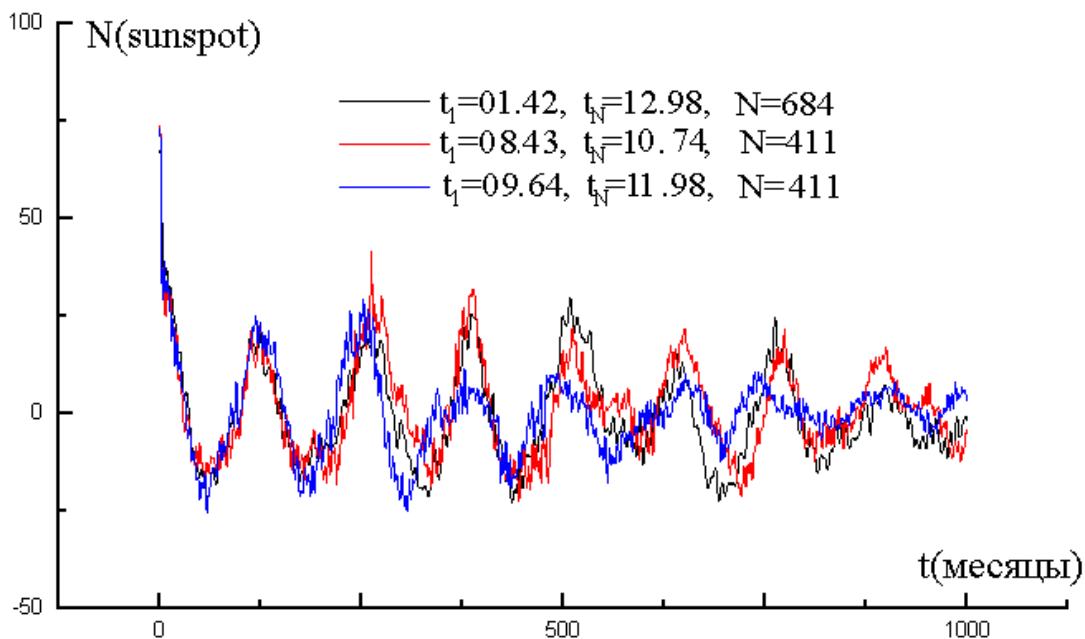


Рис. 16. Структурные элементы для первых трех вариантов

На следующем рисунке изображены структурные компоненты (рис.17), которые были построены на базе усредненных годовичных данных количества солнечных пятен. Кроме того можно заметить, что одиннадцатилетняя цикличность, представляющая из себя комбинацию интерференции линий и низкочастотных составляющих спектра.

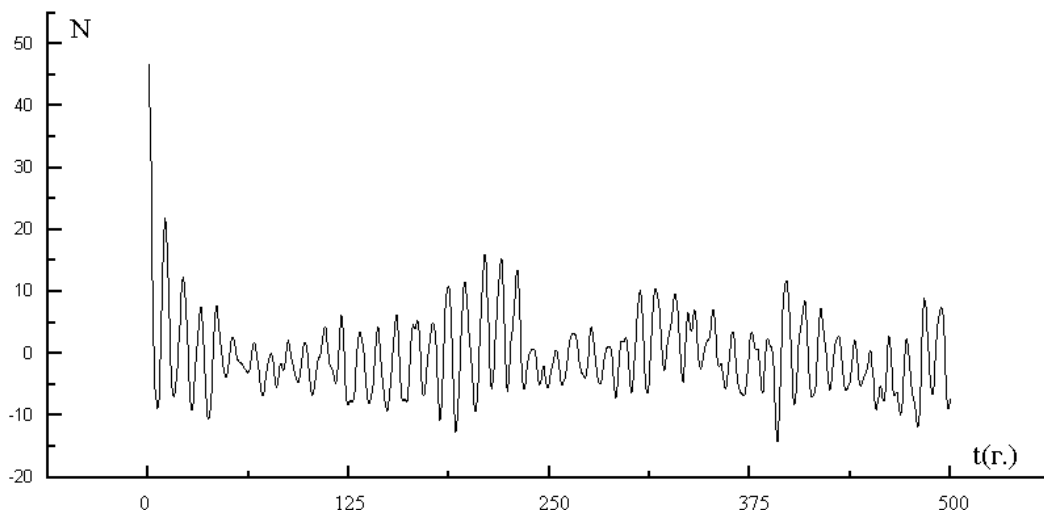


Рис. 17. Структурные элементы по усредненным годовичным данным численности солнечных пятен

Результат: Полученные результаты дают возможность выдвинуть предположение о том, что динамика количества солнечных пятен на текущий

момент испытывает существенные изменения: в течение следующих 30 лет будет прослеживаться сильное плавное понижение компоненты с низкой частотой количества пятен (периоды более 50 лет); в результате взаимодействия линий мультиплета одиннадцатилетнего цикла ожидается заметное увеличение длительности 23 цикла до 14 лет, при этом интенсивность цикла в максимуме будет сравнима с интенсивностью предыдущего 22-го цикла, или немного ниже его, но интенсивность последующего 24 цикла может резко упасть (почти в два раза).

Вывод: Полученные результаты корреляции данных можно использовать для того, чтобы спрогнозировать изменения ритмов численности солнечных пятен и определить длительность солнечного цикла.

ГЛАВА 2. ОПИСАНИЕ МЕТОДОВ РЕШЕНИЯ ЗАДАЧ КОРРЕЛЯЦИИ

2.1 Спектральный анализ

Для начала следует дать вводные определения спектра. Спектр в физике - это распределение значений физической величины (энергии) по частотам или длинам волн. «Спектр сигнала - это результат разложения сигнала на более простые составляющие в базисе ортогональных функций. В качестве разложения обычно используется преобразование Фурье». [8]

Спектральный анализ – метод обработки результатов экспериментов временных рядов, с помощью которого можно, для исследуемого сигнала, охарактеризовать частотный состав. Спектральный анализ базируется на теории равносильности представления функций в таких областях, как частотная и временная, с помощью использования Фурье преобразования. Данная деталь позволяет получать эффективные методы для получения численных характеристик исходного сигнала в частотной области тогда, когда во временной области это сделать практически невозможно, то есть если наблюдается равномерная динамика, без явных событий, трендов.

В теории цифровой обработки сигналов главный, наиболее используемый метод это спектральное преобразование, которое в простейшем виде выглядит, как разложение исследуемого периодического сигнала $f(x)$ в ряд Фурье:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(nx) + b_n \sin(nx)]. \quad (1)$$

Коэффициенты разложения получаются из следующих формул:

$$a_0 = \frac{2}{\pi} \int_{-\pi}^{\pi} f(x) dx; \quad (2)$$

$$a_n = \frac{2}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx; \quad (3)$$

$$b_n = \frac{2}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) dx; \quad (4)$$

Если перейти от использования различных тригонометрических функций к формулам Эйлера с комплексными экспонентами, то формула (1) примет вид:

$$f(x) = \sum_{n=-\infty}^{\infty} C_n e^{-inx}; \quad (5)$$

$$C_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx; \quad (6)$$

В формулах (1)-(6) $n = (1, 2, 3, \dots)$. Числа C_n носят название коэффициентов Фурье сигнала $f(x)$. Рассмотрение случайного сигнала $f(x)$ в виде комбинации, и притом линейной, гармонических функций именуется дискретным спектральным разложением. Для определения свойств линейной динамической системы (исследуемого сигнала) используется дискретный спектр.

С помощью использования формул (5) и (6) можно осуществить взаимно однозначное преобразование исходного сигнала $f(x)$ в набор коэффициентов Фурье C_n и точно таким же способом можно выполнить обратную операцию. Можно сделать вывод о том, что дискретный спектр сигнала $f(x)$ содержит всю информацию о его поведении. Считается, что в частотной области функция $f(x)$ представлена своими коэффициентами Фурье C_n , а сам исходный сигнал $f(x)$ показывает эту же информацию во временной области. Основной идеей спектрального анализа является переход от рассмотрения сигнала к функции времени $f(x)$ к рассмотрению его спектра C_n .

Формула (5) демонстрирует представление исходного сигнала $f(x)$ в виде комплексного ряда Фурье. Чтобы рассчитать коэффициенты Фурье C_n и получить данное представление, можно воспользоваться дискретным преобразованием Фурье (ДПФ) или быстрым преобразованием Фурье (БПФ),

в целях экономии времени и для сокращения количества выполняемых математических операций. Для дискретного преобразования Фурье коэффициенты Фурье рассчитываются из скалярного произведения вектора, который соответствует системе векторов e^{-inx} , при том ортонормированных, и вектора, который соответствует исследуемому сигналу $f(x)$. Данное произведение выражается с помощью формулы:

$$C_n = \frac{1}{N} \sum_{k=0}^{N-1} x_k e^{-i\left(\frac{2\pi}{N}\right)kn}, (n = 0, 1, 2, \dots, N - 1) \quad (7)$$

где x_k , ($k= 1, \dots, N-1$) – приближенные значения исследуемого сигнала $f(x)$, после выполнения преобразований Фурье. Полученное выражение называется дискретным преобразованием Фурье. Соотношение, которое выражает каждый компонент вектора $f(x)$ из выражения вида:

$$x_k = \sum_{n=0}^{N-1} C_n e^{i\left(\frac{2\pi}{N}\right)kn}, (n = 0, 1, 2, \dots, N - 1) \quad (8)$$

Формула (8) носит название обратного дискретного преобразования Фурье.

На практике дискретное преобразование Фурье редко используют для того, чтобы перевести исследуемый сигнал из временной области в частотную, в связи с тем, что требуются большие затраты времени на вычисления, чаще всего необходимо укладываться в короткие сроки для выполнения работы. Используя быстрое преобразование Фурье требуемый результат разложения можно получить гораздо быстрее, миновав большой объем вычислений и преобразований. Поэтому быстрое преобразование Фурье является основным для проведения спектрального анализа. Быстрое преобразование Фурье реализовано в виде пакета в среде Matlab.

Спектральный анализ базируется на использовании преобразований Фурье, что позволяет разложить исходный временной ряд на составляющие: спектры плотности, которые используются в расчете коэффициента

когерентности (согласованности, скоррелированности нескольких колебательных волновых процессов во времени). Кроме того, данный метод способствует определению наиболее эффективных методов получения характеристик изучаемого сигнала в частотной области тогда, когда во временной обнаружить изменения проблемно вследствие равномерной динамики исследуемого сигнала. Основным плюсом данного метода можно отметить простоту его реализации, так как он использует преобразования Фурье, которые подробно математически изучены и реализованы в различных программных обеспечениях, что позволяет существенно уменьшить объем занимаемого на вычисления времени. Но это влечет за собой существенный недостаток, данный метод не может определить быструю смену тенденций, а из-за этого возникает проблема с точностью решения. Спектральный анализ используется для решения прикладных задач: в устройстве вычислительных приборов, в геологии для исследования температурных волн, и в других областях.

2.2 Вейвлет-анализ

Разные исследователи и ученые называют вейвлет-анализ своеобразным «математическим микроскопом», который позволяет детально изучить состав и структуру неоднородного сигнала или функции.

Слово «wavelet» в переводе с английского означает «маленькая волна», или «всплеск», данный термин начал использоваться в отечественной литературе связанной с вейвлет-анализом, когда он стал широко известен, его стали применять в различных классических областях наук. Например, в ходе обработки и анализа числовых рядов разных направлений, к примеру: физических, геофизических и других подобных результатов экспериментов. На текущий момент, данный метод нашел свое широкое использование в синтезе сигналов с различной природой возникновения и структурой (например, речи). Кроме того, его также используют для анализа медицинских кардиограмм, для проведения диагностики острых лейкозов,

кроме того, вейвлет-анализ используется для решения задач обработки, сжатия, хранения и передачи больших объёмов информации, в том числе для сжатия аналогового и цифрового изображения (JPEG-2000) или видео (MPEG-4).[14]

В вейвлет-анализе огромный выбор базовых функций для выполнения преобразований, кроме того он еще и постоянно растет. На рис. 18 приведены часто используемые виды основных классических вейвлетов, с

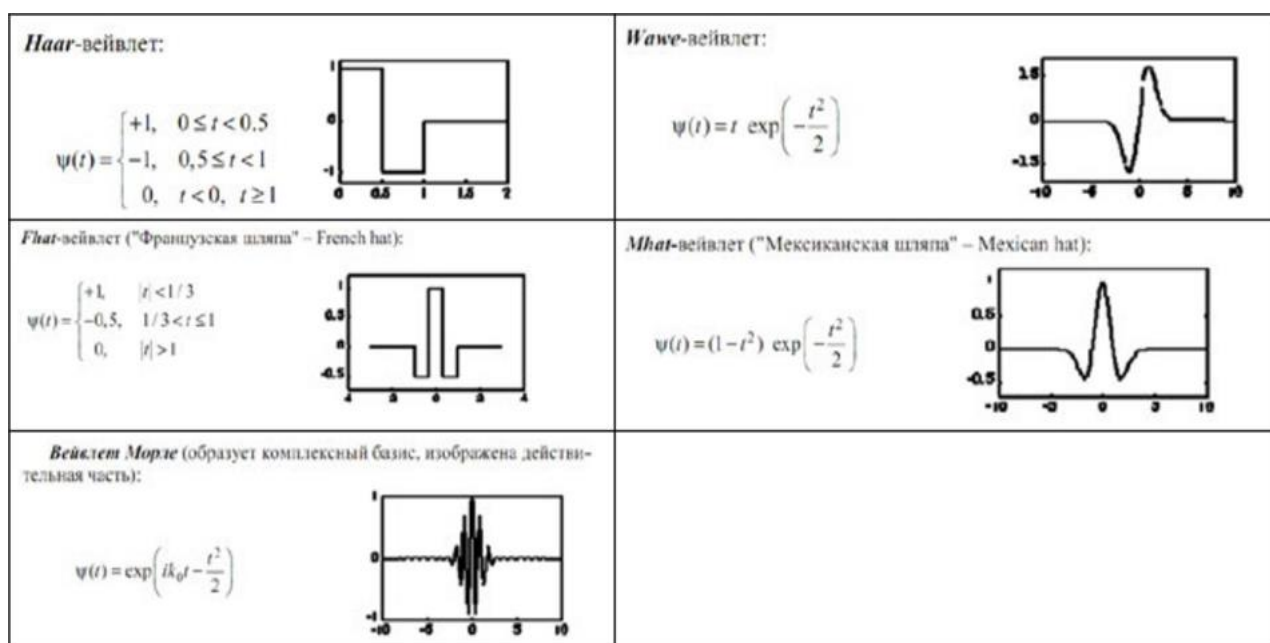


Рис. 18. Виды основных классических вейвлетов

которых начиналось развитие данного метода. Применение данных вейвлетов в настоящее время сильно упростили благодаря добавлению вейвлет-преобразований в приложения, в которых большая часть методов зашита в математических пакеты, например в среде MATLAB. [12]

Данный метод используют в тех случаях, когда результат анализа некоторого сигнала, в котором содержатся характерные частоты, должен содержать сведения и об определенных локальных координатах, при которых эти частоты влияют на сигнал. Следовательно, анализ и обработка нестационарных (по t) или неоднородных сигналов разных типов является основной задачей вейвлет анализа.

Далее рассмотрим определение, лежащее в основе вейвлет-анализа. Вейвлет-преобразованием функции $f(t) \in L^2(\mathbb{R})$ называется выражение:

$$W(a, b) = \frac{1}{|a|^{1/2}} \int_{-\infty}^{+\infty} f(t) \psi^* \left(\frac{t-b}{a} \right) dt, \quad (8)$$

где $a, b \in \mathbb{R}, a \neq 0$.

Входящая в выражение (8) функция $\psi(t)$ называется вейвлетом (анализирующим, базисным или материнским). В отличие от спектрального-анализа ядро преобразования в вейвлет-анализе не зафиксировано, поэтому вейвлет-преобразования для одной и той же функции можно получить с помощью использования разных базисных вейвлетов. Символ * обозначает комплексное сопряжение. Параметр a определяет размер вейвлета, то есть его масштаб. Параметр b задает временной интервал для исследования вейвлета и называется сдвигом.

Обратное интегральное вейвлет-преобразование задается выражением:

$$f(t) = C_{\psi}^{-1} \iint_{-\infty}^{+\infty} W(a, b) \psi \left(\frac{t-b}{a} \right) da db, \quad (9)$$

где C_{ψ} - нормирующий коэффициент:

$$C_{\psi} = \int_{-\infty}^{+\infty} \frac{|\psi|^2}{f(t)} df(t) < \infty. \quad (10)$$

Как было сказано немного ранее, ядро вейвлет-преобразования не зафиксировано, поэтому его можно выбирать на свое усмотрение из многочисленных семейств вейвлетообразующих функций.

Вейвлет-анализ предлагает большой выбор возможностей для исследования сигналов. Основная цель анализа заключается в исследовании сложных сигналов, например идеальной ударной волны, очистки сигнала от импульсных помех. В данный момент в прикладных задачах вейвлеты слабо применимы, они постепенно вводятся в теоретические задачи, кроме того тестируется их применение совместно с использованием Фурье-анализа. Данный метод имеет явное преимущество по сравнению с другими методами

анализа сигнала, так как он позволяет наиболее детально проанализировать сигнал по компонентам.

2.3 Фрактальный анализ

«Фрактал – это множество, которое обладает свойством самоподобия (объект в точности или приближенно совпадает с частью самого себя) В математике под фракталами понимают множества точек в евклидовом пространстве, имеющие дробную метрическую размерность, отличную от топологической». [13]

Фрактальные временные ряды появляются чаще всего, при измерениях и наблюдениях за различными естественными процессами, например такими, как: солнечная активность, уровень разливов рек, шумы электронных приборов, геофизическая и геомагнитная активности земной коры и других объектов, физиологические характеристики организма человека и т.д.

Наиболее важной областью применения фрактального анализа, в рамках данной диссертационной работы, является анализ временных рядов (изменения различной количественной величины во времени) и нестационарных сигналов (сигналы с различными амплитудными колебаниями). Чаще всего, информация о поведении различных сложных динамических систем получается в виде экспериментальных данных. На текущий момент хорошо известно, что фракталами могут являться графики чрезвычайно различных процессов, как стохастических (например, броуновское движение), так и определенные (например, выполнение решения уравнения логистики при определенных фиксированных значениях какого-либо параметра).

Фрактальный анализ позволяет определить наличие хаотической компоненты и определить её долю в физическом процессе.

Основной целью фрактального метода является исследование хаотичных временных рядов, поиск фрактальности (повторяющегося изменения), что не наблюдается во временных рядах дебитов закачки и

добычи, так как они не периодические. Для точной работы метода фрактального анализа в малых масштабах необходимо большое количество замеров ($10^3 - 10^4$). Хаотичные измерения присущи временным рядам в экономике и изучаются в науке – экономической физике, а исследование таких рядов не является необходимым для решения задачи взаимовлияния скважин.

Вывод:

Спектральный анализ относительно прост в использовании и быстр в вычислениях, именно его часто применяют для решения прикладных задач, но у него низкая точность в выполнении анализа сигналов. Данный метод может анализировать равномерную динамику показателей, но не может определить быструю смену тенденций, резкие скачки, и из-за этого возникают неточности, большая погрешность в вычислениях, что влечет за собой ошибочную оценку корреляции в нестационарных сигналах, которые имеют различную амплитуду, резкие изменения трендов в разные моменты времени.

Фрактальный анализ применим к любым временным интервалам. Его прогноз зависит от выбора размера исследуемого временного периода, чем он меньше, тем меньше точность прогноза, поэтому необходим большой набор исходных данных для выполнения анализа сигнала. Кроме того, фрактальный анализ хорошо подходит для решения экономических задач, но не для анализа взаимовлияния потому, что фрактал формируется после появления сильного скачка или резкого спада, то есть ему необходимо наличие хаотической компоненты для проведения анализа, что не наблюдается в поставленной задаче.

Вейвлет-анализ используется, для детального анализа сигналов, то есть при выделении в локализованных сигналах разномасштабных процессов можно рассматривать только те масштабные уровни разложения, которые представляют интерес. Вейвлет-анализ, в отличие от спектрального, имеет много разнообразных базовых функций, свойства которых ориентированы на

решение различных задач, поэтому этот метод обладает меньшей погрешностью при вычислении. Обладая небольшим объемом данных, можно выполнить вейвлет-анализ, по сравнению с фрактальным анализом, это существенный плюс. Главным недостатком этого метода является выбор базисного вейвлета, функции, с помощью которой происходит преобразование исходного сигнала. Базисных вейвлетов большое множество, и каждый выбирается для решения конкретной поставленной задачи, что делает процесс поиска необходимого вейвлета, а в следствии и решение задачи, трудоемким процессом.

Взвешивая все достоинства и недостатки выше представленных методов, наиболее подходящим для решения поставленной задачи является метод вейвлет-анализа.

ГЛАВА 3. МЕТОД ОПРЕДЕЛЕНИЯ УРОВНЯ ВЗАИМОВЛИЯНИЯ НАГНЕТАТЕЛЬНЫХ И ДОБЫВАЮЩИХ СКВАЖИН С ПРИМЕНЕНИЕМ ВЕЙВЛЕТ- АНАЛИЗА

Реализация метода вейвлет-анализа в среде разработки MATLAB

Для решения задачи используется среда разработки MATLAB, так как данное программное обеспечение обладает большой математической базой, огромным выбором настраиваемых пакетов с различным набором функций, кроме того, в MATLAB есть раздел посвященный вейвлет-анализу и вейлет-преобразованиям.

На рис. 19 представлена последовательность действий для реализации расчета взаимовлияния методом вейвлет-анализа.



Рис.19. Блок-схема расчета коэффициентов взаимовлияния методом вейвлет-анализа в программной среде Matlab

Далее рассмотрим метод более подробно, разбив его на несколько этапов:

1) *Входные данные:* Входными данными в программе будем задавать временные ряды приемистости и дебита жидкости.

$\{X_t\}$ – данные с добывающей скважины, $\{Y_t\}$ – данные с нагнетательной скважины.

2) *Преобразование исходных данных:* На данном этапе задается базисный вейвлет, выбирается порядок, до которого будет происходить разложение исходных данных на составляющие для дальнейшего анализа.

Для преобразования исходных данных в среде MATLAB используется специальная функция `wavedec`, которая выполняет многоуровневый одномерный вейвлет-анализ с использованием либо специфического вейвлета, например `db4` - вейвлет Добеши четвертого порядка, в котором порядок отвечает за структуру сигнала (чем выше порядок, тем более сглаженный сигнал можно исследовать, для каждого вейвлета количество порядков индивидуально и ограничено), либо конкретных фильтров декомпозиции вейвлетов (высокий и низкий проходы, то есть анализируются либо показатели с высокой частотой, либо с низкой).

Пакет `wavedec` включает в себя несколько параметров, которые необходимы для разложения исходного сигнала. Данные параметры имеют вид:

$$1. \varphi_{j,n}(x) = \sqrt{2^j} \varphi(2^j x - n),$$

где φ – это масштабирующая функция, число j характеризует уровень разрешения.

$$2. P_j(f) = P_{j-1}(f) + P_{j-1}^W(f),$$

где $P_j(f)$ является приближением j -го уровня, первое слагаемое в формуле есть более грубое приближение в пространстве V_{j-1} , которое было получено путем разложения исходного пространства V_j на сумму ортогональных подпространств, второе слагаемое отражает детали, которыми $P_{j-1}(f)$ отличается от $P_j(f)$.

В базисах пространств V_{j-1} и W_{j-1} имеем:

$$3. P_j(f) = \sum_{k \in Z} a_{j-1,k} \varphi_{j-1,k}(x) + \sum_{k \in Z} d_{j-1,k} \psi_{j-1,k}(x),$$

где $a_{j-1,k}$ – коэффициенты аппроксимации $(j-1)$ -го уровня разрешения,

$$a_{j-1,k} = (f, \varphi_{j-1,k}) = \int f(x) \overline{\varphi_{j-1,k}(x)} dx,$$

и $d_{j-1,k}$ – детализующие коэффициенты,

$$d_{j-1,k} = (f, \psi_{j-1,k}) = \int f(x) \overline{\psi_{j-1,k}(x)} dx,$$

в формулах выше $\psi(x) = \sum_m c_n \varphi_{1,n}$ – вейвлет, а c – константа

Полученные наборы коэффициентов обозначаются символами cA_1 и cD_1 .

Особенностью коэффициентов аппроксимации cA_1 является то, что их можно повторно разложить на еще более сглаженную часть $P_{j-2}(f)$ и детализующую часть $P_{j-2}^W(f)$:

$$P_j(f) = P_{j-2}(f) + P_{j-2}^W(f) + P_{j-1}^W(f),$$

$$P_j(f) = \sum_{k \in Z} a_{j-2,k} \varphi_{j-2,k}(x) + \sum_{k \in Z} d_{j-2,k} \psi_{j-2,k}(x) + \sum_{k \in Z} d_{j-1,k} \psi_{j-1,k}(x).$$

Таким образом, видно, что пакет `wavedec` позволяет разложить исходный сигнал до интересующего приближения и получить соответствующий набор коэффициентов отвечающих за детализацию сигнала.

Описание функции `wavedec`:

$$[C, L] = \text{wavedec}(s, N, 'wname')$$

Исходный набор параметров:

- s – одномерный сигнал;
- N – уровень разложения;
- 'wname' – название используемого вейвлета;
- C – вектор, который включает в себя наш набор полученных коэффициентов;
- L – вектор, который учитывает длины коэффициентов вектора C .

3) *Нахождение детализующих коэффициентов*: на данном шаге используется пакет `detcoef`, который находит детализующие коэффициенты одномерного сигнала задаваемого нами уровня, используя полученные ранее векторы многоуровневого разложения $[C, L]$. Формулы для нахождения детализующих коэффициентов были описаны в предыдущем пункте.

Описание функции `detcoef`:

$$[s_1, s_2, \dots, s_n] = \text{detcoef}(C, L, 1, 2, \dots, n),$$

Исходный набор параметров:

- s_n – обозначает набор компонентов соответствующего уровня разложения, порядку которого отвечает индекс и цифры в скобках из правого выражения;

- C – вектор, который включает в себя набор полученных коэффициентов;

- L – вектор, который учитывает длины коэффициентов вектора C . [12]

4) *Вычисление значения ковариации и среднего квадратического отклонения:* данные числовые характеристики рассчитаны с помощью специальных пакетов cov и std.

4.1) *Расчет ковариации:* для расчета ковариации потребуется знать математические ожидания, то есть средние значения каждой из выборок.

Математическое ожидание находится по формуле:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i,$$

где n – объем выборки, а X_i – элемент выборки, индекс обозначает номер элемента.

Зная математическое ожидание каждой из выборок можно найти значение ковариации по формуле:

$$\text{cov}(X, Y) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$$

4.2) *Расчет среднего квадратического отклонения:* определив математическое ожидание, можно рассчитать среднее квадратическое отклонение:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}$$

5) *Расчет дисперсии:* дисперсией является квадрат среднего квадратического отклонения:

$$D = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

б) *Коэффициент корреляции данных*: выходными данными для задачи является расчет коэффициента корреляции преобразованных данных. Формула, которая используется для расчета коэффициента корреляции, имеет вид:

$$r = \frac{\text{cov}(X, Y)}{\sqrt{D_X \cdot D_Y}}$$

ГЛАВА 4. ТЕСТИРОВАНИЕ АЛГОРИТМА НА ИДЕАЛЬНОМ ПРИМЕРЕ ДАННЫХ

Для расчета коэффициента корреляции было выбрано четыре примера, для каждого примера использовался набор случайно заданных данных, общим количеством 40 замеров:

- 1) *Пример со 100% взаимовлиянием.*

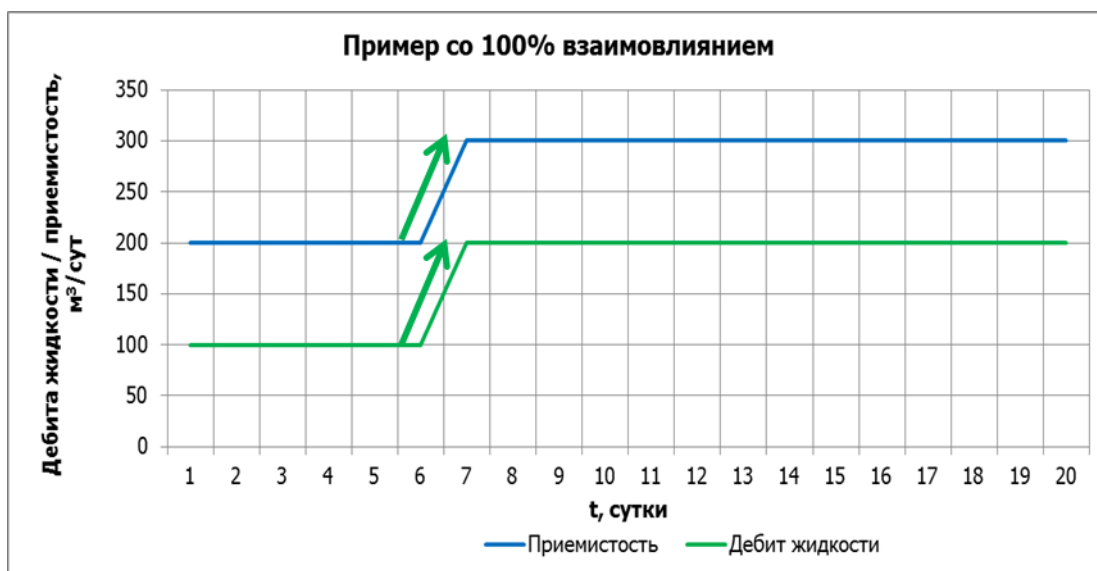


Рис.20. Пример со 100% взаимовлиянием

- 2) *Пример со сглаженной динамикой.*

Отличием от предыдущего примера является монотонный рост дебита жидкости после однократного увеличения приемистости (рис. 21).

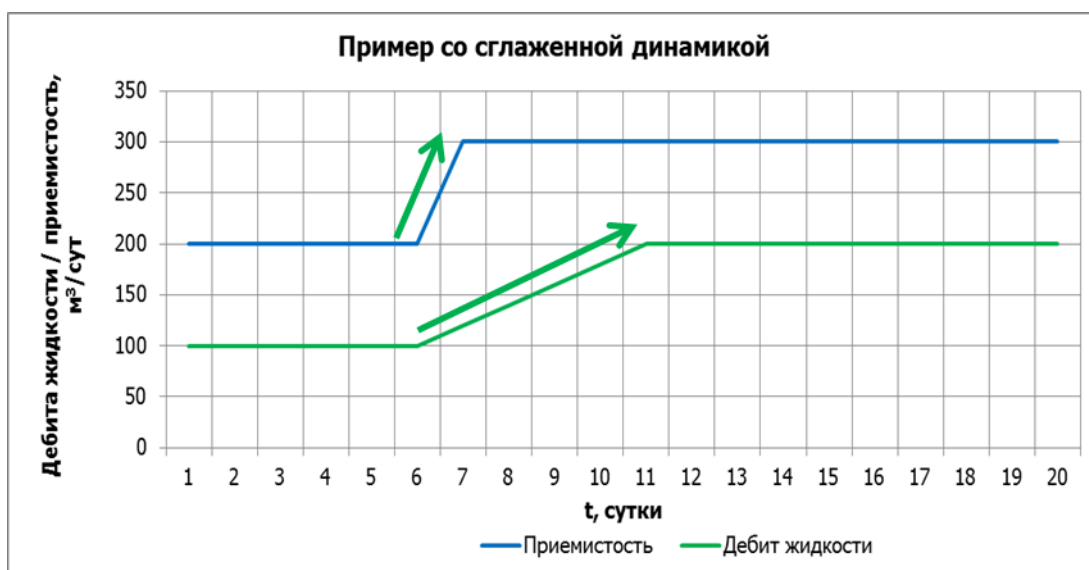


Рис.21. Пример со сглаженной динамикой

3) *Пример с отсутствием взаимовлияния.* Отличительной чертой примера является то, что дебит жидкости не реагирует на изменение приемистости.

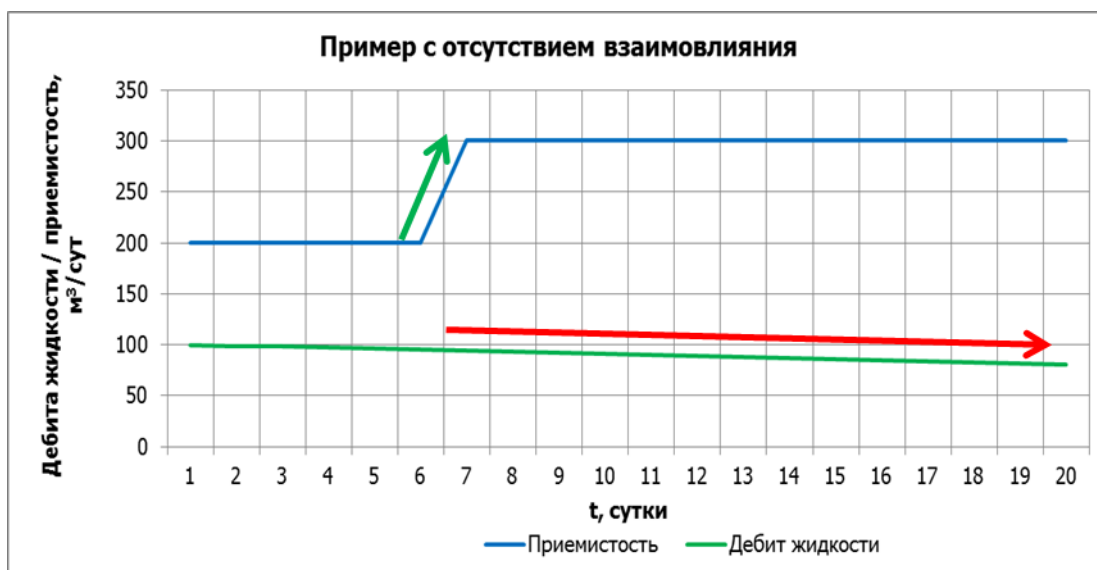


Рис.22. Пример с отсутствием взаимовлияния

4) *Пример с запаздыванием реакции.* Особенностью данного примера является наличие временного лага: отклик добывающей скважины произошел спустя некоторое время, после смены режима работы на нагнетательной.



Рис.23. Пример с запаздыванием реакции

Далее представлены результаты расчетов коэффициента корреляции для каждого примера. Значения, найденные с помощью метода вейвлет-анализа, сопоставлялись со значением коэффициента корреляции Пирсона,

который является одним из общих методов для поиска корреляции данных, он характеризует существование линейной зависимости между двумя величинами, а также он прост в реализации и занимает малый объем времени для расчета. Данное сопоставление сделано для того, чтобы продемонстрировать насколько коэффициенты корреляции, полученные с помощью вейвлет-анализа точны в различных ситуациях и чувствительны к изменениям в поведении сигналов. Расчеты проводились с помощью 6 видов ортогональных вейвлетов, три из которых представлены в различных порядках разложения.

Таблица 5. Корреляционные коэффициенты идеальных примеров

		Пример 1	Пример 2	Пример 3	Пример 4
Коэффициент корреляции Пирсона		1.00	0.88	-0.79	0.24
Метод Фурье		1.00	0.67	-0.45	0.22
Коэффициент корреляции с учетом вейвлет преобразований	Добеши 3 порядка	1.00	-0.04	-0.33	-0.44
	Добеши 4 порядка	1.00	0.75	-0.34	-0.91
	Добеши 5 порядка	1.00	0.28	-0.67	-0.43
	Симлет 2 порядка	1.00	0.97	-0.50	-0.87
	Симлет 3 порядка	1.00	-0.04	-0.33	-0.44
	Симлет 4 порядка	1.00	0.24	-0.35	-0.17
	Койфлет 2 порядка	1.00	0.37	-0.55	-0.56
	Койфлет 3 порядка	1.00	0.73	-0.38	-0.53
	Койфлет 4 порядка	1.00	0.19	0.13	-0.27
	Вейвлет Хаара	1.00	0.50	-0.50	0.82
	Дискретный вейвлет Мейера	1.00	0.26	-0.32	-0.83
	Биортогональный	1.00	-0.18	-0.02	-0.69

Для того чтобы проверить, является ли корреляция значимой, имеется наличие взаимовлияния или его отсутствие, необходимо чтобы выполнялось условие $r_{xy} > r_{крит.}$, где $r_{крит.}$ – это критическое значение корреляции, то есть если значение выше критического, то можно утверждать о наличии

взаимовлияния, если же ниже, то взаимовлияние отсутствует. Критические значения коэффициентов корреляции представлены в таблице:

**Таблица 6. Критические значения коэффициентов корреляции r-Пирсона
(r-Спирмена)**

n	P				n	P				n	P			
	0,10	0,05	0,01	0,001		0,10	0,05	0,01	0,001		0,10	0,05	0,01	0,001
5	0,805	0,878	0,959	0,991	33	0,291	0,344	0,442	0,547	61	0,213	0,252	0,327	0,411
6	0,729	0,811	0,917	0,974	34	0,287	0,339	0,436	0,539	62	0,211	0,250	0,325	0,408
7	0,669	0,754	0,875	0,951	35	0,283	0,334	0,430	0,532	63	0,209	0,248	0,322	0,405
8	0,621	0,707	0,834	0,925	36	0,279	0,329	0,424	0,525	64	0,207	0,246	0,320	0,402
9	0,582	0,666	0,798	0,898	37	0,275	0,325	0,418	0,519	65	0,206	0,244	0,317	0,399
10	0,549	0,632	0,765	0,872	38	0,271	0,320	0,413	0,513	66	0,204	0,242	0,315	0,396
11	0,521	0,602	0,735	0,847	39	0,267	0,316	0,408	0,507	67	0,203	0,240	0,313	0,393
12	0,497	0,576	0,708	0,823	40	0,264	0,312	0,403	0,501	68	0,201	0,239	0,310	0,390
13	0,476	0,553	0,684	0,801	41	0,260	0,308	0,398	0,495	69	0,200	0,237	0,308	0,388
14	0,458	0,532	0,661	0,780	42	0,257	0,304	0,393	0,490	70	0,198	0,235	0,306	0,385
15	0,441	0,514	0,641	0,760	43	0,254	0,301	0,389	0,484	80	0,185	0,220	0,286	0,361
16	0,426	0,497	0,623	0,742	44	0,251	0,297	0,384	0,479	90	0,174	0,207	0,270	0,341
17	0,412	0,482	0,606	0,725	45	0,248	0,294	0,380	0,474	100	0,165	0,197	0,256	0,324
18	0,400	0,468	0,590	0,708	46	0,246	0,291	0,376	0,469	110	0,158	0,187	0,245	0,310
19	0,389	0,456	0,575	0,693	47	0,243	0,288	0,372	0,465	120	0,151	0,179	0,234	0,297
20	0,378	0,444	0,561	0,679	48	0,240	0,285	0,368	0,460	130	0,145	0,172	0,225	0,285
21	0,369	0,433	0,549	0,665	49	0,238	0,282	0,365	0,456	140	0,140	0,166	0,217	0,275
22	0,360	0,423	0,537	0,652	50	0,235	0,279	0,361	0,451	150	0,135	0,160	0,210	0,266
23	0,352	0,413	0,526	0,640	51	0,233	0,276	0,358	0,447	200	0,117	0,139	0,182	0,231
24	0,344	0,404	0,515	0,629	52	0,231	0,273	0,354	0,443	250	0,104	0,124	0,163	0,207
25	0,337	0,396	0,505	0,618	53	0,228	0,271	0,351	0,439	300	0,095	0,113	0,149	0,189
26	0,330	0,388	0,496	0,607	54	0,226	0,268	0,348	0,435	350	0,088	0,105	0,138	0,175
27	0,323	0,381	0,487	0,597	55	0,224	0,266	0,345	0,432	400	0,082	0,098	0,129	0,164
28	0,317	0,374	0,479	0,588	56	0,222	0,263	0,341	0,428	450	0,078	0,092	0,121	0,155
29	0,311	0,367	0,471	0,579	57	0,220	0,261	0,339	0,424	500	0,074	0,088	0,115	0,147
30	0,306	0,361	0,463	0,570	58	0,218	0,259	0,336	0,421	600	0,067	0,080	0,105	0,134
31	0,301	0,355	0,456	0,562	59	0,216	0,256	0,333	0,418					
32	0,296	0,349	0,449	0,554	60	0,214	0,254	0,330	0,414					

Объем выборки во всех идеальных примерах $n = 40$, уровень доверия (вероятность того, что доверительный интервал содержит истинное значение параметра) равен 99%, статистическая вероятность $P = 1 - 0,99 = 0,01$. На пересечении этих данных находится критическое значение корреляции равное 0,403. [6]

С учетом критического значения и выполненных расчетов видно, что вейвлет Хаара показывает достаточно точную корреляцию сигналов во всех

четырёх примерах по сравнению с остальными вейвлетами и линейным коэффициентом корреляции Пирсона, а также методом Фурье. Следует отметить, что вейвлет Хаара, оказался более достоверным в четвертом примере, где есть запаздывание реакции и выраженное взаимовлияние между сигналами. Вейвлеты: Добеши 4 порядка, Симлет 2 порядка и Койфлет 3 порядка, показали верные значения корреляции во втором и третьем примере, но в четвертом, как и коэффициент корреляции Пирсона и метод Фурье, показали ошибочные значения, то есть отсутствие влияния.

Полученные результаты позволяют отсеять вейвлеты, которые не подходят для решения задачи о взаимовлиянии, кроме того, следует обратить внимание на вейвлет Хаара, который показывает отличные результаты во всех четырех идеальных примерах. Далее будут использоваться в рассмотрении примеров реального месторождения вейвлеты вида: вейвлет Хаара, вейвлет Добеши 4 порядка, Симлет 2 порядка и Койфлет 3 порядка.

ГЛАВА 5. ТЕСТИРОВАНИЕ АЛГОРИТМА НА РЕАЛЬНОМ ПРИМЕРЕ ДАННЫХ

Применим разработанный алгоритм для данных добывающих и нагнетательных скважин реального месторождения. С добывающих скважин берутся показатели дебита жидкости, а с нагнетательных приемистость. Оба ряда соответствуют одному и тому же временному интервалу, временной шаг берется равным 1 месяцу, показатели дебита жидкости и приемистости соответствуют среднемесячным замерам.

В качестве примеров выбраны скважины:

- 1) № 2969Г – добывающая скважина, соседняя нагнетательная скважина № 2982;

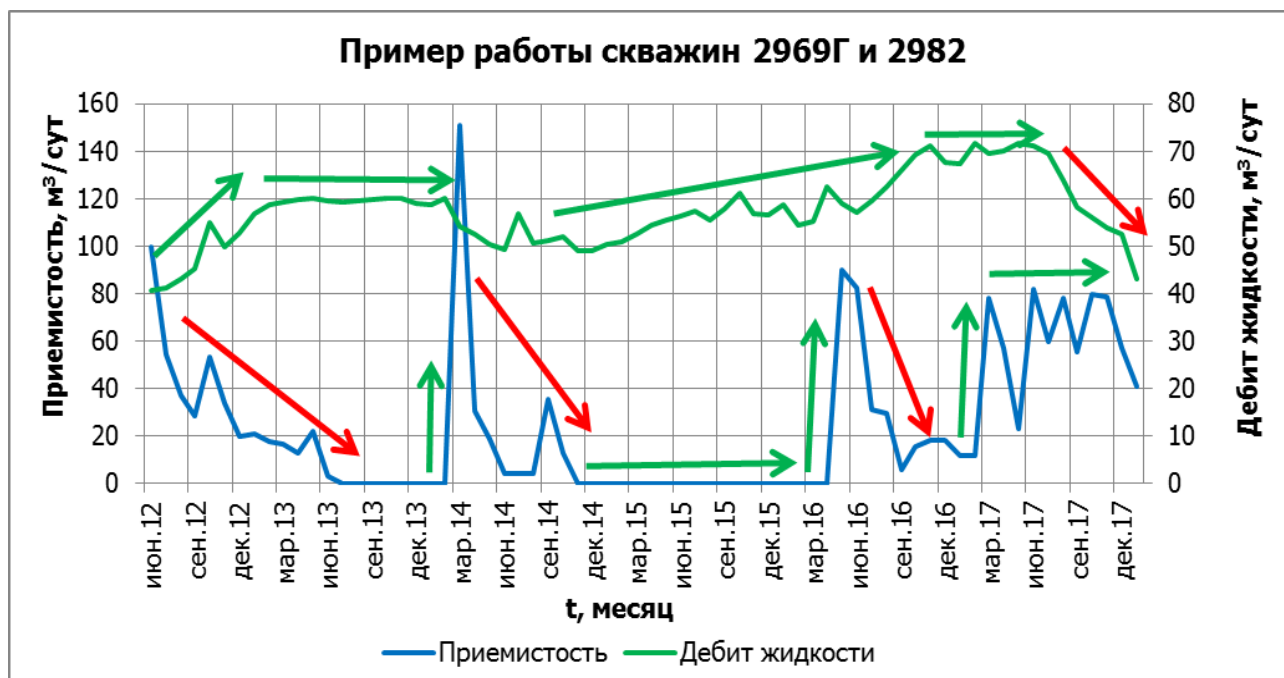


Рис.24. Пример работы скважин 2969Г и 2982

2) № 1779Г – добывающая скважина, № 7551 – соседняя нагнетательная скважина;

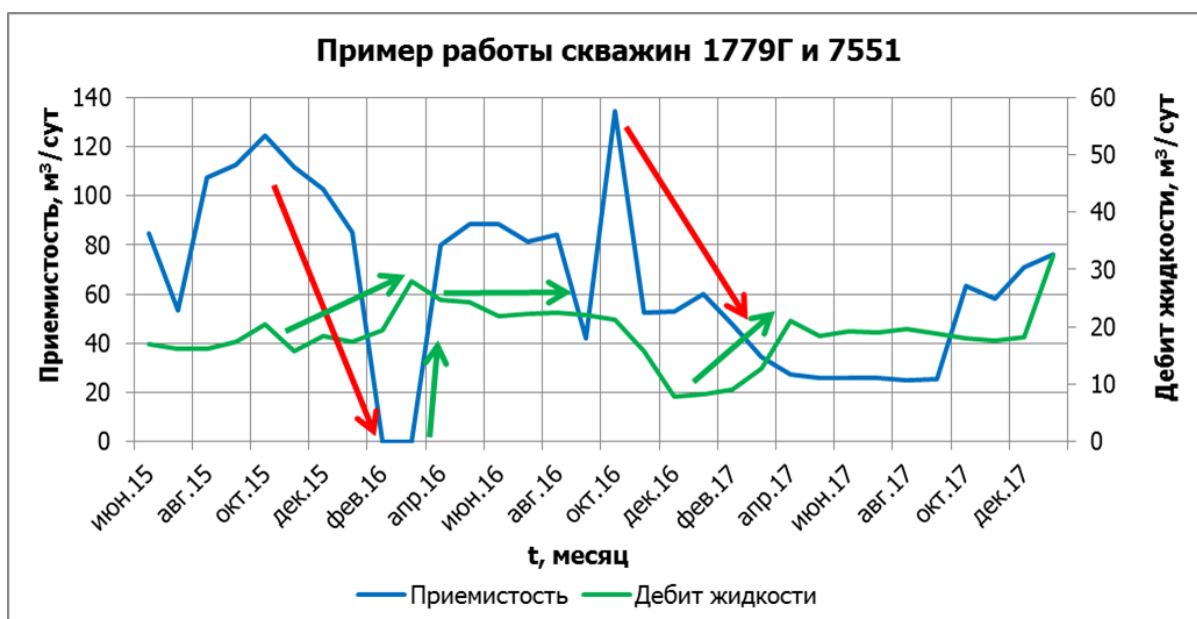


Рис.25. Пример работы скважин 1779Г и 7551

3) № 2082Г – добывающая скважина, № 2081 – соседняя нагнетательная скважина.

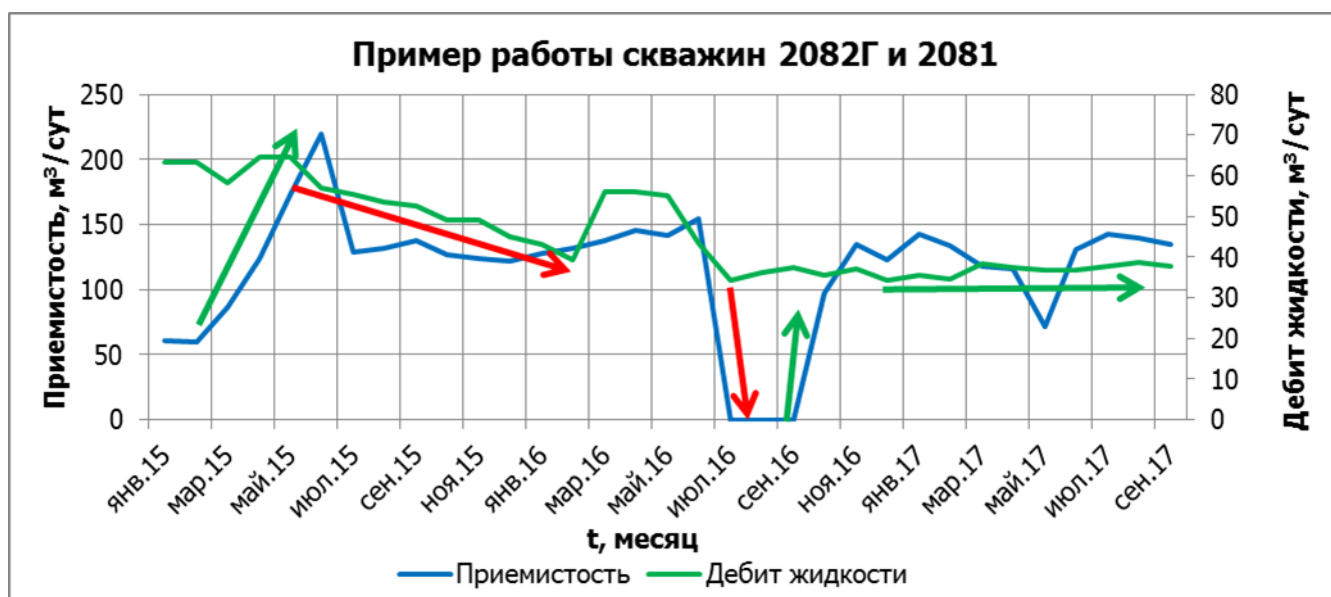


Рис.26. Пример работы скважин 2082Г и 2081

4) № 6886Г – добывающая скважина, № 6410У – соседняя нагнетательная скважина.

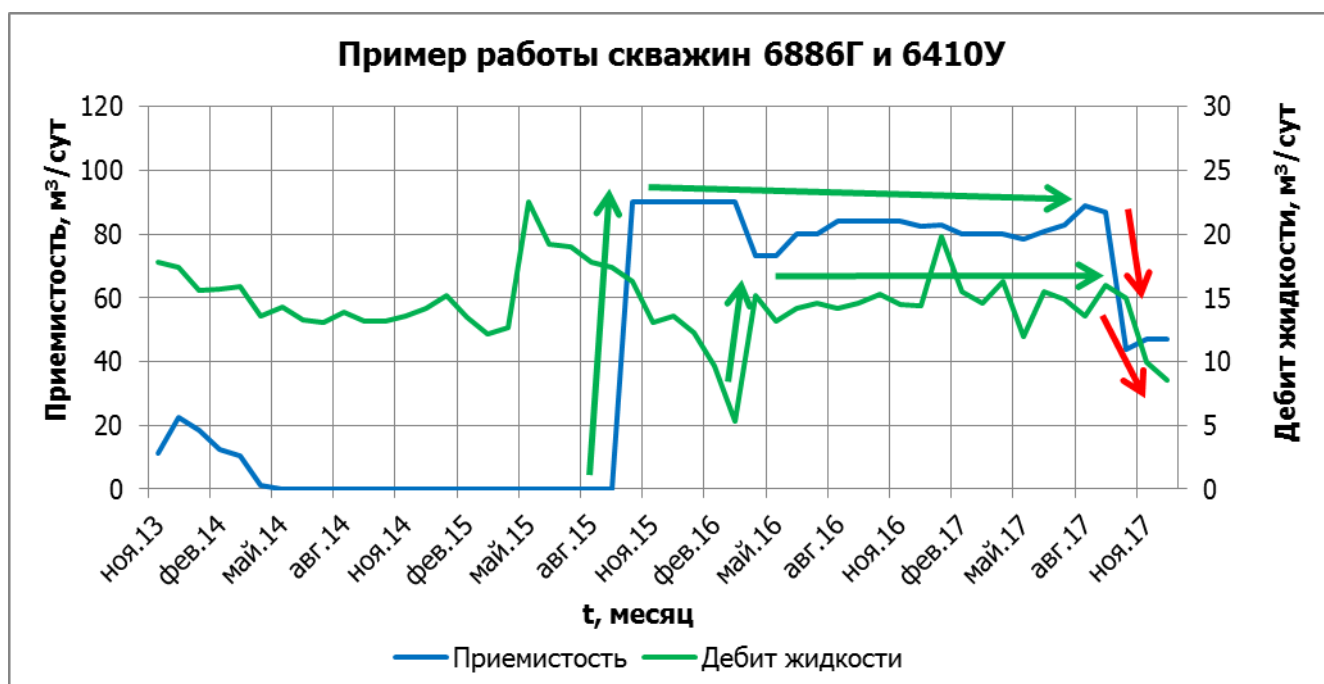


Рис.27. Пример работы скважин 6886Г и 6410У

В виде таблицы, представлены результаты расчетов значения коэффициента корреляции для каждого примера.

Таблица 7. Корреляционные коэффициенты реальных примеров

№ Скважины добывающая/нагнетательная	Коэффициент корреляции Пирсона	Метод Фурье	Коэффициент корреляции с учетом вейвлет преобразований			
			Добеши 4 порядка	Симлет 2 порядка	Койфлет 3 порядка	Вейвлет Хаара
2969Г/2982	-0.024	-0.072	0.070	0.080	-0.037	-0.770
1779Г/7551	0.039	0.121	-0.070	0.103	-0.070	-0.344
2082Г/2081	0.230	0.542	0.574	-0.441	0.425	-0.628
6886Г/6410У	-0.229	0.683	0.096	-0.208	0.714	0.719

Объем выборки для первого примера составил $n = 136$ элементов, соответствующее критическое значение корреляции равно 0,225. Объем второй выборки составляет $n = 64$ элемента, соответствующее критическое значение корреляции равно 0,320. Объем выборки для третьего примера составил $n = 66$, критическое значение корреляции равно 0,315. Объем выборки для четвертого примера составляет $n = 100$, критическое значение корреляции равно 0,256. В результате выполненных расчетов для первого и

второго примера, в которых отсутствует взаимовлияние скважин, все методы показали верные значения корреляции, которые подтверждают отсутствие влияния. В третьем примере вейвлет Добеши 4 порядка, Койфлет 3 порядка и метод Фурье показали неверное наличие взаимовлияния между скважинами. По результатам четвертого примера, вейвлет Хаара, Койфлет 3 порядка и метод Фурье показали верное наличие взаимовлияния между нагнетательной и добывающей скважиной.

Исходя из полученных результатов, можно сделать вывод, что вейвлет Хаара подходит для дальнейшего применения в расчетах взаимовлияния скважин, так как его показатели оказались верны во всех четырех реальных и четырех идеальных примерах, что показывает его полное соответствие поставленной задаче.

Почему именно ступенчатая функция, вейвлет Хаара дала наилучшие результаты, возможно, это связано с тем, что на реальном месторождении весь добывающий фонд работает механизированным способом, то есть спущен насос. У любого насоса есть определенная наработка, межремонтный период, после которого его меняют. Оператор должен поменять насос на насос такой же производительности, либо, если существуют предпосылки для оптимизации, то на более производительный, либо для диоптимизации на менее. В такие моменты происходит ступенчатая смена дебита жидкости, ровно до следующего момента пока на скважине не произойдет очередная поломка, либо наступит очередная дата наработки насоса, когда его нужно будет сменить. Поэтому каждый раз происходит серия ступенчатых смен дебита жидкости, возможно именно это связано с тем, что ступенчатый вейвлет Хаара показал лучшие результаты в исследовании.

ЗАКЛЮЧЕНИЕ

Таким образом, поставленная цель создания алгоритма для автоматизированного расчета взаимовлияния скважин, достигнута.

При реализации алгоритма, исходя из поведения динамики скважин следует отметить, что необходимо индивидуально подбирать базисный вейвлет, но для выбранных идеальных примеров и примеров реального месторождения показал свое полное соответствие методике – вейвлет Хаара.

Данный алгоритм позволяет, с высокой точностью, определить наличие или отсутствие взаимовлияния между скважинами, а также он оптимизирует время работы пользователя, то есть минимизирует затраты требуемого на работу времени.

Считаю данную тему перспективной для дальнейшего исследования: сравнение результатов работы разработанного алгоритма с трассерными исследованиями и гидропрослушиванием, а также другими методами, исследования взаимовлияния между скважинами, используемыми в практике.

СПИСОК ЛИТЕРАТУРЫ

1. Антипов О.И., Добрянин А.В., Неганова Е.В., Неганов В.А. Экономические науки. 2010. Т. 66. № 5. С. 260-270.
2. Астафьева Н.М. Вейвлет-анализ: основы теории и примеры применения. Анализ временных метеорологических рядов [Электронный ресурс]. Режим доступа: http://edu.alnam.ru/book_away.php?id=16 (Дата обращения 21.01.2018).
3. Блаттер К. Вейвлет анализ. Основы теории. М.: Техносфера, 2004.
4. Бокс Дж., Дженкинс Г., Анализ временных рядов прогноз и управление. Выпуск 1. Изд-во Мир, 1974. – 408 с.
5. Бурнаев Е.В., Оленев Н.Н. Мера близости для временных рядов на основе вейвлет коэффициентов. //Тр. XLVIII научн. Конф. МФТИ. Долгопрудный: ФУПМ, 2005. С.108-110.
6. Гмурман В.Е. Теория вероятностей и математическая статистика: Учеб. Пособие для вузов/В.Е. Гмурман.-9-е изд., стер. - М.: Высш. шк., 2003. - 479 с.
7. Добеши И. Десять лекций по вейвлетам. М.: НИЦ “Регулярная и хаотическая динамика”, 2004.
8. Ломоносова Л.С., Фалькова О.Б. Спектральный анализ. Изд-во Государственное научно-техническое издательство литературы по черной и цветной металлургии, 1958, 364 с.
9. Орлов Ю., Осминин. К. Нестационарные временные ряды. Методы прогнозирования с примерами анализа финансовых и сырьевых рынков., Изд-во Либроком, 2017. – 384 с.
10. Петерс Э. Фрактальный анализ финансовых рынков. Применение теории хаоса в инвестициях и экономике. Изд-во «Интернет-трейдинг», 2004. – 304 с.
11. Сериков А.Е. Социология: методология, методы и математическое моделирование. Социология: 4М., 2006. С. 162-183.

12. Смоленцев Н.К. Основы теории вейвлетов. Вейвлеты в MATLAB. - М.: ДМК Пресс. - 304 с.
13. Уэлстид С. Фракталы и вейвлеты для сжатия изображения в действии. М.: Триумф, 2003.
14. Фатьянов М.П. Применение метода кросс-вейвлетов для анализа финансовых рядов // Молодой ученый. — 2014. — №3. — С. 83-88.[Электронный ресурс]. Режим доступа: <https://moluch.ru/archive/62/9304/> (дата обращения: 06.03.2018).
15. Христофоров А.В., Христофорова Н.Н., Бурганов Б.Т. Температурные волны в природных средах: Фурье и вейвлет-анализ термограмм скважин / Научно-технический журнал Георесурсы.- 2(17).- 2005.- с.2-6.
16. Abarbanel H. D. I. Analysis of Observed Chaotic Data. // Springer, New York, 1996.
17. Costa M. Golderberger A.L. Peng C.K. Multiscale Entropy Analysis of Complex Physiologic Time Series // Phys. Rev. lett., 2002, V.89, №6.
18. Scafetta N. and B. J. West Solar Flare Intermittency and Earth's Temperature Anomalies // Phys. Rev. lett., 2003.
19. S. Rehman, A. H. Siddiqui, N.M. Al-Abbadi, “Wavelet based correlation of time series of Saudi Meteorological Data”, Chaos, Solitons and Fractals, v. 39(4), p. 1764-1789, 2009.
20. Stoica P., Moses R., Spectral analysis of signals // Prentice Hall; 1 edition, 2005, 452p.
21. Wolf D. Noise in Physical systems // Shpringer, New-York, 1978.