

ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное образовательное учреждение  
высшего профессионального образования  
**Национальный исследовательский университет**  
**«Высшая школа экономики»**  
Московский институт электроники и математики

Факультет Прикладной математики и кибернетики  
Кафедра Кибернетики

Дипломная работа  
по специальности 230401.65 «Прикладная математика»

**Исследование зависимости климатических параметров  
от орбитальных параметров с использованием  
нейросетевого моделирования.**

Студент группы М-96

Исаев А.Е.

Руководитель

доцент, к.т.н.

Боголюбов Д.П.

Зав. кафедрой

профессор, д.т.н.

Афанасьев В.Н.

Москва 2014

Реферат

Дипломная работа по специальности 230401.65 «Прикладная математика» студента группы М-96 Исаева Алексей Евгеньевича на тему:

Исследование зависимости климатических параметров от орбитальных параметров с использованием нейросетевого моделирования.

83с., 41 рис., 36 источников.

ЦИРКУЛЯЦИОННЫЕ ПРОЦЕССЫ В АТМОСФЕРЕ, ПРИХОДЯЩАЯ СОЛНЕЧНАЯ РАДИАЦИЯ, ВРЕМЕННОЙ РЯД, СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ, ПРОГНОЗИРОВАНИЕ.

Целью данной дипломной работы является исследование зависимости климатических параметров планеты Земля (в частности, давления) от орбитальных параметров, выраженных проходящей солнечной радиацией, по заказу научного консультанта. Цель исследования – обогащение имеющихся сведений об атмосферных процессах, происходящих на нашей планете. Исследование проводится на примере Северо-Атлантического течения, точнее, двух связанных с ним географических объектов: Исландии и Азорских островов. Дополнительной задачей работы является попытка прогнозирования климатических параметров на основе уже имеющихся данных по предыдущим периодам времени.

Исследование опиралось как на стандартные математические, статистические и аналитические методы, так и на систему поддержки принятия решений, основанную на нейросетевом моделировании и разработанную аспирантом МИЭМ НИУ ВШЭ Бухаровым Олегом Евгеньевичем.

В работе проводится анализ имеющихся данных и подтверждение на их основе существующей теории, что даёт возможность для дальнейших исследований в данной области.

## **Annotation.**

Degree work on speciality 230401.65 «Applied mathematics» by student of M-96 group Isaev Alexey Evgenyevich on topic:

### **Examination of the dependence of climatic parameters on orbital parameters with the use of neural network modeling**

83 pages, 41 pictures, 36 sources.

The main purpose of this work is the analysis of dependency between climatic parameters of Earth (in our case, it is atmospheric pressure) and orbital parameters, which are represented by received solar radiation. The particular subject of the analysis is North-Atlantic Oscillation, in particular it's two objects: Iceland and Azore Islands. Additional task is forecasting of climatic phenomenons, based on previous data.

The work implies mathematical, analitical and statistical methods as well as decision helping support system, which is based on neural network modelling.

The result of the work is confirmation of current teories in climatology, what gives us more opportunities for subsequent researches in this area.

## Содержание.

Введение	6
1. Постановка задачи	7
1.1. Введение в теорию циркуляционных процессов.	7
1.2. Методика проведения расчётов.	8
1.3. Результаты, достигнутые теорией.	9
1.3.1 Годовая солнечная радиация	9
1.3.2 Сезонная солнечная радиация	14
1.3.3 Межгодовая изменчивость приходящей радиации	21
1.4 Условия задачи	28
2. Модель прогнозирования	29
2.1 Предпосылки к использованию модели прогнозирования	29
2.2 Общие сведения об СППР.	30
2.2.1 Классификация СППР	30
2.2.2 Основные определения	33
2.2.3 Классификация временных рядов	34
2.2.4 Анализ общих методов поддержки принятия решений	36
2.2.5 Актуальность создания алгоритма прогнозирования временных рядов.	40
2.3 Методы реализации алгоритмов прогнозирования временных рядов	43
2.3.1 Использование интервальных нейронных сетей	43
2.3.2 Интервальные нейронные сети: основные определения и принцип работы.	44
2.3.3 Использование генетических алгоритмов	47
2.3.4 Иллюстрация работы генетического алгоритма	52

2.4 Реализация прогнозирования временных рядов	57
2.4.1 Функциональная схема системы прогнозирования временных рядов и детали реализации	57
2.4.2 Алгоритм работы оболочки системы	59
3. Компьютерный эксперимент	61
3.1 Исходные данные по исследуемым объектам.	61
3.1.1 Исландия	62
3.1.2 Азорские острова	65
3.1.3 Разность – Азорские минус Исландия	68
3.2 Результаты первичной обработки данных	71
3.3 Поиск связи годового хода давления и радиации	72
3.4 Многолетний ход по среднемесячным значениям	77
Заключение	81
Список использованных источников	82

## Введение

Целью данной дипломной работы является исследование зависимости климатических параметров планеты Земля (в частности, давления) от орбитальных параметров, выраженных приходящей солнечной радиацией, по заказу научного консультанта. Цель исследования – обогащение имеющихся сведений об атмосферных процессах, происходящих на нашей планете. Исследование проводится на примере Северо-Атлантического течения, точнее, двух связанных с ним географических объектов: Исландии и Азорских островов. Дополнительной задачей работы является попытка прогнозирования климатических параметров на основе уже имеющихся данных по предыдущим периодам времени.

В статье В.М. Фёдорова «Солярные реперы в общей циркуляции атмосферы» сообщается о существовании пространственных соответствий в широтной локализации особенностей распределения приходящей на верхнюю границу атмосферы солнечной радиации и особенностей общей циркуляции атмосферы: средних положений экваториальной депрессии, субтропических зон высокого и субполярных зон низкого давления, полярных фронтов, умеренных и тропических воздушных масс. Причины найденных пространственных (топологических) соответствий определяются особенностями в распределении приходящей солнечной радиации, связанными с формой Земли, наклоном оси ее вращения и возмущенным орбитальным движением.

Таким образом, данное направление предоставляет огромное количество задач, актуальность которых обусловлена необходимостью прогнозирования погодных явлений, на протяжении всей истории человечества играющих значительную роль в его жизнедеятельности, а также невозможностью обработки данных без помощи математического аппарата.

Для анализа имеющихся исходных данных была использована система поддержки принятия решений, разработанная аспирантом МИЭМ НИУ ВШЭ Бухаровым Олегом, методы прогнозирования временных рядов, статистические и аналитические методы. Основное использование программы будет производиться для решения дополнительной задачи.

## **1 Постановка задачи.**

### **1.1 Введение в теорию циркуляционных процессов.**

Согласно существующей теории, Солнце является основным источником энергии атмосферных процессов. Приходящая к Земле солнечная радиация определяет зональность в распределении среднеширотных температур как в приземном слое атмосферы, так и в поверхностном слое океана. Пространственная и временная неоднородность в распределении приходящей радиации является основной причиной циркуляционных процессов в атмосфере и особенностей общей циркуляции атмосферы. Важнейшей особенностью общей циркуляции атмосферы (ОЦА) является географическая локализация климатологических полярных фронтов (разделяющих умеренные и тропические воздушные массы), экваториальной депрессии (и внутритропической зоны конвергенции - ВЗК), субтропических зон высокого и субполярных зон низкого давления. Попытки количественно объяснить географическую привязанность широтных зон повышенного и пониженного давления делались уже давно, однако удовлетворительного ответа получено не было. В современных эмпирических схемах общей циркуляции атмосферы географическое положение зон давления либо принимается как данное, либо объясняется без количественной оценки меридиональным распределением источников и стоков тепла в атмосфере. В связи с этим, нерешенный вопрос о причинах такого распределения сужает круг возможных приложений моделей и схем ОЦА, например, в отношении прогноза распределения давления и развития циркуляционных и вихревых процессов в атмосфере.

## 1.2 Методика проведения расчётов.

По данным астрономических эфемерид (JPL Planetary and Lunar Ephemerides) DE-405/406 рассчитывались значения приходящей солнечной радиации (при отсутствии атмосферы) за тропические годы, полугодия и сезоны года в различные широтные зоны (протяженностью в  $5^\circ$  широты) земного эллипсоида в интервале от 3000 г. до н.э. до 2999 г. н.э. Точность эфемерид по расстоянию между Землей и Солнцем составляет  $10^{-9}$  а.е (или 0,1496 км), по времени 1 с (или 0,0000115 сут). Поверхность Земли аппроксимировалась эллипсоидом (GRS80), с длинами полуосей равными 6378137 м (большие) и 6356752 м (малая). Расчеты производились с использованием ряда формул, среди которых основным было выражение:

$$I_{nm}(\varphi_1, \varphi_2) = \int_{t_1}^{t_2} \left( \int_{\varphi_1}^{\varphi_2} \sigma(H, \varphi) \left( \int_{-\pi}^{\pi} \Lambda(H, t, \varphi, \alpha) d\alpha \right) d\varphi \right) dt, \quad (1)$$

где  $I$  – приходящая солнечная радиация за элементарный  $n$ -й фрагмент  $m$ -го тропического года (Дж);  $\sigma$  – площадной множитель ( $\text{м}^2$ ) с помощью которого вычисляется площадной дифференциал  $\sigma(H, \varphi) d\alpha d\varphi$  – площадь бесконечно малой прямоугольной ячейки эллипсоида;  $\alpha$  - часовой угол,  $\varphi$  – географическая широта, выраженные в радианах;  $H$  – высота поверхности эллипсоида относительно поверхности Земли (м);  $\Lambda(H, \varphi, t, \alpha)$  – инсоляция в заданный момент в заданном месте поверхности эллипсоида ( $\text{Вт}/\text{м}^2$ ),  $t$  - время (с). Шаги при интегрировании составляли: по долготе  $1^\circ$ , по широте  $1^\circ$ , по времени  $1/360$  часть продолжительности тропического года. При расчетах учитывались изменения расстояния между Землей и Солнцем и продолжительности периода обращения Земли (продолжительности тропического года) в связи с периодическими возмущениями орбитального движения Земли. На основе полученных данных были определены характеристики годовой, сезонной (по полугодиям) и межгодовой изменчивости приходящей солнечной радиации по 5-градусным широтным зонам.



## 1.3 Результаты, достигнутые теорией.

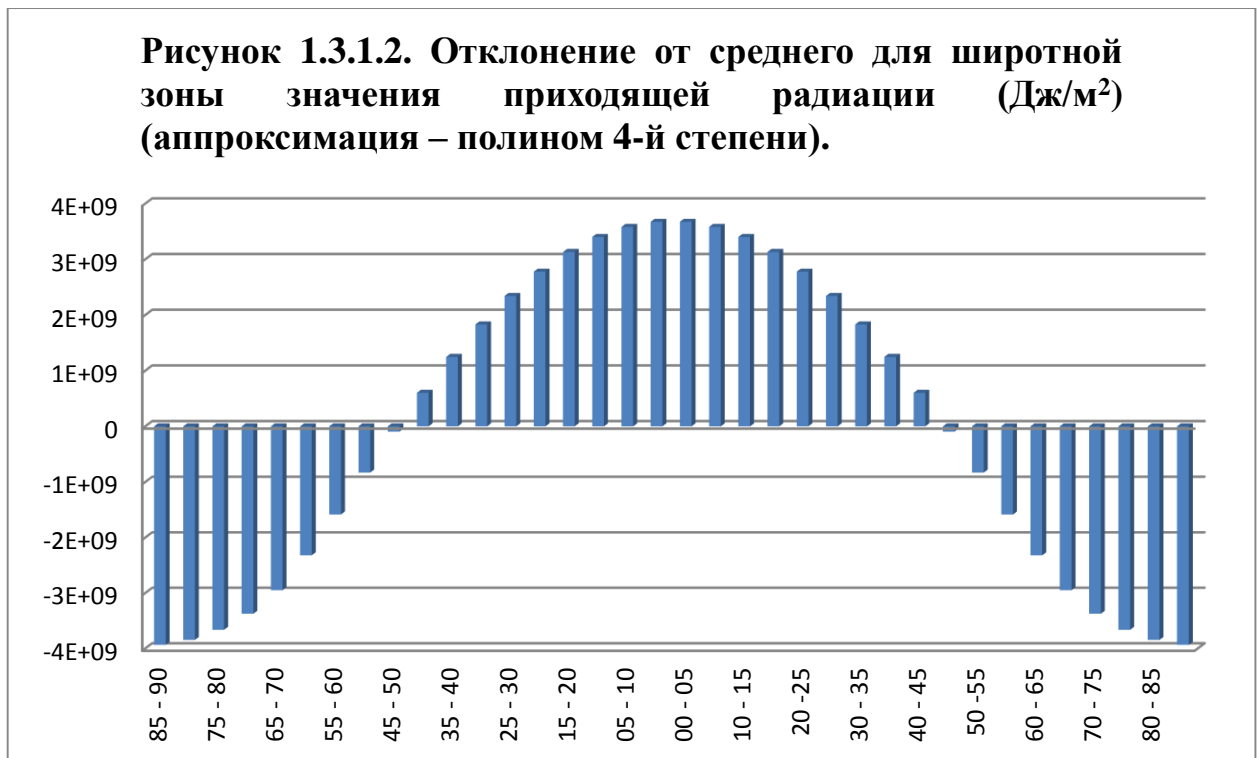
### 1.3.1 Годовая солнечная радиация.

Максимальные значения в осредненном по 5-градусным широтным зонам распределении приходящей за тропический год на верхнюю границу атмосферы солнечной радиации характерны для экваториальных зон ( $0^\circ \pm 5^\circ$ ), где они составляют  $1,31E+10$  Дж/м<sup>2</sup> (рис. 1.3.1.1). Минимальные значения отмечаются в полярных широтных зонах –  $5,54E+09$  Дж/м<sup>2</sup>. Средняя для широтной зоны величина приходящей за тропический год солнечной радиации на верхнюю границу атмосферы равна  $9,47E+09$  Дж/м<sup>2</sup>.



Полученным характером распределения приходящей за год солнечной радиации определяется симметрия средней структуры атмосферы (осредненный меридиональный разрез) относительно экваториальной плоскости. Сезонная и межгодовая изменчивость существенно изменяют полученную картину годового распределения. Характер распределения приходящей радиации по отношению к средне зональной величине радиации ( $9,47E+09$  Дж/м<sup>2</sup>) представлен на рис. 1.3.1.2.

Из полученного распределения видно, что от широтных зон  $45^{\circ} - 50^{\circ}$  в каждом полушарии к полюсам отмечаются отрицательные отклонения, в широтном диапазоне от  $40^{\circ} - 45^{\circ}$  с.ш. – до  $40^{\circ} - 45^{\circ}$  ю.ш. – положительные отклонения от среднего. Характер распределения входящей на верхнюю границу атмосферы солнечной радиации является основой приходной составляющей радиационного баланса Земли и определяет географическое положение глобальных источников и стоков тепла. Отражением полученного характера распределения входящей радиации относительно среднего значения является распределение средне зонального потока энергии в атмосфере.

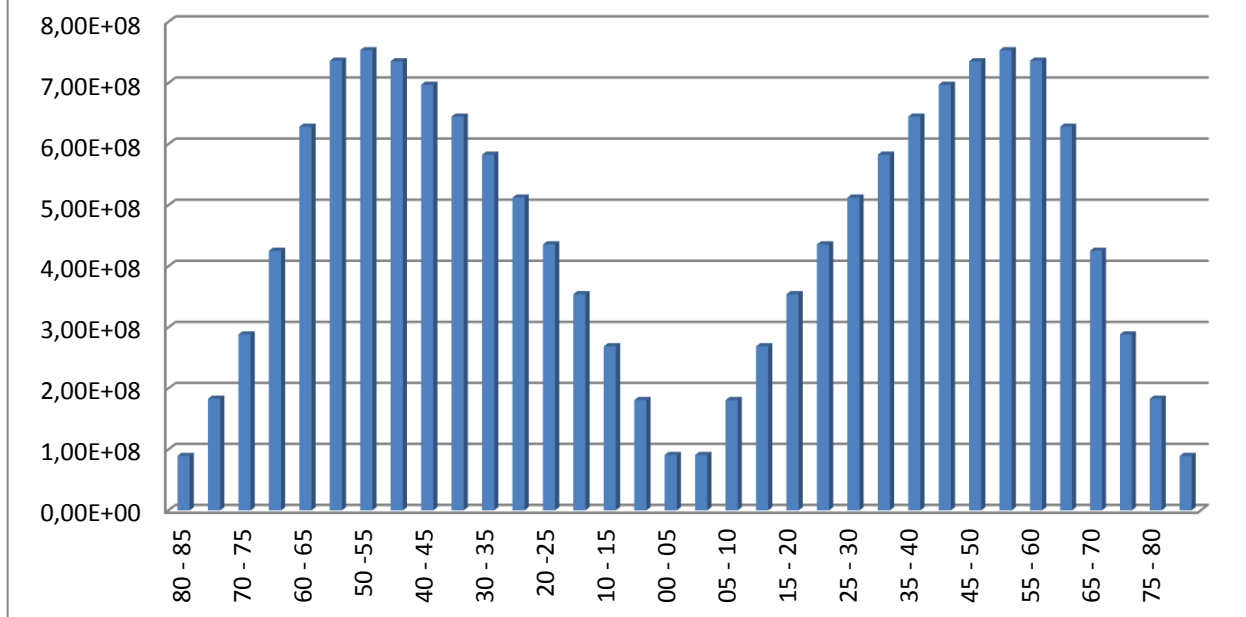


Из приведенных графиков (рис. 1.3.1.1 и 1.3.1.2) следует, что максимальные значения входящей на верхнюю границу атмосферы солнечной радиации приходится на широтный диапазон от  $5^{\circ}$  с.ш. до  $5^{\circ}$  ю.ш. Известно, что с этим широтным диапазоном связано среднее положение экваториальной депрессии и ВЗК. В зонах  $45^{\circ} - 50^{\circ}$  широты отмечается смена знака отклонения входящей радиации от среднего, т.е. они отделяют области с различными по знаку отклонениями от среднего. Приблизительно с этими широтами связана

географическая локализация климатологических полярных фронтов (и соответствующих фронтальных зон и струйных течений). Области полярнофронтальных возмущений являются одними из источников генерации энергии атмосферных процессов.

На основе среднеширотных значений (рис. 1.3.1.1) были выполнены приблизительные расчеты межширотного градиента, приходящей на верхнюю границу атмосферы, солнечной радиации (рис. 1.3.1.3). Значения межширотного градиента рассчитывались последовательным встречным вычитанием. Так, значения солнечной радиации, приходящей в полярные зоны ( $85^{\circ} - 90^{\circ}$ ), вычиталось из значений соответствующих соседних зон ( $80^{\circ} - 85^{\circ}$ ). Затем значения зон ( $80^{\circ} - 85^{\circ}$ ) также вычитались из соответствующих соседних зон ( $75^{\circ} - 80^{\circ}$ ) и т.д., до соответствующих экваториальных зон ( $0^{\circ} - 5^{\circ}$ ). Полученная широтная структура градиента приходящей солнечной радиации (и связанного с ней тепла) характеризуется наличием двух максимумов в зонах  $50^{\circ} - 55^{\circ}$  градусов каждого полушария и трех минимумов. Один минимум локализован в экваториальной области от  $5^{\circ}$  с.ш. до  $5^{\circ}$  ю.ш., два других минимума отмечаются в полярных зонах  $85^{\circ} - 90^{\circ}$  каждого полушария. При этом уменьшение градиента в направлении от максимумов в сторону полюсов происходит более интенсивно, чем уменьшение в направлении от максимумов к экватору. Значение градиента в экваториальной зоне минимумов ( $5^{\circ}$  с.ш. –  $5^{\circ}$  ю.ш.) составляет  $9,09\text{E}+07$  Дж/м<sup>2</sup>. Приблизительно такая же величина межширотного градиента отмечается для полярных зон ( $85^{\circ} - 90^{\circ}$ ) –  $8,93\text{E}+07$  Дж/м<sup>2</sup>. Максимальное значение (характерное для широтных зон  $50^{\circ} - 55^{\circ}$  широты в каждом полушарии) составляет приблизительно  $7,54\text{E}+08$  Дж/м<sup>2</sup>. Следует отметить (рис. 1.3.1.3), что области, в которых значения межширотного градиента превышают  $5,50\text{E}+08$  Дж/м<sup>2</sup>, соответствуют западному переносу, а области с меньшими значениями – восточному.

**Рисунок 1.3.1.3. Межширотный градиент приходящей солнечной радиации (Дж/м<sup>2</sup>).**



Сравнение полученного характера распределения межширотного градиента приходящей солнечной радиации с широтными особенностями ОЦА показывает, что с максимальными (по модулю) значениями межширотного градиента в приходящей за тропический год солнечной радиации, связана географическая локализация климатологических полярных фронтов, а с минимальными – среднее географическое положение экваториальной депрессии и ВЗК. Известно, что климатологический полярный фронт – как поверхность разрыва, лежащая между массами воздуха тропического и полярного (умеренных широт) происхождения – рассматривается как характерная особенность ОЦА. Подобно этому и ВКЗ – основная зона, где на нижних уровнях наблюдается сходимости воздушных течений, возникающих в южном и северном полушариях, расположенная вокруг экватора и почти всегда опоясывающая земной шар – также относится к особенностям ОЦА.

Таким образом, причина географической локализации экваториальной депрессии, ВЗК, полярных фронтов, полярных и тропических воздушных масс связана с формой Земли, определяющей особенности широтного распределения приходящей на верхнюю границу атмосферы за тропический год солнечной радиации. Происхождение же арктического и антарктических фронтов, связано,

вероятно, с широтной протяженностью линии раздела (наибольших контрастов) материков и океанов.

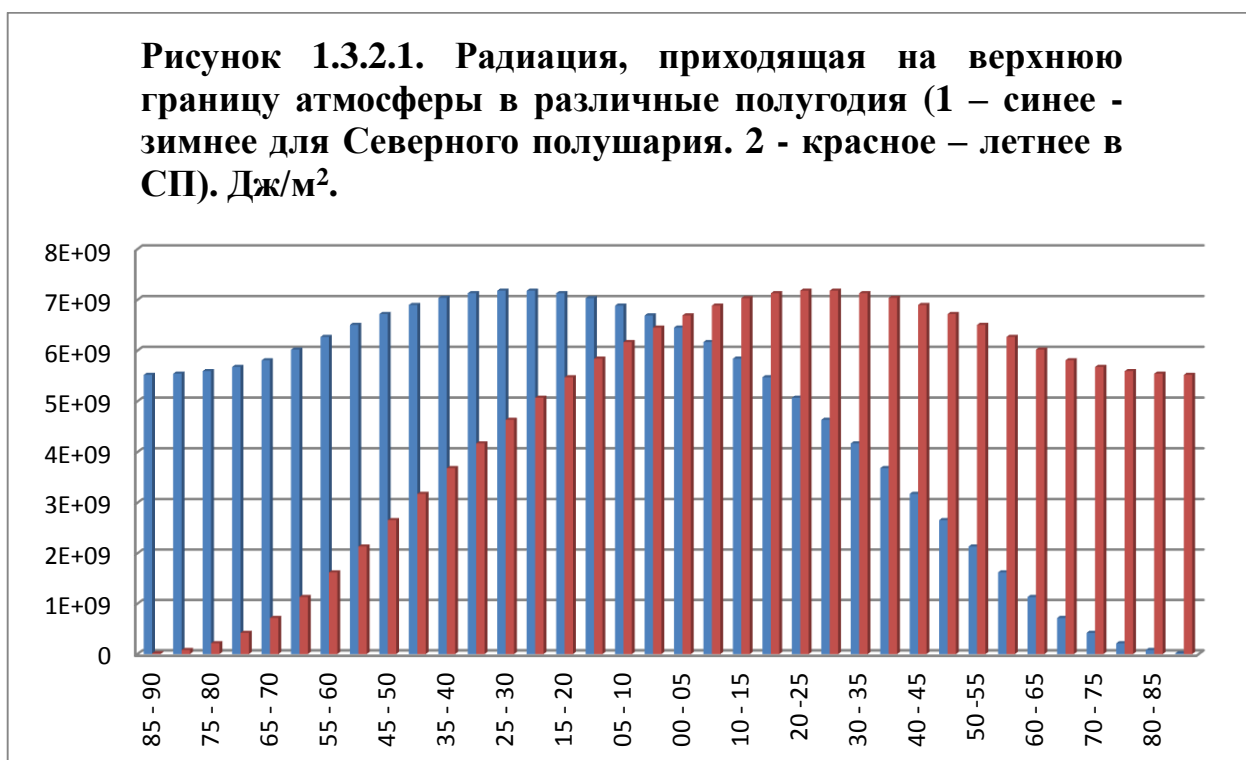
### 1.3.2 Сезонная солнечная радиация.

Среди особенностей ОЦА так же выделяются субтропические области высокого и субполярные области низкого давления. Причины образования этих зон связываются с особенностями циклонической деятельности. Считается, что подвижные антициклоны, возникающие на полярном фронте, при движении с запада на восток, смещаются в более низкие широты, где они усиливаются и образуют в каждом полушарии субтропическую зону высокого давления с осью около  $35^{\circ}$  широты. Циклоны, возникающие в тех же средних широтах у восточных берегов материков, при движении к востоку, отклоняются к более высоким широтам и, локализируются там, образуя в каждом полушарии зону низкого давления в зоне  $60^{\circ}$  –  $65^{\circ}$  градусов широты. Образование подвижных циклонов и антициклонов у восточных берегов материков в зоне полярного фронта связано, во-первых, с тем, что в этой зоне отмечается максимальное различие температурных и барических свойств воздушных масс (умеренного и тропического пояса). Во-вторых, в зимнее время год здесь усиливаются и температурные и барические различия между материком и океаном (т.е. наибольшие контрасты температуры здесь сочетаются с расходимостью изогипс и адвекцией холода).

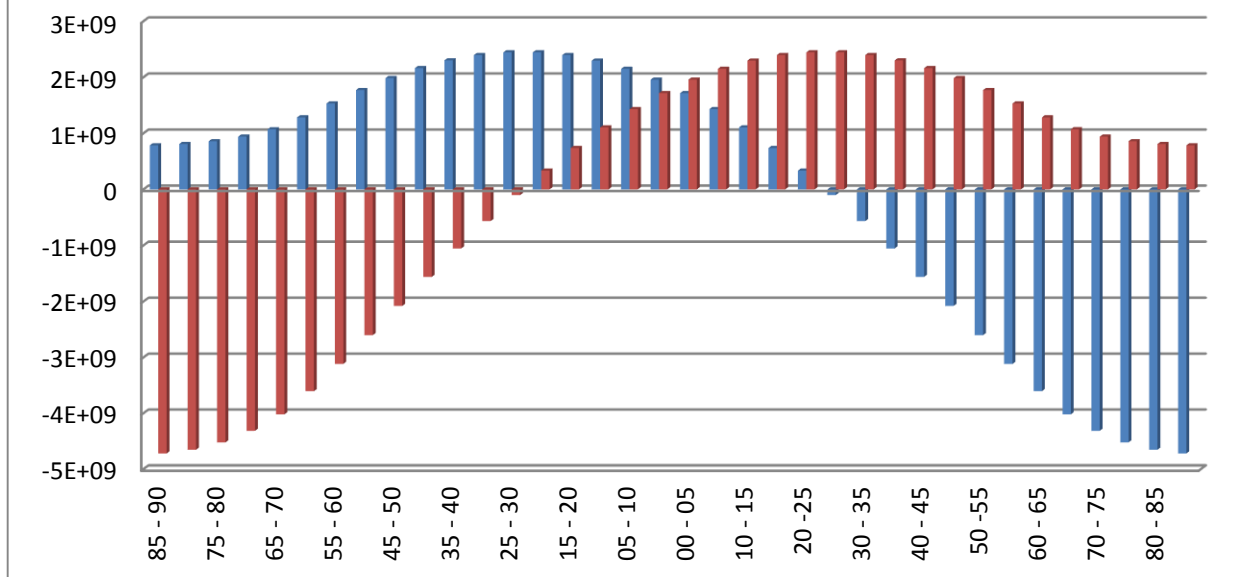
Отмечаемые траектории движения циклонов и антициклонов объясняются изменением отклоняющей силы вращения Земли с широтой. Вывод о том, что движущиеся циклоны вследствие увеличения с широтой параметра Кориолиса получают ускорение в направлении более высоких широт, а антициклоны – к экватору получен Россби и подтвержден дальнейшими исследованиями. Также известно, что если область низкого давления имеет холодное ядро, то циклонический поток вокруг ядра будет усиливаться с высотой, и наоборот. Аналогично, если область высокого давления имеет теплое ядро, то антициклоническое течение будет возрастать с высотой и наоборот. Если же центры температуры и давления не совпадают, то циклоны будут двигаться с высотой в сторону холодного воздуха, а антициклоны – в направлении теплого воздуха. В середине прошлого столетия также предпринимались попытки связать локализацию субтропической области высокого давления с радиационным и тепловым балансом

атмосферы на основе анализа карт радиационного баланса и движения воздуха на высотах. Однако, как будет показано ниже, первопричины географической локализации зон высокого и низкого давления имеют более сложную физическую природу, связанную главным образом с особенностями сезонного распределения проходящей солнечной радиации. Рассчитанные нами значения солнечной радиации проходящей на верхнюю границу атмосферы за полугодия графически представлены на рис. 1.3.2.1.

Максимальное поступление солнечной радиации отмечается в зимнее для северного полушария полугодие в диапазоне  $20^{\circ} - 30^{\circ}$  ю.ш., в летнее для северного полушария полугодие в диапазоне  $20^{\circ} - 30^{\circ}$  с.ш. Поступление солнечной радиации на верхнюю границу атмосферы в этих широтных зонах составляет  $7,18E+09$  Дж/м<sup>2</sup>. На основе полученных среднеширотных значений солнечной радиации рассчитывалось отклонение от средней для широтной зоны величины (рис. 1.3.2.2).



**Рисунок 1.3.2.2. Отклонение от среднего (полином 4-й степени) для зимнего полугодия (северное полушарие). Синее – зимнее полугодие в СП, красное – летнее. Дж/м<sup>2</sup>**



В зимнее для северного полушария полугодие в широтной области, расположенной выше  $25^{\circ} - 30^{\circ}$  с.ш., отмечаются отрицательные отклонения от среднего. В летнее для северного полушария полугодие отрицательные значения характерны для зоны  $25^{\circ} - 30^{\circ}$  ю.ш. и более высоких широтных зон этого полушария. Таким образом, в летнее для полушария полугодие с широты приблизительно около  $25^{\circ}$  зимнего полушария происходит смена знака отклонения от среднего в сезонном распределении приходящей на верхнюю границу атмосферы солнечной радиации по широтам. В зимнее полугодие для полушария смена знака приходится на широтную зону  $25^{\circ} - 30^{\circ}$  этого зимнего полушария. То есть, широтные зоны  $25^{\circ} - 30^{\circ}$  широты отражают сезонную смену в широтной локализации областей источников и стоков приходящей радиации (тепла).

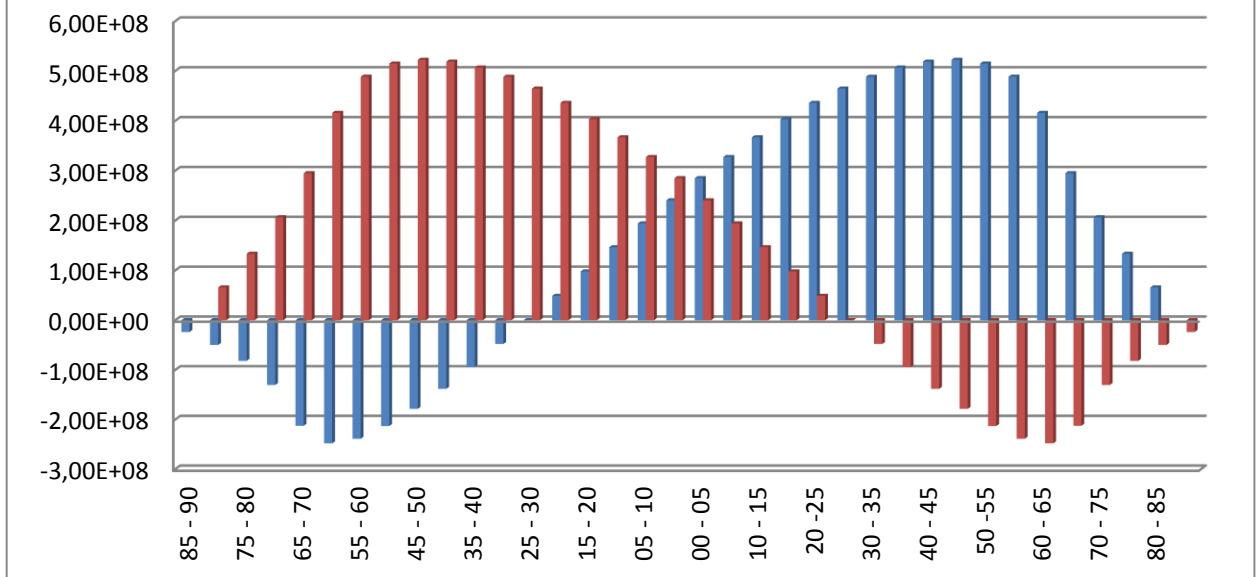
Зоны перекрытия, как следует из графика (рис. 1.3.2.2), отражают сезонную смену в поступлении радиации при сохранении положительных значений отклонения от среднего в зимний и летний сезоны. Этот широтный диапазон (от  $25^{\circ} - 30^{\circ}$  с.ш. до  $25^{\circ} - 30^{\circ}$  ю.ш.) приблизительно соответствует географическому положению циркуляционных ячеек Хэдли, представляющими собой соленоидально прямые циркуляции. Сохранение ячеек Хэдли обеспечивается поступлением тепла в



восходящей ветви в зоне сходимости пассатов (ВЗК) в сочетании с радиационными потерями в нисходящей субтропической ветви. Меридиональные течения, соединяющие эти две ветви (пассаты и противоположные им противоположные им течения в верхней тропосфере) направлены в сторону низкого давления. Считается, что здесь генерируется кинетическая энергия горизонтальных ветров, большая часть которой рассеивается под воздействием трения в нижней ветви, при меньших потерях в верхней. Ячейки Хэдли, рассматриваемые как средние циркуляции, представляют собой сравнительно простые тепловые машины. И, исходя из полученного распределения, функционирование этих тепловых машин связано с сезонными изменениями в поступлении солнечной радиации на верхнюю границу атмосферы. Область ячейки Хэдли является, таким образом, вторым источником генерации энергии в атмосфере.

По результатам распределения сезонного прихода солнечной радиации рассчитывались межширотные градиенты ее изменчивости (рис. 1.3.2.3). В этом случае градиент рассчитывался последовательным вычитанием. Для расчета межширотного градиента зимнего в северном полушарии полугодия из значения радиации в южной полярной зоне  $85^{\circ} - 90^{\circ}$  ю.ш. вычиталось значение радиации в зоне  $80^{\circ} - 85^{\circ}$  ю.ш., затем из значения радиации зоны  $80^{\circ} - 85^{\circ}$  ю.ш. вычиталось значение радиации в зоне  $75^{\circ} - 80^{\circ}$  ю.ш. и т.д. до северной полярной зоны. Для летнего в северном полушарии полугодия из значения радиации в северной полярной зоне  $85^{\circ} - 90^{\circ}$  с.ш. вычиталось значение соседней зоны  $80^{\circ} - 85^{\circ}$  с.ш. и т.д., до южной полярной зоны.

**Рисунок 1.3.2.3 Межширотный градиент синее – зимнее полугодие для СП. красное – летнее полугодие для СП. Градиент – неоднородность – неустойчивость.**



Из полученного распределения видно, что максимальные значения межширотного градиента связаны с летними для полушарий полугодиями и локализируются в зоне  $60^{\circ} - 65^{\circ}$  летнего полушария. Величина градиента в этом случае составляет  $-2,5E+08$  Дж/м<sup>2</sup>. Максимальные градиенты связаны с зимними сезонами и широтными зонами  $45^{\circ} - 50^{\circ}$  зимнего полушария. Величина межширотного градиента определяется значением  $5,22E+08$  Дж/м<sup>2</sup> (рис. 1.3.2.3). Вблизи  $25^{\circ}$  широты отмечается смена знака межширотного градиента. При этом, значение приходящей радиации в летнее полугодие в северном полушарии в полярную зону этого полушария приблизительно равно радиации поступающей в зону  $25^{\circ} - 30^{\circ}$  южного полушария. И наоборот, значение радиации в полярной зоне южного полушария в летнее полугодие приблизительно равно величине радиации характерной для зоны  $25^{\circ} - 30^{\circ}$  северного полушария. В зимнее для северного полушария полугодие в южном полушарии межширотный градиент в широтных зонах  $25^{\circ} - 30^{\circ}$  ю.ш. и расположенных южнее имеет отрицательные значения. Начиная с  $30^{\circ} - 35^{\circ}$  южной широты и до северной полярной зоны значения градиента положительные. В летнее в северном полушарии полугодие значения градиента от зоны  $25^{\circ} - 30^{\circ}$  с.ш. и севернее отрицательные, а во всех зонах начиная

с  $30^{\circ} - 35^{\circ}$  и южнее – положительные. Широтный диапазон от  $25^{\circ} - 30^{\circ}$  с.ш. до  $25^{\circ} - 30^{\circ}$  ю.ш. характеризуется положительными значениями межширотного градиента во все сезоны (ячейки Хэдли). Широтные области от  $25^{\circ} - 30^{\circ}$  ю.ш. до  $60^{\circ} - 65^{\circ}$  ю.ш. в зимнее для северного полушария полугодие и от  $25^{\circ} - 30^{\circ}$  с.ш. до  $60^{\circ} - 65^{\circ}$  с.ш. в летнее для северного полушария полугодия характеризуются изменением значений межширотного градиента радиации от нуля до минимальных отрицательных. Эти широтные области соответствуют обратным циркуляционным ячейкам Ферреля (Ferrel) в ранних схемах ОЦА.

Из графика (рис. 1.3.2.3) также видно, что градиенты приходящей на верхнюю границу атмосферы солнечной радиации в зимние полугодия в каждом полушарии по модулю значительно (более чем вдвое) превосходят летние.

Таким образом, зонально симметричное распределение радиации характерное для года, в масштабах сезонов сменяется на зонально асимметричное распределение (рис. 1.3.2.1 – 1.3.2.3). При этом широтные зоны  $25^{\circ} - 30^{\circ}$  и  $60^{\circ} - 65^{\circ}$  в каждом полушарии в различные полугодия соответствуют в ОЦА границам зонального расчленения нижних слоев атмосферы (прежде всего, направления основных переносов) и являются солярными реперами, отражающими особенности широтного распределения приходящей солнечной радиации в связи с сезонным изменением ее поступления. Географическая локализация зон высокого и низкого давления, а также климатологических полярных фронтов определяется значениями межширотного градиента приходящей за полугодия солнечной радиации на верхнюю границу атмосферы. Следовательно, основным фактором локализации субтропических областей высокого и субполярных областей низкого давления является наклон оси вращения Земли, с которым связано изменение в приходе солнечной радиации в каждое полушарие от зимнего сезона к летнему. С особенностями широтного распределения зон высокого и низкого атмосферного давления связаны барические градиенты, определяющие преобладающие направления переносов воздушных масс. В области низких широт, ограниченной субтропическими антициклонами преобладает восточный перенос. В области, расположенной между зоной субтропических антициклонов и субполярной зоной

низкого давления – западный. Максимальный поток момента количества движения к полюсам локализован в широтных диапазонах, разделяющих области восточного и западного переноса, что приблизительно соответствует положению субтропических зон высокого давления.

### 1.3.3 Межгодовая изменчивость приходящей радиации.

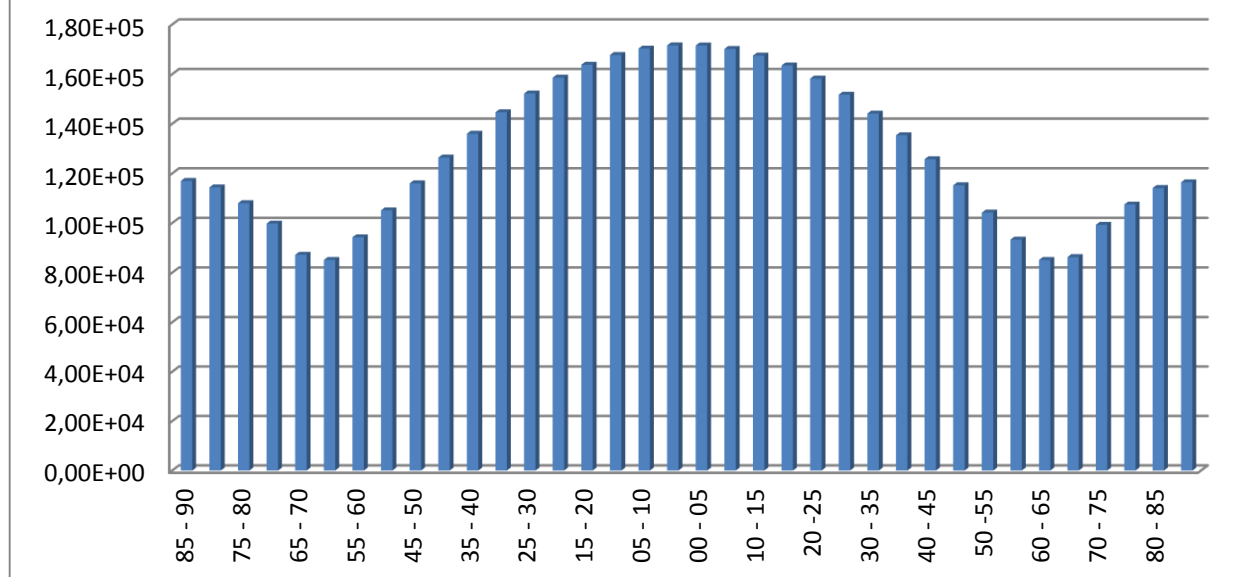
В работах связанных с разработкой схем и моделей ОЦА отмечается, что для более глубокого понимания механизма ОЦА необходимо решить две задачи. Во-первых, найти причину или причины межгодовой изменчивости и, во-вторых, понять, как эта временная изменчивость влияет на сезонные изменения и осредненную по времени циркуляцию.

Среднее значение амплитуды межгодовой изменчивости, рассчитанное нами для широтной зоны протяженностью в  $5^\circ$ , составляет  $1,3E+05$  Дж/м<sup>2</sup>. Общая энергия атмосферы приблизительно оценивается в  $6,5E+06$  Дж/м<sup>2</sup>. Таким образом, в среднем амплитуда межгодовой изменчивости приходящей на верхнюю границу атмосферы солнечной радиации составляет приблизительно 2% от общей энергии атмосферы (размах – 4%). Максимальные значения амплитуды межгодовой изменчивости составляют 2,7% и размах – 5,4% от энергии атмосферы. Отношение амплитуды межгодовой изменчивости в летнее полугодие составляет 1,7%, в зимнее – 0,8%. Вероятно, возможно резонансное усиление малых вариаций межгодовой изменчивости приходящей радиации в связи с тем, что в этой изменчивости отмечается четкая периодичность близкая и кратная к периоду годовых колебаний атмосферы. Другой причиной усиления (и диффузии) малых периодических вариаций солнечной радиации в климатической системе может быть стохастический резонанс – отклик бистабильной или метастабильной нелинейной системы на слабый периодический сигнал при шумовом воздействии определенной мощности.

Распределение рассчитанных (для тропических лет) значений амплитуды межгодовой изменчивости солнечной радиации (для верхней границы атмосферы) по широтным зонам представлено на рис. 1.3.3.1.

Величины межгодовой изменчивости находятся в пределах от  $8,0E+04$  Дж/м<sup>2</sup> до  $1,75E+05$  Дж/м<sup>2</sup>. Максимальные амплитуды, равные  $1,75E+05$  Дж/м<sup>2</sup>, отмечаются для экваториальных зон  $0^\circ - 5^\circ$  в каждом полушарии. Минимальные значения характерны для зон  $60^\circ - 65^\circ$  в каждом полушарии ( $-4,34E+04$  Дж/м<sup>2</sup>).

**Рисунок 1.3.3.1 Средняя для широтных зон амплитуда межгодовой изменчивости приходящей за год радиации, Дж/м<sup>2</sup>**

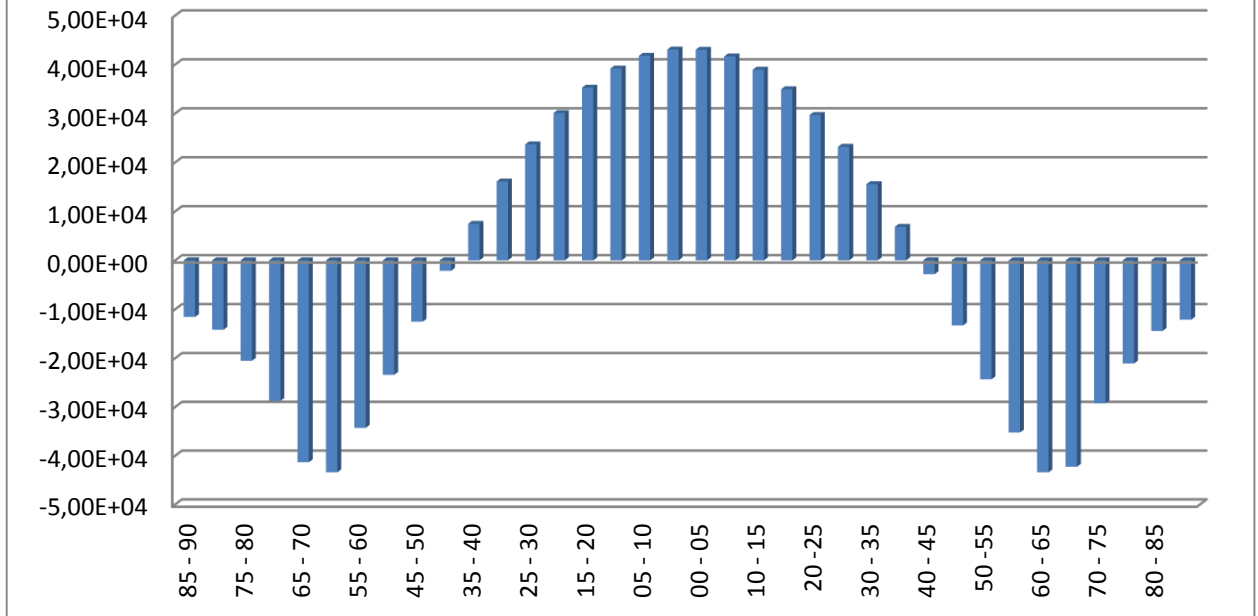


Анализ распределения отклонений амплитуды межгодовой изменчивости от среднего для широтной зоны значения показывает, что положительные отклонения характерны для широтного диапазона от 45° – 50° с.ш. до 45° – 50° ю.ш. (рис. 1.3.3.2). В зонах 40° – 45° градусов широты и в более высоких широтах в каждом полушарии отмечаются отрицательные отклонения амплитуды межгодовой изменчивости. Максимальные значения отклонений характерны для экваториальных зон 0° – 5° каждого полушария (4,3E+04 Дж/м<sup>2</sup>). Минимальные значения приходятся на широтные зоны 60° – 65° в каждом полушарии (-4,34E+04 Дж/м<sup>2</sup>).

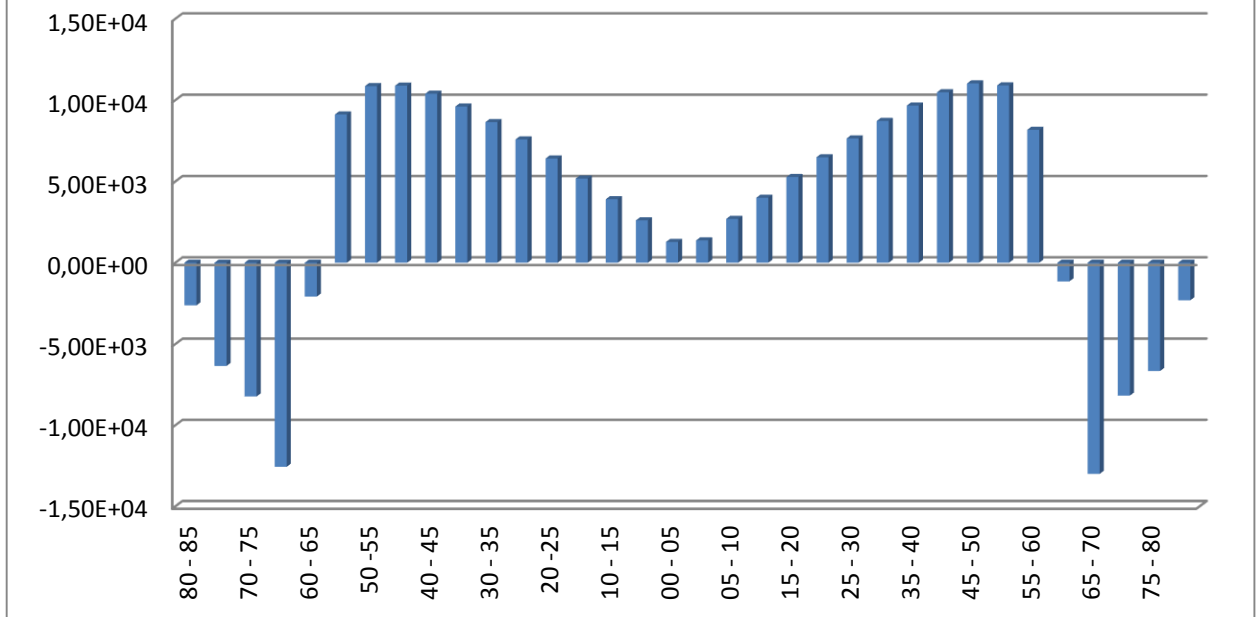
Межширотный градиент амплитуды межгодовой изменчивости рассчитывался аналогично межширотному градиенту приходящей на верхнюю границу атмосферы за тропический год солнечной радиации. Из полученного распределения видно, что в зонах 60° – 65° и в более высокоширотных зонах каждого полушария градиент имеет отрицательные значения (рис. 1.3.3.3). В зонах от 55° – 60° с.ш. до 55° – 60° ю.ш. – градиент положителен. Минимальные значения в этой области связаны с экватором – в зоне 0° – 5° ю.ш. значение градиента составляет 1,28E+03 Дж/м<sup>2</sup>. В зоне 0° – 5° с.ш. значение 1,37E+03 Дж/м<sup>2</sup>. Максимальные значения приходятся на

широтные зоны  $45^{\circ} - 50^{\circ}$  в каждом полушарии и составляют  $1,09E+04$  Дж/м<sup>2</sup> в южном полушарии и  $1,10E04$  Дж/м<sup>2</sup> – в северном.

**Рисунок 1.3.3.2 Отклонение от среднего значения амплитуды межгодовой изменчивости.**



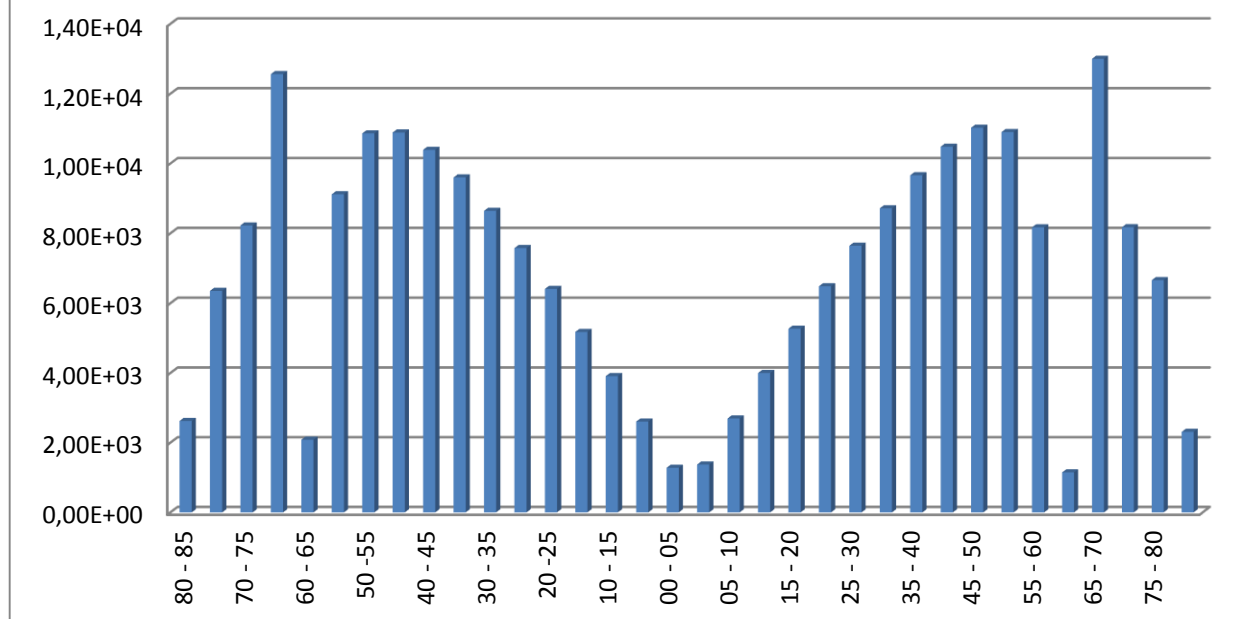
**Рисунок 1.3.3.3 Межширотный градиент изменения амплитуды межгодовой изменчивости.**



Минимальные, по модулю, значения градиента амплитуды межгодовой изменчивости проходящей радиации характерны для экваториальной области ( $0^\circ \pm 5^\circ$ ), полярных областей ( $85^\circ - 90^\circ$  в каждом полушарии) и для широтных зон  $60^\circ - 65^\circ$  каждого полушария (рис. 1.3.3.4). Такой характер распределения определяется картиной наложения амплитуд межгодовой изменчивости различных временных циклов в межгодовой изменчивости проходящей радиации. Максимальные амплитуды 19-ти летнего нутационного цикла отмечаются в полярных зонах, максимальные амплитуды 2-х и 3-х летней изменчивости – в экваториальной области. Минимальная изменчивость амплитуд приходится, таким образом, на зоны максимальных значений амплитуд 19-ти, 2-х и 3-х летней изменчивости (полярные и экваториальная), а также на зону минимальных значений амплитуд межгодовой изменчивости (рис. 1.3.3.1). Изменчивость амплитуды в нутационном цикле уменьшается от максимальных значений в полярных областях до минимальных значений в зонах  $60^\circ - 65^\circ$ . Вариации амплитуды в 2-х и 3-х летних циклах также сокращаются от максимальных значений в экваториальных районах к минимальным в зонах  $60^\circ - 65^\circ$ . Следовательно, широтные зоны  $60^\circ - 65^\circ$  являются зонами сопряжения (схождения в минимумах) минимальных вариаций амплитуд межгодовой изменчивости проходящей радиации во временных циклах, определяемых периодическими возмущениями орбитального движения Земли.



**Рисунок 1.3.3.4 Модуль межширотного градиента изменения амплитуды межгодовой изменчивости.**



Таким образом, солярными реперами особенностей ОЦА являются характеристики распределения приходящей на верхнюю границу атмосферы солнечной радиации, связанные с широтными зонами локализации особенностей ОЦА. Особенности в распределении солнечной радиации отражают широтную изменчивость в годовом поступлении к Земле, сезонную изменчивость в различных полушариях и межгодовую изменчивость солнечной радиации.

Локализация экваториальной депрессии, ВЗК, климатологических полярных фронтов, тропических и полярных воздушных масс связана в основном с характеристиками годового распределения приходящей на верхнюю границу солнечной радиации и ее межгодовой изменчивости. Исключение составляет только один сезонный репер.

Так среднее положение экваториальной депрессии и ВЗК характеризуется пространственным соответствием годовым солярным характеристикам: диапазону максимальных значений и диапазону минимальных значений межширотного градиента приходящей за тропический год радиации; диапазону максимальных значений амплитуды межгодовой изменчивости, диапазону минимальных (по модулю) значений межширотного градиента амплитуды межгодовой изменчивости

и диапазону максимальных отклонений от среднего амплитуды межгодовой изменчивости солнечной радиации .

Для климатологических полярных фронтов характерно пространственное соответствие диапазонам максимальных значений межширотного градиента приходящей за год радиации, диапазону максимальных значений межширотного градиента в зимнее для полушария полугодие и диапазонам максимальных значений межширотного градиента амплитуды межгодовой изменчивости.

Географическая локализация тропических воздушных масс пространственно соответствует диапазону положительных значений отклонения от среднего годового значения приходящей радиации и диапазону положительных отклонений от среднего значения амплитуды межгодовой изменчивости.

Положение полярных воздушных масс связано с широтными диапазонами отрицательных значений отклонения от среднего годового значения приходящей радиации и отрицательных отклонений от среднего значения амплитуды межгодовой изменчивости.

Локализация субтропических зон высокого давления связана только с сезонными особенностями в распределении радиации. Положение субполярных зон низкого давления характеризуется пространственным соответствием с характеристиками приходящей радиации в летнее полугодие и характеристиками межгодовой изменчивости.

Субтропические зоны высокого давления пространственно соответствуют широтным диапазонам максимальных значений приходящей за летнее и зимнее полугодия радиации, диапазонам смены знака отклонений от среднего для полугодия значения приходящей радиации и диапазонам смены знака межширотного градиента приходящей за полугодия радиации.

Субполярные зоны низкого давления пространственно связаны с диапазонами минимальных значений межширотного градиента приходящей в летнее для полушария полугодие радиации, с диапазонами минимальных значений амплитуды межгодовой изменчивости, с диапазонами минимальных (по модулю) значений межширотного градиента амплитуды межгодовой изменчивости и с диапазоном

минимальных отклонений от среднего значения амплитуды межгодовой изменчивости.

#### **1.4 Условия задачи.**

Нам даны данные по давлению и приходящей радиации на Азорских островах и Исландии. Поставленные перед нами задачи:

1. Найти связь годового хода давления и радиации.
2. Составить многолетний ход по среднемесячным значениям.

## **2 Модель прогнозирования.**

### **2.1 Предпосылки к использованию модели прогнозирования**

Для обработки имеющихся у нас данных с целью прогнозирования дальнейшей ситуации вычислительных мощностей человеческого мозга определённно недостаточно. Это приводит нас к необходимости использования дополнительных вычислительных мощностей в виде системы поддержки принятия решения. Перед данными системами ставится задача помочь людям, принимающим решение, в условиях большого количества данных и сложных реакций управляемой системы на проявление внешних факторов. Детально и качественно анализируя предметную область, системы поддержки принятия решений должны оперативно предоставлять пользователю полную и объективную информацию о наблюдаемой или управляемой системе, спрогнозированные показатели, варианты оптимальных или необходимых решений для разных условий. Большинство таких систем находится в постоянном развитии - их последующее состояние зависит от ряда предыдущих состояний и от последовательности их прохождения. В системах могут наблюдаться тенденции к периодическим изменениям и цикличности. Исходя из этого, при анализе системы, помимо статистического разнообразия значений параметров, следует учитывать их взаимосвязи со временем. Таким образом, информацию о состоянии системы в прошлом и настоящем можно рассматривать как наборы временных рядов.

Чтобы проанализировать предметную область и сформировать решение, системы поддержки принятия решений (СППР) могут применять различные методы, такие как: поиск информации, знаний или закономерностей в базах данных, рассуждение на основе прецедентов, ситуационное моделирование и ситуационный анализ, интеллектуальный анализ информации, моделирование субъективных представлений экспертов о ситуации, нейронные сети, генетические алгоритмы и др.

Все вышеупомянутые возможности делают СППР отличным выбором для реализации задачи прогнозирования климатических явлений.

## **2.2 Общие сведения об СППР.**

### **2.2.1 Классификация СППР**

Вследствие отсутствия строгого единого определения для систем поддержки принятия решений, мы также сталкиваемся с отсутствием общепринятой классификации. Наиболее общая классификация ранжирует СППР в соответствии с основными выполняемыми ими операциями, от ориентированных на данные систем до модель-ориентированных систем (Alter S.L. 1980). Данная классификация была основана на анализе 56 различных СППР и включает в себя следующие семь типов :

- Файловая система, предоставляющая пользователю удобный доступ к элементам данных.
- Системы анализа данных, позволяющие обрабатывать данные с использованием автоматизированных приложений, специализированных для определенных задач или общих методов.
- Системы анализа данных, которые обеспечивают доступ к ряду ориентированных на поиск решений баз данных и малых моделей.
- Бухгалтерские и финансовые модели, целью которых является просчет последствий возможных действий.
- Наглядные (репрезентативные) модели, которые оценивают последствия действий на основе имитационных моделей.
- Оптимизационные модели, целью которых является выработка основных поведенческих принципов, путем создания оптимального решения в соответствии с серией ограничений.
- Модели-советчики, которые выполняют логическую обработку, позволяющую найти конкретное решение для хорошо структурированной или глубоко изученной задачи.

Кроме данной системы типов исследователями СППР были разработаны и другие, основанные на различных критериях классификации: по типу взаимодействия с пользователем, по типу оказываемой поддержки, и другим.

Следует заметить, что исходя из проведенных исследований, количество производимой человечеством информации увеличивается более чем в два раза за

период времени равный двум годам. В связи с этим все труднее искать и получать нужные данные из общего информационного потока. Для того чтобы справиться с этой задачей сейчас активно развивается новая область Data Mining, которая стремится решить проблему поиска новых знаний, необходимой информации и взаимного влияния различных факторов в больших массивах неструктурированных данных. Также для решения задачи извлечения информации из огромных объемов неструктурированных данных значительного многообразия активно развивается новый класс подходов, методов и инструментов - "Большие данные" (Big Data). Данный класс противопоставляет себя стандартным системам управления базами данных за счет использования средств массово-параллельной обработки постоянно увеличивающихся в масштабе неструктурированных данных.

Одним из наиболее сложных и трудоемких процессов является процесс получения знаний на первых шагах создания интеллектуальных экспертных систем (ЭС) и СППР. Этот процесс не всегда достигает желаемого результата в силу того, что для слабоструктурированных предметных областей базы данных часто содержат разнотипную, неполную, а иногда и нечеткую, и даже противоречивую информацию. Класс слабоструктурируемых проблем (ill-structured) затрагивает широкую область реальных ситуаций, которым присущи следующие признаки: неточные двусмысленные определения, изменяющиеся понятия, большая зависимость ситуаций от контекста, наличие противоречивости и неполноты имеющихся данных, разнотипность неоднозначность, ненадежность и неопределенность информации. Исходя из вышесказанного, привлечение интеллектуальных методов Data Mining для решения существующих проблем является одним из перспективных решений. В рамках области Data Mining исследователи различают 6 различных типов задач:

- поиск зависимостей,
- классификация по группам,
- выявление аномалий и отклонений,
- регрессионный анализ,
- обобщение,

- нахождение шаблонов возникновения подпоследовательностей в последовательностях данных.

Последний тип задач все чаще привлекает к себе новых исследователей, так как он находит широкое применение в анализе действий пользователей интернета, нахождении лингвистических зависимостей в текстах на естественных языках, анализе истории болезней для определения возможных заболеваний, определении аминокислотной или нуклеотидной последовательности ДНК, а также в предсказании биржевых показателей.



### **2.2.2 Основные определения**

Под термином оболочки СППР в данной работе понимается "пустая" экспертная система, т.е. экспертная система без привязки к конкретной предметной области и базы знаний. Преимущество оболочки в том, что для создания из нее готовой экспертной системы она не требует изменения кода. Необходимо лишь ее обучение специалистом по предметной области или на основе заполнения базы знаний. Однако, в случае если предметная область плохо укладывается в используемую в оболочке модель, обучить ее будет трудно.

Временной ряд – собранный в разные моменты времени статистический материал о значении каких-либо параметров исследуемого процесса. Учитывается взаимосвязь изменений со временем или порядок статистического разнообразия и характеристики выборки.

### 2.2.3 Классификация временных рядов

Одной из наиболее перспективных и сложных задач является прогнозирование временных рядов - наблюдаемых в разные моменты времени статистических данных о значении каких-либо переменных изучаемого процесса. Каждая единица статистических данных называется отсчетом или измерением, также допустимо называть его уровнем на соответствующий ему момент времени. Принципиальным отличием временного ряда от обычной выборки заключается в том, что помимо статистических характеристик и разнообразия выборки при анализе учитываются связь данных со временем. Задачи прогнозирования временных рядов встречаются во многих сферах жизнедеятельности человека.

Существует несколько различных классификаций временных рядов. Рассмотрим основные из них:

1. По характеру представления измерений:

- ряды средних величин;
- относительных показателей;
- абсолютных показателей.

2. По типу временного параметра:

- Моментные временные ряды. Для временных рядов такого типа измерения отвечают за состояние наблюдаемых величин в определенные моменты времени.

- Интервальные временные ряды. Для интервальных рядов измерения отвечают за состояние наблюдаемых величин за определенные временные промежутки.

3. По длине промежутка между моментами времени, к которым относятся измерения:

- полные (равноотстоящие) - когда измерения следуют друг за другом через равные промежутки времени или моменты начала и окончания измерений отстоят друг от друга всегда на равный временной интервал;

- неполные (неравноотстоящие) - в том случае, если описанные выше условия не выполняются.

4. По способу формирования данных временные ряды делятся на:

- детерминированные - являющиеся значениями какой-либо некоторой неслучайной функции от времени (ряд упорядоченной по месяцам информации о продолжительности месяца в днях);

- случайные - являются результатом реализации некоторого случайного процесса.

5. В зависимости от наличия основной тенденции:

- стационарные ряды - не меняющие свои основные характеристики со временем (математическое ожидание постоянно и автокорреляционная функция процесса зависит только от величины интервала между измерениями);

- нестационарные - не отвечающие вышеприведенному критерию.

Говоря о прогнозировании временных рядов, в данной работе имеется ввиду задача изучения и анализа временного ряда наблюдений над некоторыми параметрами исследуемого процесса для обучения на их основе интервальной нейронной сети. В дальнейшем с ее помощью находить интервалы значений прогнозируемой величины на заданный момент в будущем (к моменту в времени  $t + \Delta t$ ). Величина  $\Delta t$  именуется глубиной прогнозирования.

## **2.2.4 Анализ общих методов поддержки принятия решений**

Для анализа большого объема информации в интересах прогнозирования поведения управляемой системы и выработки управляющего решения разработаны и используются различные методы. Рассмотрим основные из них:

1. Информационный поиск - процесс поиска среди большого количества текстов или документов таких, которые содержат информацию по интересующему предмету, соответствуют заданному условию поиска или содержат в себе необходимые данные, факты, сведения.

Обычно, информационный поиск состоит из трех шагов:

- Формулировка критериев поиска и создание запроса;
- Выявление документов, в которых может содержаться интересующая информация;
- Детальный поиск в отобранных документах и оценка результатов поиска.

2. Деревья принятия решений - представляет собой граф с одной начальной вершиной. На ребрах которого записаны атрибуты, от которых зависит целевая функция. В конечных вершинах (из которых не выходят ребра) указаны значения целевой функции, а в остальных вершинах указываются атрибуты, по которым различаются случаи. Чтобы получить решение с использованием дерева, необходимо пройти по нему до конечной вершины (листа) и подсчитать соответствующее значение с помощью целевой функции.

3. Кластерный анализ - набор различных алгоритмов классификации, позволяющих на основе сходства частных признаков или набора признаков разбивать на отдельные группы наблюдаемые наборы признаков (ситуации). В случае правильного применения алгоритма кластеризации, после его работы справедливо следующее утверждение - для любого набора признаков, набор оказавшийся с ним в одном кластере более схож с ним, чем любой набор из другого кластера. А значит и решение, принимаемое по наблюдаемому набору признаков должно быть схоже с решениями в данном кластере.

4. Рассуждение на основе прецедентов - подход предполагающий решение новых задач за счет использования и/или адаптации решений для уже известных задач.

Обычно, формирование решения на основе прецедентов состоит из четырех шагов:

- Выбор схожего прецедента из базы данных;
- Адаптация решения по найденному прецеденту к решаемой задаче;
- Проверка созданного решения;
- Добавление в базу данных информации по решенной задаче вместе с ее решением.

5. Имитационное моделирование - это метод, направленный на изучение наблюдаемой реальной системы путем замены ее достаточно точной моделью. Проводя эксперименты с моделью, исследователь (лицо, принимающее решение) может без каких-либо рисков изучить какие факторы и как влияют на систему, а также смоделировать результат того или иного решения с целью выявить наилучшее.

6. Генетические алгоритмы - алгоритмы для решения сложных неформализованных задач, применяются при очень больших размерах задач и при отсутствии упорядоченности в исходных данных. Данные алгоритмы подражают эволюционной теории Дарвина, осуществляя поиск решения путем последовательного улучшения наборов потенциальных решений. Улучшение каждого следующего набора происходит за счет скрещивания и мутаций лучших представителей предыдущего набора решений.

7. Когнитивное моделирование - метод, применяющийся для поиска решения в слабоструктурированных системах с многоаспектными взаимосвязанными процессами, без достаточного количества информации о развитии процессов.

В данной модели информация о системе представляется в виде набора понятий и связывающей их причинно-следственной сети, называемой когнитивной картой. Данная карта является отражением субъективных представлений эксперта о законах и закономерностях, присущих моделируемой системе. Для получения

оптимальной стратегии управления при помощи когнитивной карты используют методы аналитической обработки, направленные на исследование структуры системы и получение прогнозов её поведения при различных стратегиях управления.

8. Статистические методы прогнозирования - математические методы прогнозирования, осуществляющие прогнозирование на основе объективных данных о наблюдаемом процессе.

Статистические методы ставят перед собой задачу нахождения некоторого функционала от наблюдаемых параметров процесса, значения которого являлись бы прогнозом. От этого функционала требуется минимизация математического ожидания квадрата нормы разности между его значением в некоторый будущий момент времени и реальным значением прогнозируемой величины. В зависимости от конкретного метода прогнозирования на функционал накладываются различные ограничения.

9. Нейронные сети - методы, моделирующие биологическую модель нервных систем для решения задач прогнозирования, классификации или управления. Достигается это путем воспроизведения достаточно сложных зависимостей с большим числом переменных на входе.

Основным преимуществом нейронных сетей является способность к обучению. Благодаря этому нейронные сети способны автоматически приобретать знания, но не способны эти знания объяснить. Обученная нейронная сеть представляется пользователю некоторым черным ящиком.

Одной из наиболее популярных нейросетевых моделей является многослойный персептрон, основанный на послойном взаимодействии искусственных нейронов. Каждый нейрон слоя получает на вход сумму взвешенных выходов нейронов предыдущего слоя. На выходе каждый нейрон имеет значение своей функции активации от входа.

Для обучения многослойного персептрона применяют алгоритм обратного распространения ошибки, основанный на методе градиентного спуска.

10. Системы с нечёткой логикой - системы, работающие с лингвистическими переменными и аппаратом нечеткой логики, допускающим принятие функцией принадлежности значений в интервале от 0 до 1 и логические операции над нечеткими множествами.

Системы с нечеткой логикой хорошо подходят для объяснения получаемых с их помощью результатов, но не могут сами приобретать знания на основе которых должно приниматься решение.

Системы с нечеткой логикой имеет смысл использовать в следующих случаях:

- для сложных процессов, в отсутствии других стандартных математических моделей;
- если знания эксперта о процессах или системе можно сформулировать только в лингвистической форме.

Основными недостатком систем с нечеткой логикой является то, что набор нечетких правил и вид функции принадлежности задается субъективно человеком-экспертом и может оказаться неправильным или неполным.

11. Экспертные методы прогнозирования - методы, опирающиеся в прогнозировании на оценки людей-экспертов.

Одним из распространенных экспертных методов является метод Дельфи. Суть метода состоит в опросе различных экспертов и обобщения их мнений в единую оценку. Для прогнозирования этим методом каких-либо показателей необходимо выделить экспертную группу специалистов в этой области, провести несколько опросов, анкетирований или мозговых штурмов, проверить согласованность мнений экспертов, проанализировать полученные выводы и сформировать итоговое решение.

## 2.2.5 Актуальность создания алгоритма прогнозирования временных рядов.

Рассмотрим подробнее наиболее распространенные статистические методы прогнозирования временных рядов:

Таблица 2.2.5.1.

### Распространенные статистические методы прогнозирования временных рядов

Статистический метод	Предпосылки для применения	Пример модели
Экстраполяционные методы прогнозирования	Во временном ряду наблюдается тренд или долгосрочные тенденции	$y = a_0 + a_1t + a_2t^2 + \dots + a_nt^n + \varepsilon$ (линейная регрессия)
Методы корреляционного анализа данных	Наблюдается зависимость между значениями разных временных рядов	$X_1 = \mu_1 + \rho \frac{\sigma_1}{\sigma_2} (X_2 - \mu_2)$
Модели, в параметры которых закладывается один из видов интервенции (устойчивое скачкообразное, устойчивое постепенное, скачкообразное временное)	Временной ряд с резкими изменениями тенденции процесса под внешним или внутренним воздействием (обычно внешним)	$y = a_0 + a_1t + a_2t^2 + w * 1 \{t_1 < t < t_2\}$ (степенной полином со скачкообразным временным воздействием)
GARCH модели - обобщенная авторегрессионная условно гетероскедастичная модель	Наблюдения с большими и малыми отклонениями от средних имеют тенденцию к образованию кластеров	$y_t = x'_t \beta + u_t$ $\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i u_{t-i}^2 + \sum_{j=1}^q \gamma_j \sigma_{t-j}^2$ GARCH(p,q)



Приведем краткую сравнительную характеристику методов, позволяющих строить оценки параметров или осуществлять прогнозирование, исключив из рассмотрения остальные:

**Таблица 1.2.5.2**

**Сравнение методов прогнозирования временных рядов**

Метод	Требование к квалификации пользователя	Возможность самостоятельного извлечения знаний из данных	Конечное количество возможных решений
Деревья принятия решений	небольшие	отсутствует	да
Кластерный анализ	средние	присутствует	да
Рассуждения на основе прецедентов	средние	в малой степени	да
Статистические методы прогнозирования	большие	в малой степени	нет
Нейронные сети	небольшие	присутствует	нет
Экспертные методы	средние	присутствует	нет

Рассмотрев существующие методы можно прийти к следующим выводам:

- Существующие методы испытывают трудности в нахождении переменных оказывающих наибольшее влияние на прогнозируемую величину. Особенно при наличии большого числа таких переменных и сложных нелинейных зависимостей между ними и прогнозируемой величиной.

- Редкие методы, помимо статистических, позволяют работать с недоопределенными данными.
- Не существует легкого в применении универсального (не зависящего от конкретной предметной области) метода, требующего небольших временных затрат.

## **2.3 Методы реализации алгоритмов прогнозирования временных рядов**

### **2.3.1 Использование интервальных нейронных сетей**

В работе доказывается, что любая непрерывная функция  $n$  переменных может быть с наперед заданной точностью приближена нейронными сетями с использованием любой непрерывной функции активации одного переменного.

Являясь моделью сложной многомерной нелинейной регрессии, нейронная сеть потенциально превосходит по точности вышеперечисленные методы, а также обладает рядом преимуществ:

а) Возможность работы с шумовыми и неинформативными входными сигналами – нейронная сеть самостоятельно может определить их непригодность для решения поставленной задачи и явно отбросить их, обнулив соответствующие коэффициенты.

б) Возможность получать в качестве входных данных разнотипную информацию – непрерывнозначную и дискретнозначную, качественную и количественную, что представляется нелегкой задачей для многих методов.

в) Возможность одновременно решать несколько задач при наличии у нейронной сети нескольких выходов.

г) Существуют алгоритмы, позволяющие решать обратные задачи, используя нейронные сети, обученные для решения прямой задачи. Например, можно на выходе нейронной сети подключить новую нейронную сеть и обучить её таким образом, чтобы на выходе были входные данные для первой нейронной сети. Таким образом, можно получить решение задачи, которую остальные методы могут решить лишь в обратную сторону.

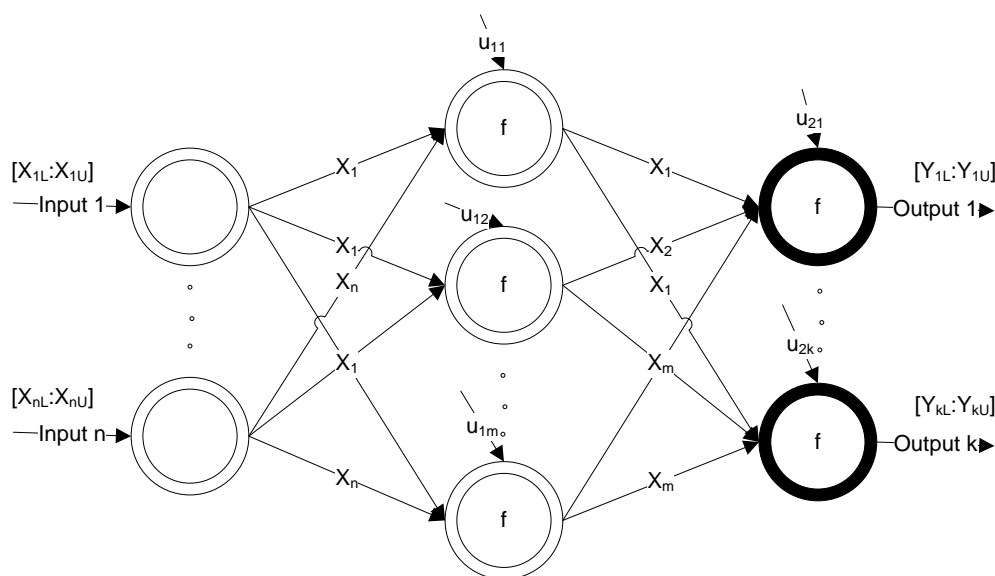
д) Нейронная сеть имеет меньшие требования к квалификации пользователя, чем, например, сложные статистические модели способные давать аналогичные результаты.

е) Изначально задав синаптические веса нейронной сети, можно воссоздать и проверить предполагаемые статистические модели, а также улучшить их путем тренировки сети.

### 2.3.2 Интервальные нейронные сети: основные определения и принцип работы.

В случае прогнозирования интервальных значений необходимо использовать интервальные нейронные сети. Как наглядно проиллюстрировано в работе профессора Ishibuchi, при применении «обычных» НС могут возникнуть ошибки прогнозирования, связанные с превышением нижней границы прогнозируемого интервала над верхней.

Интервальная нейронная сеть – это система соединённых и взаимодействующих между собой интервальных нейронов, имеющих на входе и на выходе значения заданные в виде интервала (не одно значение, а континуальное множество значений в промежутке между парой значений, задающей границы интервала).



**Рисунок 2.3.2.1** Схема интервальной нейронной сети с одним скрытым слоем ( $n$  – число входных нейронов,  $m$  – число скрытых нейронов,  $k$  – число выходных нейронов)

На каждый из входных нейронов подается интервальное значение, на основе которого планируется получить прогноз. У входных нейронов на выходе находится то же значение, что и на входе. Для всех остальных нейронов, на выходе мы получаем интервальные значения, получаемые по формуле:

$$Y = f\left(\sum_{i=1}^N w_i X_i + u_i\right),$$

где  $Y$  - выход нейрона сети,  $w_i$ - вес  $i$ -го входа (для каждого нейрона свой),  $f$ - функция активации,  $N$  - число входов нейрона,  $u_i$  - вес смещения (у каждого нейрона свой),  $X_i$  – интервальное значение, поданное на  $i$ -ый вход нейрона. Значение  $Y$  вычисляется с учетом правил интервальной арифметики:

$$A * B = [a_L, a_U] * [b_L, b_U] = [\min\{a * b \mid a \in [a_L, a_U], b \in [b_L, b_U]\}, \max\{a * b \mid a \in [a_L, a_U], b \in [b_L, b_U]\}],$$

где  $*$  – любая операция. Например:

$$A + B = [a_L, a_U] + [b_L, b_U] = [a_L + b_L, a_U + b_U],$$

$$A - B = [a_L, a_U] - [b_L, b_U] = [a_L - \max\{b_L, b_U\}, a_U - \min\{b_L, b_U\}],$$

$$A \cdot w = [a_L, a_U] \cdot w = \begin{cases} [a_L \cdot w, a_U \cdot w], & w \geq 0 \\ [a_U \cdot w, a_L \cdot w], & w < 0 \end{cases}$$

$$A \cdot B = [a_L, a_U] \cdot [b_L, b_U] = [\min\{a_L b_L, a_U b_U, a_L b_U, a_U b_L\}, \max\{a_L b_L, a_U b_U, a_L b_U, a_U b_L\}].$$

Использование интервальных значений позволяет прогнозировать и использовать в качестве входных параметров:

- Стандартные интервальные параметры (цена открытия/закрытия, покупки/продажи),

- Интервальные значения, позволяющие сократить количество входных значений (вместо ежеминутного курса пускать на вход сети интервал от минимальной до максимальной цены за час или день),
- Величины, значения которых из-за неточности измерительных приборов искажены ошибками округления,
- Случайные величины (принимающие значения на интервале).
- Обычные – не интервальные параметры, приравняв нижнюю и верхнюю границу интервала.

### 2.3.3 Использование генетических алгоритмов

Если на вход нейронной сети подать значения всех известных параметров, то это значительно замедлит скорость ее работы, а так же увеличится вероятность нахождения сетью реально не существующих зависимостей. С другой стороны, подавая на вход нейронной сети лишь часть параметров, кажущихся лицу, принимающему решение, наиболее значимыми, можно упустить из виду параметры, реально вносящие вклад в значение прогнозируемой величины.

С целью решить эту задачу было разработано несколько алгоритмов. Наиболее популярный для применения в нейронных сетях – алгоритм, основанный на последовательном увеличении количества входных параметров при помощи добавления наиболее значимого, пока добавление нового параметра будет давать заметный результат. На первом шаге алгоритма для каждого параметра обучается своя сеть с единственным входом. Параметр, на основе которого сеть дает наименьшую ошибку при прогнозировании, считается наиболее значимым и добавляется к числу значимых. На каждом следующем шаге для каждого параметра, не входящего в число значимых, обучается сеть, которой на вход подаются все значимые плюс данный параметр. Параметр, "чья" сеть дает наименьшую ошибку, при прогнозировании добавляется к числу значимых. Алгоритм повторяется до тех пор, пока добавление нового значимого параметра уменьшает ошибку прогнозирования.

Корреляционный анализ не подходит в качестве метода обработки входных параметров для решения поставленной задачи, так как корреляция отражает лишь линейную зависимость величин, но не отражает их функциональной связности. Например, если вычислить коэффициент корреляции между величинами  $A = \sin(x)$  и  $B = \cos(x)$ , то он будет близок к нулю, то есть (линейная) зависимость между величинами отсутствует. Между тем, величины  $A$  и  $B$  очевидно связаны функционально по закону  $\sin^2(x) + \cos^2(x) = 1$ .

В качестве метода решения данной задачи нами был выбран генетический алгоритм. Генетический алгоритм - адаптивный метод поиска, который все чаще используется для решения задач функциональной оптимизации. Он основан на

генетических процессах биологических организмов: биологические популяции развиваются в течение нескольких поколений, подчиняясь законам естественного отбора и принципу "выживает наиболее приспособленный" (survival of the fittest), открытому Чарльзом Дарвином. Аналогично этому процессу, генетические алгоритмы способны "развивать" - улучшать решения реальных задач, если те закодированы соответствующим образом. Они оперируют наборами "особей" - популяцией, где каждая особь представляет собой возможное решение поставленной задачи. Каждой особи присваивается мера ее приспособленности в соответствии с тем, насколько хорошо представленное ей решение задачи. Оценивание качества решения осуществляется при помощи функции приспособленности (целевой функции).

Наиболее приспособленные особи отбираются для создания нового поколения (воспроизведения потомства) с помощью перекрестного скрещивания с другими особями популяции. В результате этого появляются новые особи, которые сочетают в себе некоторые характеристики, унаследованные ими от родительских особей. Из-за того что вероятность наименее приспособленных особей быть отобранными для воспроизведения потомков меньше чем у остальных, следует, что те свойства, которыми они обладали, будут постепенно исчезать из популяции в процессе эволюции. Так воспроизводится каждое новое поколение допустимых решений, выбирая более качественных представителей предыдущего поколения, скрещивая их и получая множество новых решений. Каждое новое поколение, по отношению к предыдущему, содержит большее количество характеристик, которыми обладают хорошие члены предыдущего поколения. Благодаря этому, из поколения в поколение, характеристики более качественных особей распространяются по всей популяции. Таким образом, скрещивание наиболее приспособленных особей позволяет исследовать наиболее перспективные направления пространства поиска. В конечном итоге, популяция будет сходиться в среднем к оптимальному решению задачи.



**Таблица 2.3.2.1 Сравнительная характеристика методов**

Метод	Нахождение линейной зависимости	Нахождение функциональной зависимости	Скорость получения первых значимых результатов
Последовательное увеличение числа параметров	да	да	Низкая
Корреляционный анализ	да	нет	Высокая
Генетический алгоритм	да	да	Высокая

Опишем более подробно принцип работы генетического алгоритма для взаимодействия с нейронными сетями.

Начальные поколения наборов параметров (особей) для генетического алгоритма определяются случайным образом. Далее, наиболее приспособленными особями считаются наборы, обученные на которых сети дают минимальные ошибки. Новое поколение особей получается путем скрещивания наиболее приспособленных особей предыдущего поколения и мутации.

Отбор особей для скрещивания осуществляется с использованием "рулеточного отбора". Данный вид отбор представляет собой вероятностный метод случайного выбора особей, где каждая особь имеет априорную вероятность быть выбранной, соотносящуюся со значением ее функции приспособленности. Каждой особи ставится в соответствие относительная приспособленность - число из диапазона  $[0;1]$ , определяемое отношением значения ее целевой функции к сумме значений целевых функций всех особей популяции. После определения отношений отрезок  $[0;1]$  разбивается на число отрезков равное количеству особей в популяции так, что длина каждого отрезка соответствует относительной приспособленности

соответствующей ему особи. Необходимое количество раз запускается генератор непрерывной случайной величины, равномерно распределенной на интервале [0,1] (базовой случайной величины). Особи для скрещивания выбираются в соответствии с отрезками, на которые попали сгенерированные величины.

Преимуществом данного метода является возможность использования для следующего поколения всех особей, сохраняя при этом высокие шансы за более приспособленными. Это позволяет периодически разбавлять генофонд и, благодаря этому, предотвращать сходимость к локальному экстремуму целевой функции.

Скрещивание (кроссовер) происходит следующим образом - случайно определяется точка внутри хромосомы (набора параметров), называемая точкой разрыва, в которой обе хромосомы делятся на две части. Соединив первую часть одной хромосомы со второй частью другой и наоборот, в результате получают две новые хромосомы. На каждом цикле, для предотвращения преждевременной сходимости, с некоторой низкой вероятностью осуществляется мутация особей, которая заключается в изменении отдельного гена-параметра в особи на другой из общего набора параметров.



**Рисунок 2.3.3.1 Иллюстрация процедуры кроссовера**



**Рисунок 2.3.3.2 Иллюстрация процедуры мутации**

Получение новых поколений продолжается до тех пор, пока популяция не сойдется или не будет превышено заданное пользователем максимальное количество поколений.

### 2.3.4 Иллюстрация работы генетического алгоритма

Продemonстрируем принцип работы генетического алгоритма на следующем простом искусственном примере:

Рассмотрим задачу нахождения хромосомы, которая имеет максимальное количество единиц. Пусть хромосомы состоят из 12 генов, а каждая популяция состоит из 8 хромосом. Следовательно, наилучшая хромосома будет состоять из 12 единиц. Рассмотрим решение этой задачи с использованием генетического алгоритма.

Создание или генерация начальной популяции. Для этого потребуется сгенерировать 8 двоичных последовательностей, каждая длиной по 12 бит. Их можно получить, например, разыгрывая базовую случайную величину 96 раз (при получении значения из интервала  $[0;0.5]$  приписывается значение 1, в случае значения из интервала  $(0.5,1]$  - 0) или использованием любого генератора двоичных случайных чисел. Сформируем исходную популяцию, используя любой из предложенных методов:

$$\begin{array}{ll} \text{ch}_1 = [010111101010]; & \text{ch}_2 = [110000101101]; \\ \text{ch}_3 = [111000001001]; & \text{ch}_4 = [100010000001]; \\ \text{ch}_5 = [001101011000]; & \text{ch}_6 = [100010000101]; \\ \text{ch}_7 = [001000001010]; & \text{ch}_8 = [110011010110]. \end{array}$$

Оценим приспособленность хромосом в созданной популяции. Так как в нашем простом примере мы ищем такую хромосому, которая содержит максимальное число единиц. Следовательно, функция принадлежности должна зависеть от количества единиц в хромосоме. Будем обозначать функцию принадлежности символом  $F$ . А за ее значение брать сумму генов хромосомы. В этом случае значения для каждой хромосомы из полученной популяции будут следующие:

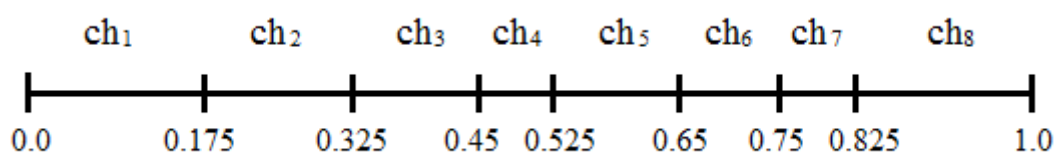
$$F(\text{ch}_1) = 7; \quad F(\text{ch}_2) = 6;$$

$$\begin{array}{ll}
 F(\text{ch}_3) = 5; & F(\text{ch}_4) = 3; \\
 F(\text{ch}_5) = 5; & F(\text{ch}_6) = 4; \\
 F(\text{ch}_7) = 3; & F(\text{ch}_8) = 7.
 \end{array}$$

Хромосомы  $\text{ch}_1$  и  $\text{ch}_8$  имеют наибольшие значения функции приспособленности. В данной популяции они являются основными претендентами на решение задачи. Если условие остановки алгоритма не выполняется, то на следующем шаге производится отбор особей текущей популяции для последующего скрещивания.

Селекция производится рулеточным методом. Для каждой из 8 хромосом исходной популяции получаем сектора единичного отрезка, выраженные отношением значения функции приспособленности отдельной особи к сумме значений функции приспособленности всех особей (рисунок 2.3.4.1).

$$\begin{array}{ll}
 v(\text{ch}_1) = 0.175; & v(\text{ch}_2) = 0.15; \\
 v(\text{ch}_3) = 0.125; & v(\text{ch}_4) = 0.075; \\
 v(\text{ch}_5) = 0.125; & v(\text{ch}_6) = 0.1; \\
 v(\text{ch}_7) = 0.075; & v(\text{ch}_8) = 0.175.
 \end{array}$$



**Рисунок 2.3.4.1** Разбиение отрезка для рулеточного отбора

Розыгрыш с помощью построенного разбиения сводится к случайному выбору числа из интервала  $[0,1]$ , указывающего на соответствующий сектор отрезка, т.е. на соответствующую этому отрезку конкретную хромосому.

Разыграем 8 случайных чисел:

0.79    0.44    0.9    0.74    0.04    0.86    0.48    0.23

Соответственно выбраны хромосомы:

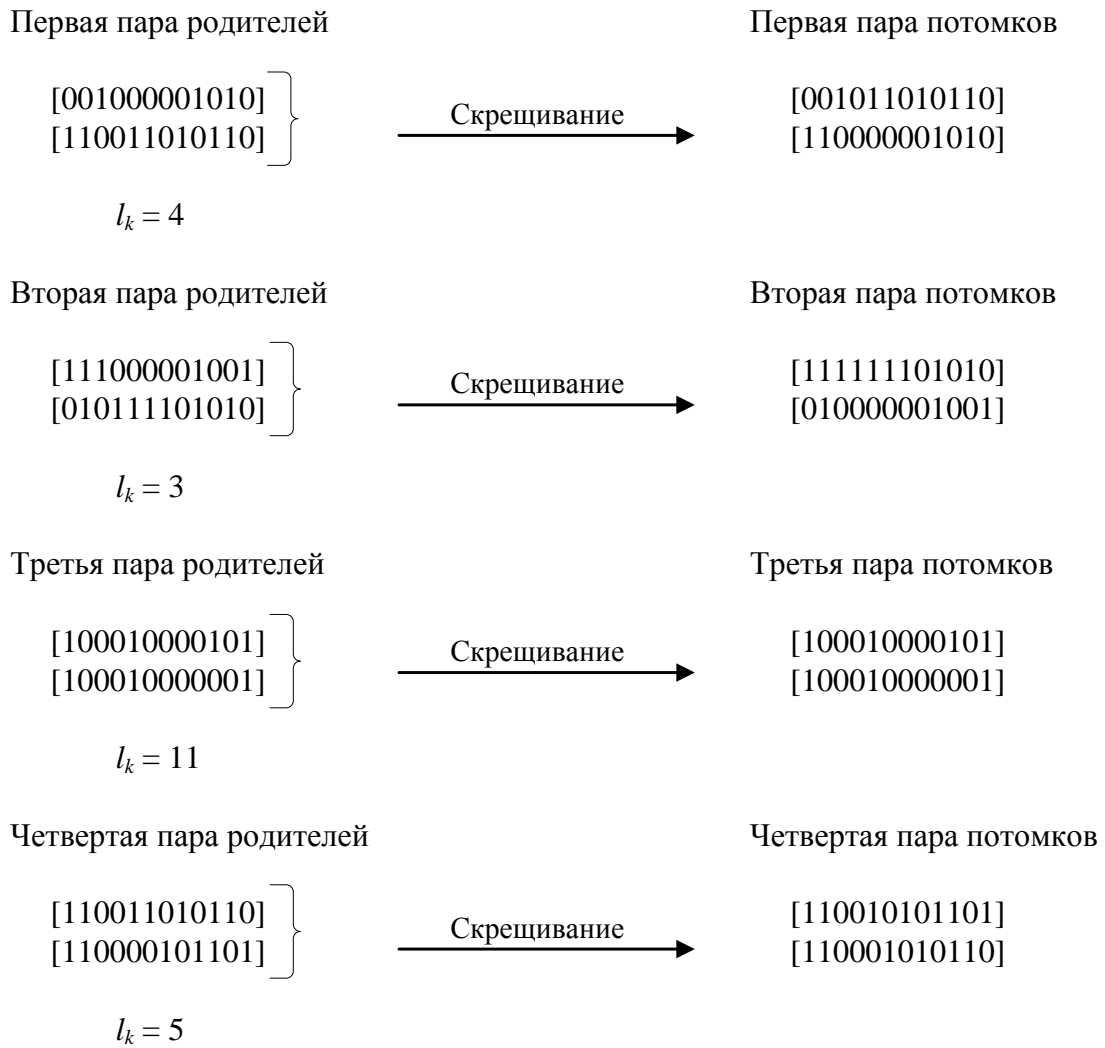
ch<sub>7</sub>    ch<sub>3</sub>    ch<sub>8</sub>    ch<sub>6</sub>    ch<sub>1</sub>    ch<sub>8</sub>    ch<sub>4</sub>    ch<sub>2</sub>

Как видно, хромосома ch<sub>8</sub> была выбрана дважды, а хромосома ch<sub>1</sub> - единожды. Эти хромосомы имеют наибольшее значение функции принадлежности. Однако, также выбраны и хромосомы ch<sub>4</sub> и ch<sub>7</sub>, имеющие наименьшие значения функции принадлежности. На основе всех отобранных хромосом будет создаваться следующее поколение.

Применение операторов скрещивания. В рассматриваемом примере не будем подвергать отобранные для селекции хромосомы мутации, и всех их добавим в популяцию хромосом, для скрещивания. Случайным образом сформируем из этих хромосом четыре пары родителей:

ch<sub>7</sub> и ch<sub>8</sub>            ch<sub>3</sub> и ch<sub>1</sub>            ch<sub>6</sub> и ch<sub>4</sub>            ch<sub>8</sub> и ch<sub>2</sub>

Также случайным образом определим точки скрещивания для каждой из пар. Для первой пары  $l_k = 4$ , для второй  $l_k = 3$ , для третьей  $l_k = 11$ , для четвертой  $l_k = 5$ . Сам процесс скрещивания протекает так, как показано на рисунке 2.3.4.2. По завершению работы оператора скрещивания получаем 4 пары потомков.



**Рисунок 2.3.4.2 Процесс скрещивания хромосом**

Если бы при отборе пар для скрещивания были объединены хромосомы  $ch_8$  с  $ch_8$  и  $ch_7$  с  $ch_2$  вместо  $ch_7$  с  $ch_8$  и  $ch_8$  с  $ch_2$ , а другие остались бы в прежних парах без изменения, то скрещивание  $ch_8$  с  $ch_8$  дало бы пару идентичных родительским хромосом при любой разыгранной точки скрещивания. Отметим, что для хромосом, имеющих наибольшее значение функции принадлежности, такая ситуация более вероятна. Таким образом чем больше у хромосомы значение функции приспособленности, тем больше шансов оказаться в следующей популяции.

После скрещивания отобранных особей мы получим (согласно рисунку 7) новую популяцию потомков:

$$\begin{array}{ll} \text{Ch}_1 = [001011010110]; & \text{Ch}_2 = [110000001010]; \\ \text{Ch}_3 = [111111101010]; & \text{Ch}_4 = [010000001001]; \\ \text{Ch}_5 = [100010000101]; & \text{Ch}_6 = [100010000001]; \\ \text{Ch}_7 = [110010101101]; & \text{Ch}_8 = [110001010110]. \end{array}$$

С целью отличить вновь полученные хромосомы от хромосом предыдущей популяции, будем начинать их обозначение с заглавной буквы С.

Согласно генетическому алгоритму мы вновь возвращаемся ко второму этапу - оценке приспособленности хромосом из только что сформированной популяции. Значения функций приспособленности хромосом новой популяции:

$$\begin{array}{ll} F(\text{Ch}_1) = 6; & F(\text{Ch}_2) = 4; \\ F(\text{Ch}_3) = 9; & F(\text{Ch}_4) = 3; \\ F(\text{Ch}_5) = 4; & F(\text{Ch}_6) = 3; \\ F(\text{Ch}_7) = 7; & F(\text{Ch}_8) = 6. \end{array}$$

Видно, что у популяция потомков среднее значение функции приспособленности выше, чем у родительской популяции. Заметим, что после скрещивания была получена хромосома  $\text{Ch}_3$  значением функции приспособленности которой больше, чем у любой хромосомы из родительской популяции. Но могло произойти и обратное, если бы при случайном выборе родительских хромосом наиболее приспособленная не попала в выборку. Несмотря на это средняя приспособленность новой популяции в редких случаях оказывается ниже предыдущей, благодаря тому, что хромосомы с большими значениями функции приспособленности имеют больше шансов появиться в следующих поколениях.



## 2.4 Реализация прогнозирования временных рядов

В качестве программы, осуществляющей прогнозирование временных рядов была взята программа, разработанная аспирантом МИЭМ НИУ ВШЭ Бухаровым Олегом Евгеньевичем на базе архитектуры Nvidia CUDA (Compute Unified Device Architecture). Выбор был обусловлен новизной, качеством и доступностью программы, а также тем, что на текущий момент её разработчик решал с её помощью задачи схожей тематики для того же заказчика.

### 2.4.1 Функциональная схема системы прогнозирования временных рядов и детали реализации

При объединении генетического алгоритма и интервальной нейронной сети мы получаем универсальную систему прогнозирования, работающую за приемлемое время. Функциональная схема создаваемой системы представлена на рисунке 2.4.1.1.

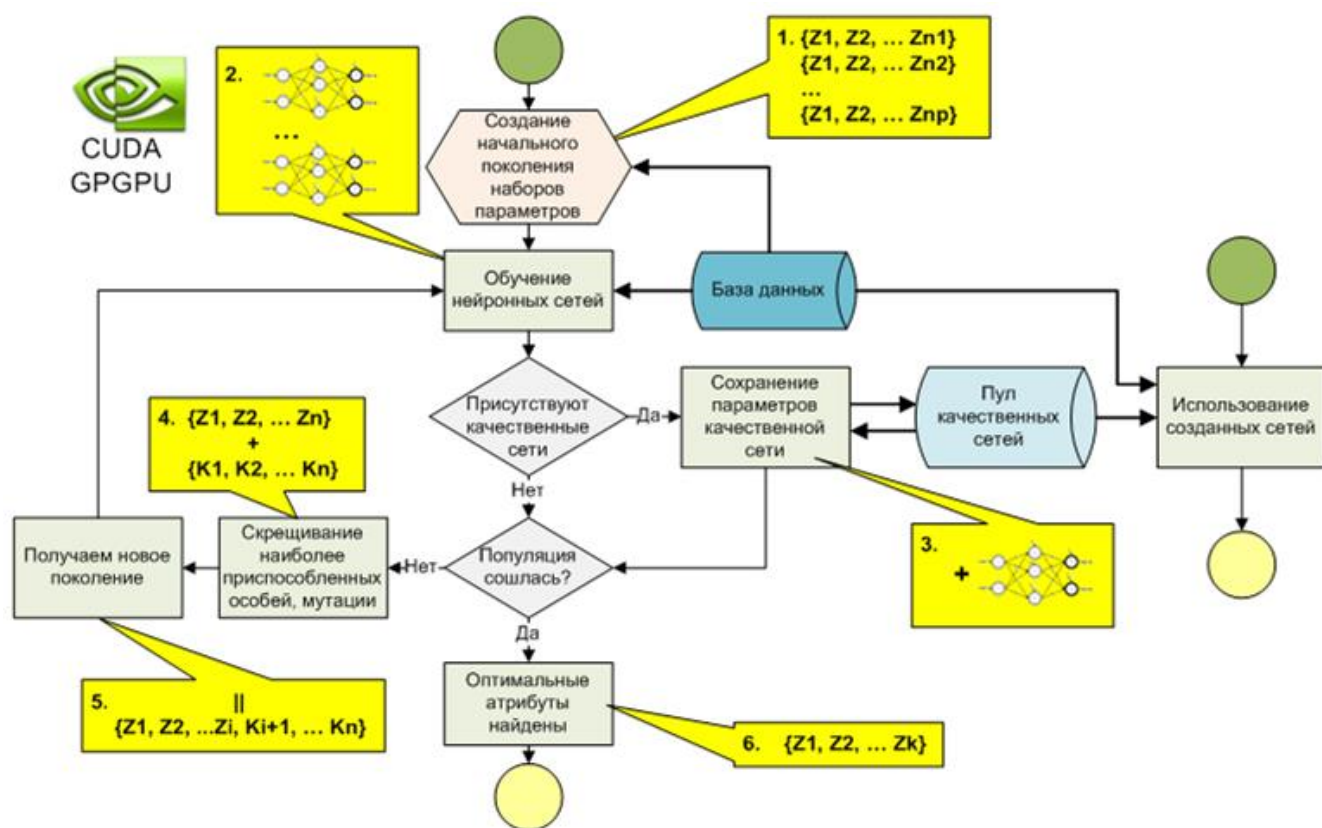


Рисунок 2.4.1.1 Функциональная схема оболочки системы

Перед запуском алгоритмов нахождения оптимальной нейронной сети для прогнозирования интересующего параметра пользователь может задать системе следующие настройки:

- Название и адрес CSV-файла содержащего временные ряды для обучения сетей
- Название и адрес CSV-файла содержащего временные ряды на основе которых предполагается осуществлять прогнозирование.
- Путь к директории для сохранения качественных сетей.
- Номера столбцов атрибутов и прогнозируемой величины
- Диапазон количества атрибутов используемых для прогнозирования
- Диапазон размеров временных окон для атрибутов, по которым будет осуществляться прогнозирование (на основании какого количества последовательных значений временного ряда строить прогноз)
- Размер временного окна прогнозируемой величины (сколько последовательных значений необходимо спрогнозировать)
- Размер временного лага между окном атрибутов и окном прогнозируемой величины (может быть отрицательным, например, для оценивания показателей одной величины по показателям других, находящихся в одном временном интервале с оцениваемой величиной)
- Максимальные и минимальные значения исследуемых величин (для нормализации)
- Размер популяций для генетического алгоритма
- Количество внутренних нейронов для нейронной сети
- Количество циклов обучения для нейронных сетей
- Количество обучаемых нейронных сетей на одном наборе параметров
- Максимальное число итераций

## 2.4.2 Алгоритм работы оболочки системы:

1. Первым шагом алгоритма поиска оптимальной для прогнозирования нейронной сети является создание начальных поколений. Для каждого значения диапазона используемых атрибутов и диапазона временных окон формируется заданное пользователем количество (размер популяции) случайных наборов параметров, на основании которых будет строиться прогноз.

2. На следующем шаге для каждого набора параметров система обучает множество нейронных сетей, прогнозирующих искомую величину. Обучение происходит параллельно с использованием архитектуры CUDA для распараллеливания вычислений на графических процессорах, что существенно сокращает время работы алгоритма.

3. Если среди обученных сетей существует сеть, работающая с меньшей погрешностью, чем остальные сети, параметры такой сети необходимо сохранить в пул качественных сетей. Как только в пуле появится первая сеть, параллельно с алгоритмом нахождения оптимальных параметров нейронной сети может работать алгоритм прогнозирования, выбирая из пула сеть с наименьшей погрешностью.

4. Если популяция не сошлась или максимальное число итераций не превышено, в каждом поколении определяется набор лучших особей, имеющих наименьшую ошибку прогнозирования, для скрещивания.

5. На основе отобранных лучших особей генерируются новые поколения путем скрещивания и мутаций. Выполнение алгоритма продолжается с пункта 2.

6. Программа выходит из цикла либо если превышено максимальное количество итераций, либо в случае сходимости популяции. Если популяция сошлась, то на выходе мы получаем оптимальный набор атрибутов, а основанная на нем нейронная сеть уже находится в пуле качественных сетей.

Как видно из схемы, одним из ключевых преимуществ системы является возможность осуществлять прогнозирование уже на ранних этапах работы генетического алгоритма (параллельно с его работой). Данная возможность обусловлена добавлением в пул сетей, имеющих минимальную ошибку прогнозирования, на каждом шаге генетического алгоритма. Таким образом, по мере работы генетического алгоритма пул будет обновляться сетями, способными более точно предсказывать прогнозируемые величины.

Вторым и основным преимуществом оболочки является распараллеливание самой трудоемкой процедуры данного решения - обучения множества нейронных сетей. Благодаря использованию архитектуры CUDA скорость работы данного алгоритма значительно выше, чем у аналогичной реализации только с участием CPU.

Эффективное использование связки CPU и GPU достигается за счет распараллеливания алгоритма на GPU с использованием нескольких тысяч отдельных нитей, за счет использования преимуществ иерархии нитей CUDA, подготовки данных на CPU параллельно с обработкой на GPU.

Учитывая, что каждый нейрон сети может рассчитываться независимо от остальных нейронов его слоя, а также тот факт, что каждая отдельная нейронная сеть активно взаимодействует со своими параметрами (синаптическими весами), для создаваемой оболочки был разработан и реализован следующий метод распараллеливания: помимо одновременного обучения нейронных сетей одного поколения каждая сеть распараллеливается по нейронам внутри отдельного блока. Этот метод позволяет эффективно нагрузить GPU множеством нитей, а также использовать быструю распределенную память при проведении расчетов с сетями.

### **3 Компьютерный эксперимент.**

#### **3.1 Исходные данные по исследуемым объектам.**

Анализ месячных карт аномалий приземного давления и температуры показывает, что отрицательные аномалии температуры в умеренных широтах зимой связаны с положительной аномалией давления в этих же широтах и наоборот. Кроме того, обнаруживается, что при положительной аномалии давления в высоких широтах в умеренных и низких широтах имеет место отрицательная аномалия. Наоборот, при отрицательной аномалии давления в высоких широтах в умеренных и низких широтах имеет место положительная аномалия давления.

На Атлантике это приводит к известной связи между колебаниями интенсивности азорского антициклона и исландского циклона. На основе этих колебаний определялась интенсивность циркуляции над Атлантикой. Вполне понятно, что если на район исландского циклона накладывается положительная аномалия давления, а на район азорского антициклона – отрицательная, то градиент приземного давления между этими центрами будет ослаблен, а при противоположном расположении аномалии он будет усилен.

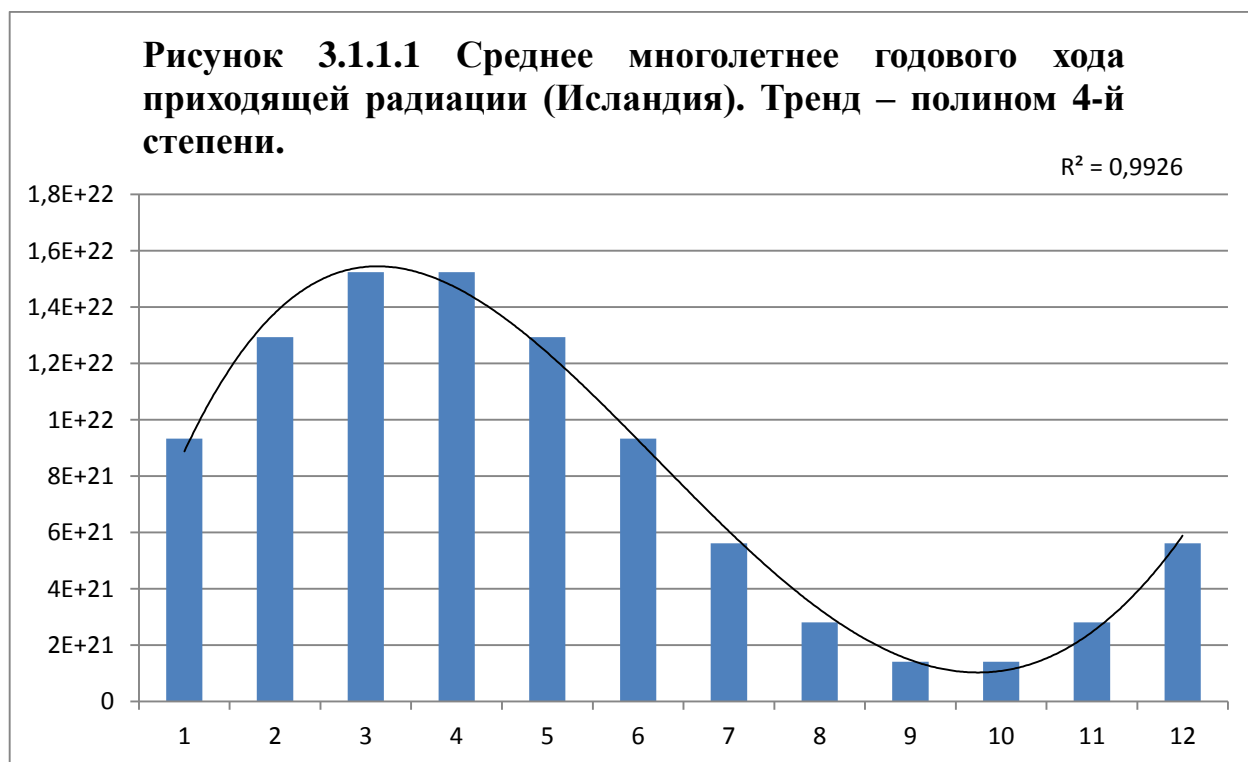
### 3.1.1 Исландия.

Годовой ход атмосферного давления в многолетнем среднем определяется годовым ходом проходящей на верхнюю границу атмосферы (или, точнее, на поверхность эллипсоида в отсутствии атмосферы соответствующей широтной зоны). Коэффициент корреляции 0,93.

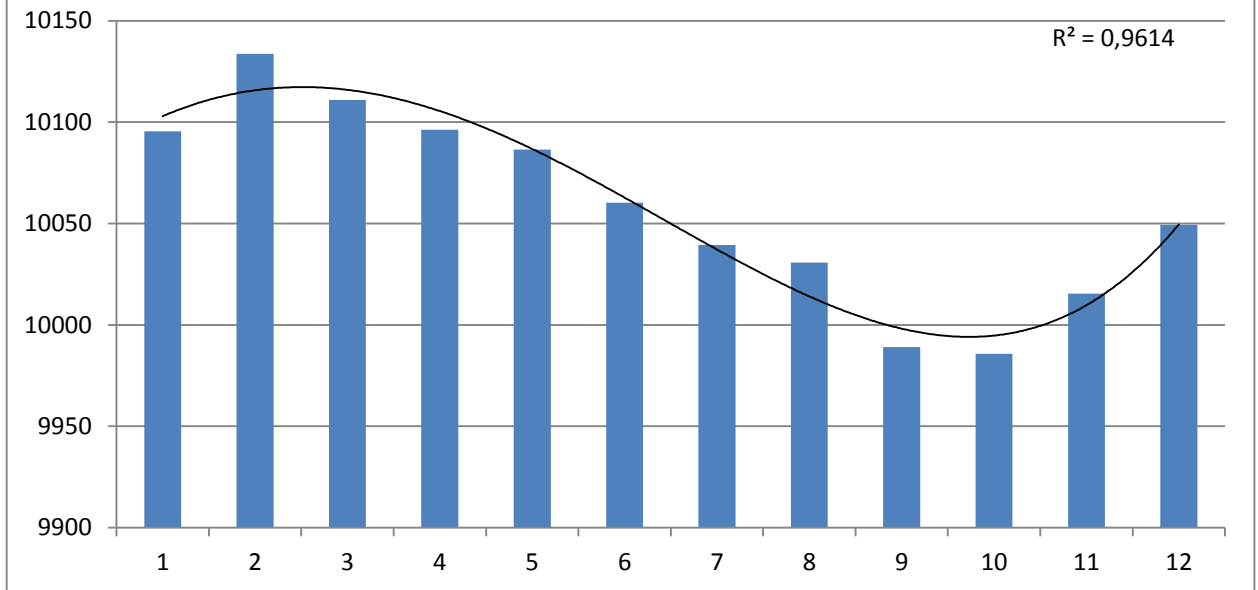
В отдельные годы годовой ход атмосферного давления связан с соответствующим годовым ходом проходящей солнечной радиации с коэффициентом корреляции 0,55.

Амплитуда межгодовой изменчивости атмосферного давления в годовом ходе имеет отрицательную корреляционную связь с многолетним средним годовым ходом амплитуды межгодовой изменчивости проходящей радиации.  $K = -0,93$ .

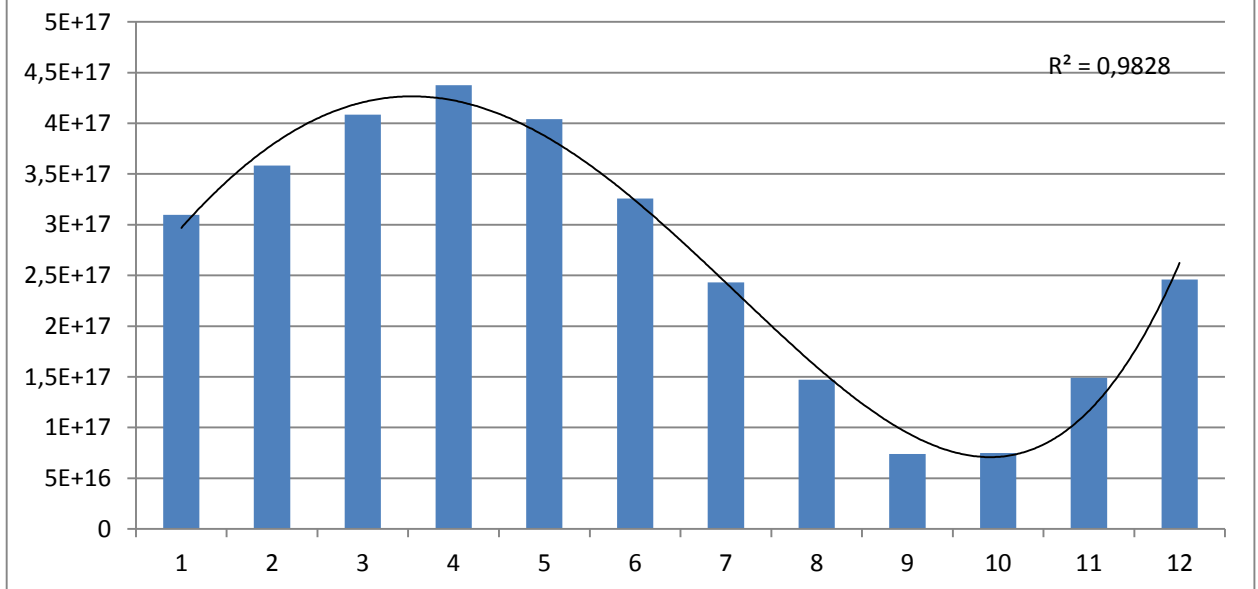
Амплитуда межгодовой изменчивости проходящей радиации меньше при положении Земли вблизи перигелия орбиты т.е. в зимнее полугодие в северном полушарии. В это время отмечаются максимальные амплитуды межгодовой изменчивости атмосферного давления в годовом ходе.



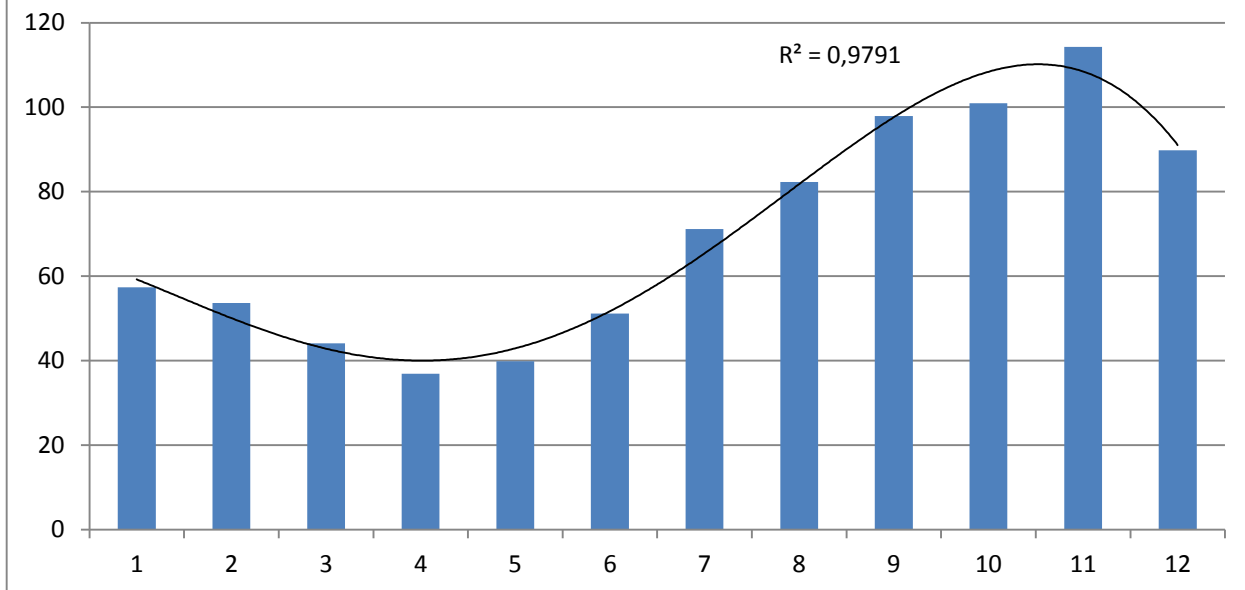
**Рисунок 3.1.1.2 Среднее многолетнее годового хода атмосферного давления (Исландия). Тренд - полином 4-й степени.**



**Рисунок 3.1.1.3 Амплитуда межгодовой изменчивости в годовом ходе приходящей радиации (Исландия). Тренд полином – 4-й степени**



**Рисунок 3.1.1.4 Амплитуда межгодовой изменчивости в годовом ходе атмосферного давления (Исландия). Тренд – полином 4-й степени.**





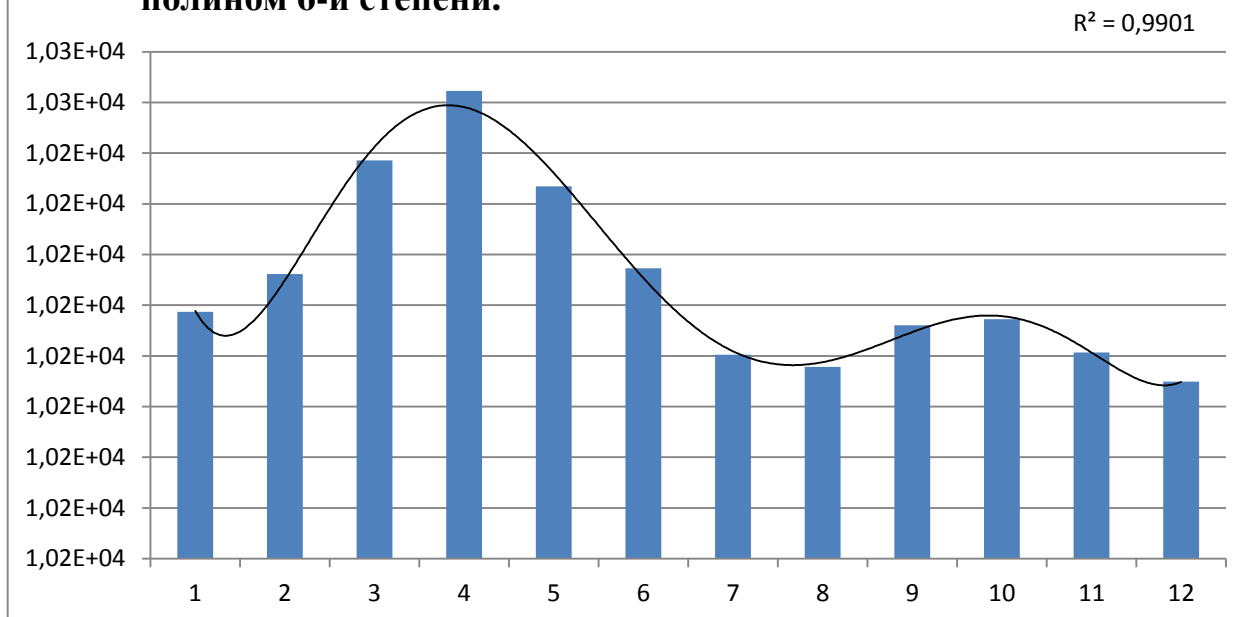
### 3.1.2 Азорские острова.

Годовой ход атмосферного давления в многолетнем среднем определяется годовым ходом проходящей на верхнюю границу атмосферы (или, точнее, на поверхность эллипсоида в отсутствии атмосферы соответствующей широтной зоны). Коэффициент корреляции 0,80.

В отдельные годы годовой ход атмосферного давления связан с соответствующим годовым ходом проходящей солнечной радиации с коэффициентом корреляции 0,33.

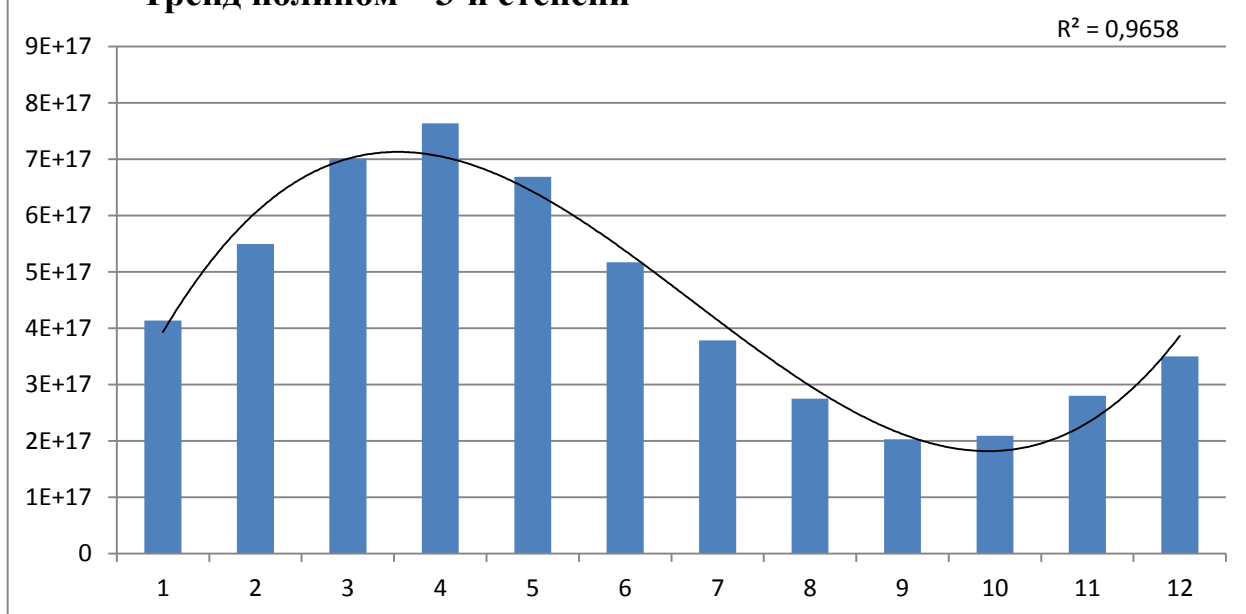


**Рисунок 3.1.2.2 Среднее многолетнее годового хода атмосферного давления (Азорские острова). Тренд - полином 6-й степени.**

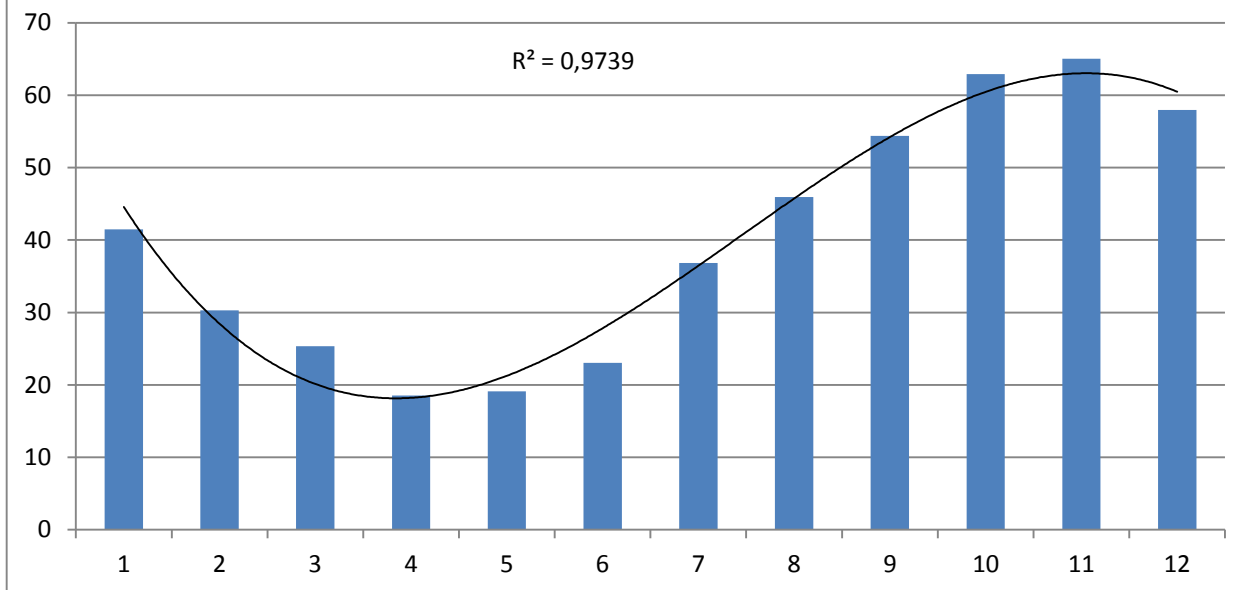


Амплитуда межгодовой изменчивости атмосферного давления в годовой ходе имеет отрицательную корреляционную связь с многолетним средним годовым ходом амплитуды межгодовой изменчивости приходящей радиации.  $K = -0,90$ .

**Рисунок 3.1.2.3 Амплитуда межгодовой изменчивости в годовом ходе приходящей радиации (Азорские острова). Тренд полином – 3-й степени**



**Рисунок 3.1.2.4 Амплитуда межгодовой изменчивости в годовом ходе атмосферного давления (Азорские острова). Тренд – полином 3-й степени.**

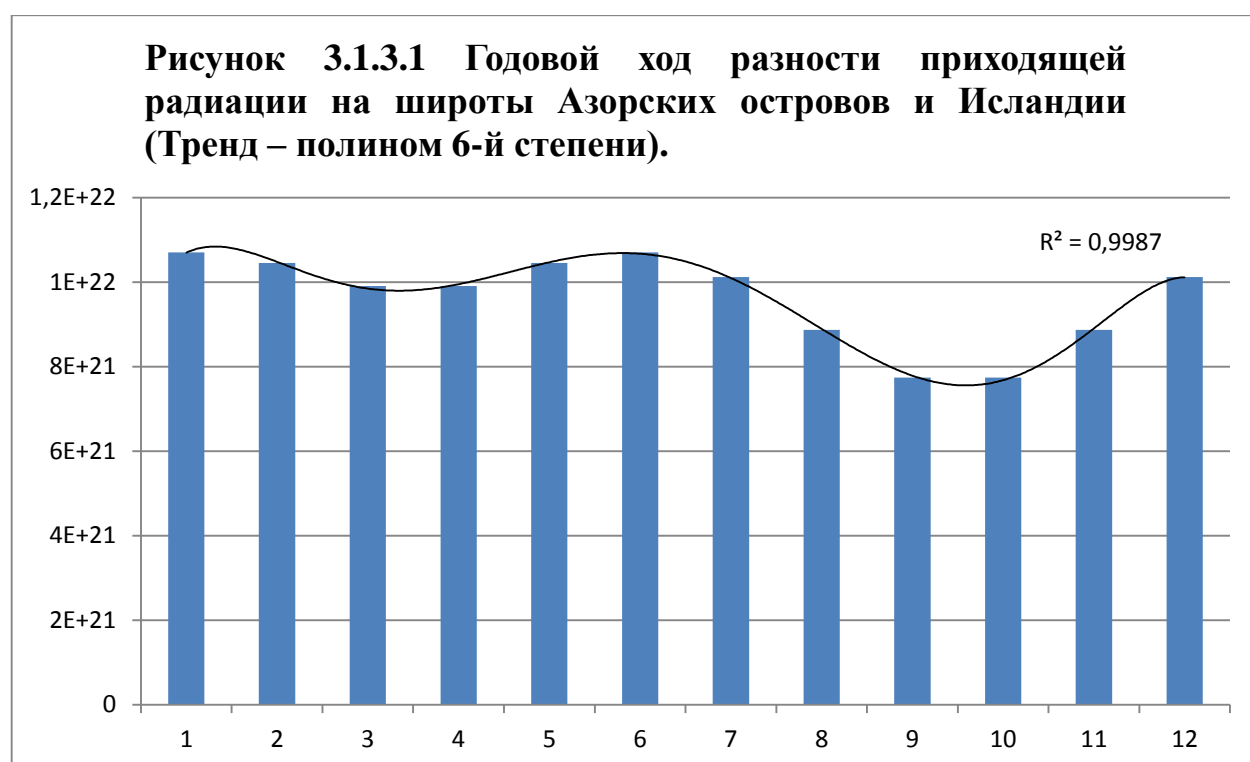


### 3.1.3 Разность – Азорские минус Исландия.

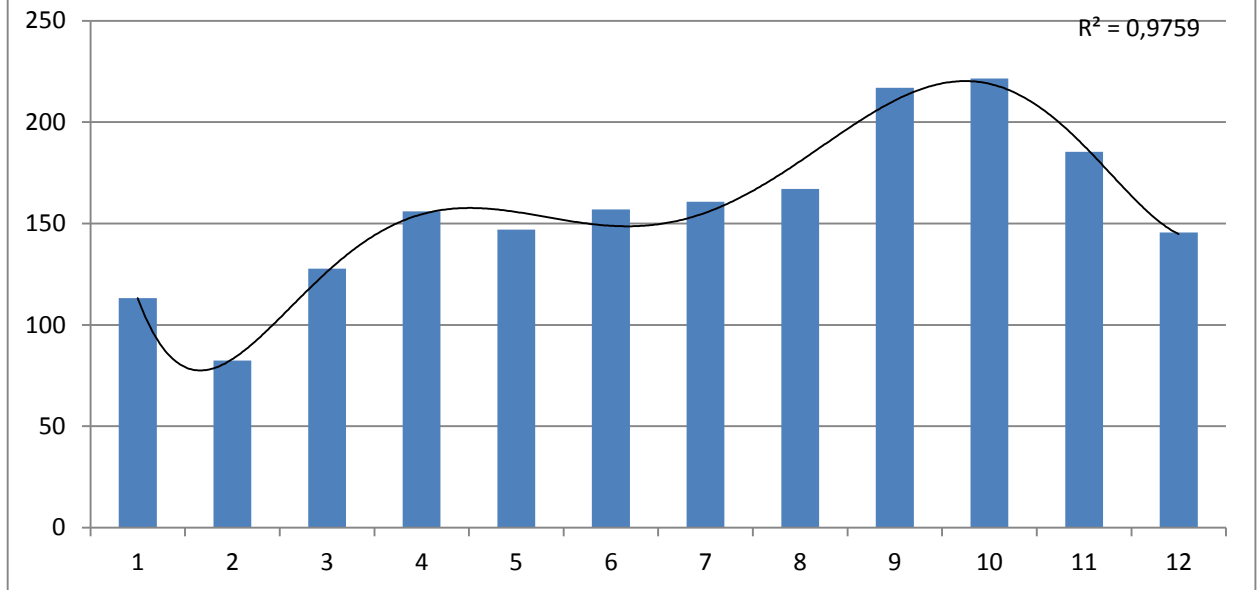
Годовой ход атмосферного давления в многолетнем среднем определяется годовым ходом проходящей на верхнюю границу атмосферы (или, точнее, на поверхность эллипсоида в отсутствии атмосферы соответствующей широтной зоны). Коэффициент корреляции 0,84.

В отдельные годы годовой ход атмосферного давления связан с соответствующим годовым ходом проходящей солнечной радиации с коэффициентом корреляции 0,33.

Амплитуда межгодовой изменчивости атмосферного давления в годовой ходе имеет отрицательную корреляционную связь с многолетним средним годовым ходом амплитуды межгодовой изменчивости проходящей радиации.  $K = -0,89$ .

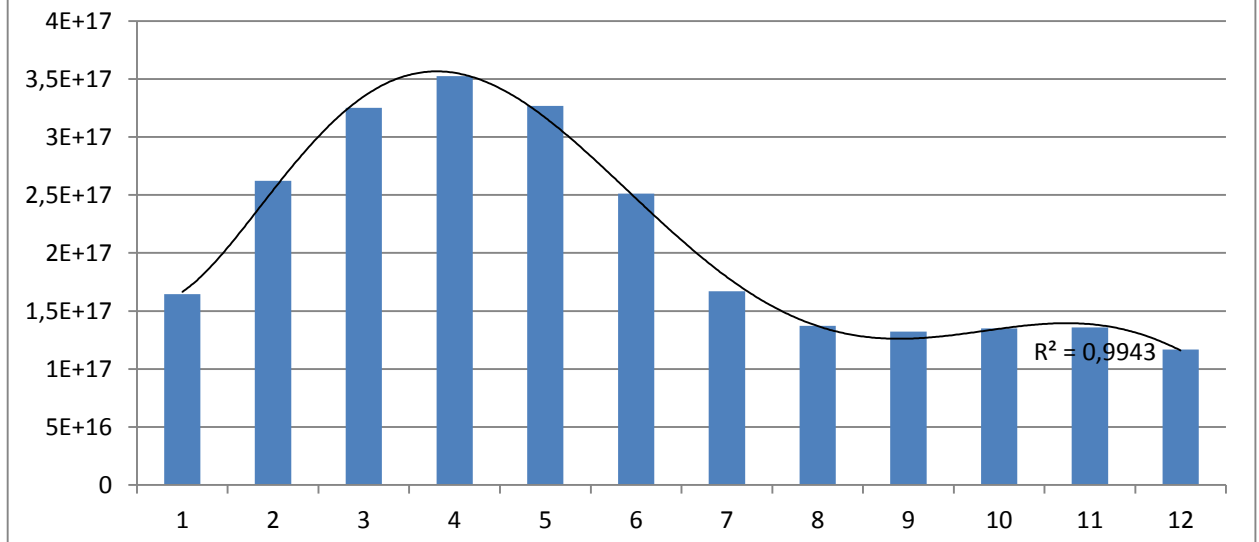


**Рисунок 3.1.3.2** Годовой ход разности атмосферного давления на Азорских островах и в Исландии (Тренд – полином 6-ц степени).

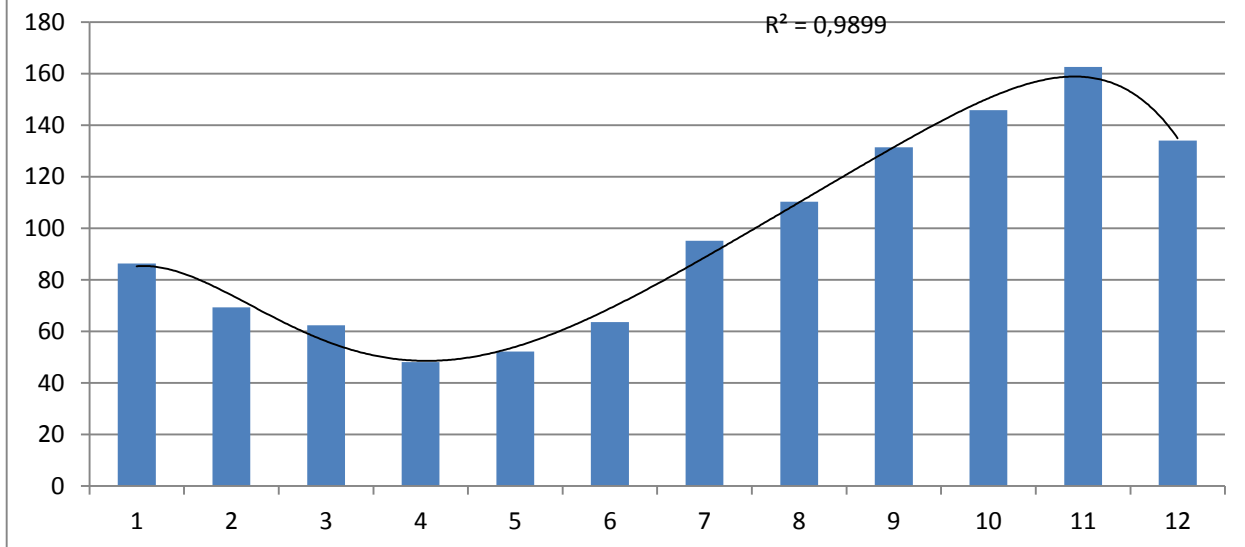


Амплитуда межгодовой изменчивости приходящей радиации меньше при положении Земли вблизи перигелия орбиты т.е. в зимнее полугодие в северном полушарии. В это время отмечаются максимальные амплитуды межгодовой изменчивости атмосферного давления в годовом ходе.

**Рисунок 3.1.3.3** Годовой ход амплитуды межгодовой изменчивости разности приходящей радиации на широту Азорских островов и Исландии (Тренд – полином 6-й степени)



**Рисунок 3.1.3.4** Годовой ход амплитуды межгодовой изменчивости разности атмосферного давления на Азорских островах и в Исландии (Тренд – полином 6-й степени)



### 3.2 Результаты первичной обработки данных

Начальные попытки прогнозирования с помощью системы поддержки принятия решений привели к весьма высоким коэффициентам ошибок, как можно увидеть из таблицы 3.2.1. Данная аномалия явно свидетельствует о необходимости предварительного аналитического разбора данных для поиска закономерностей и построения более точной модели.

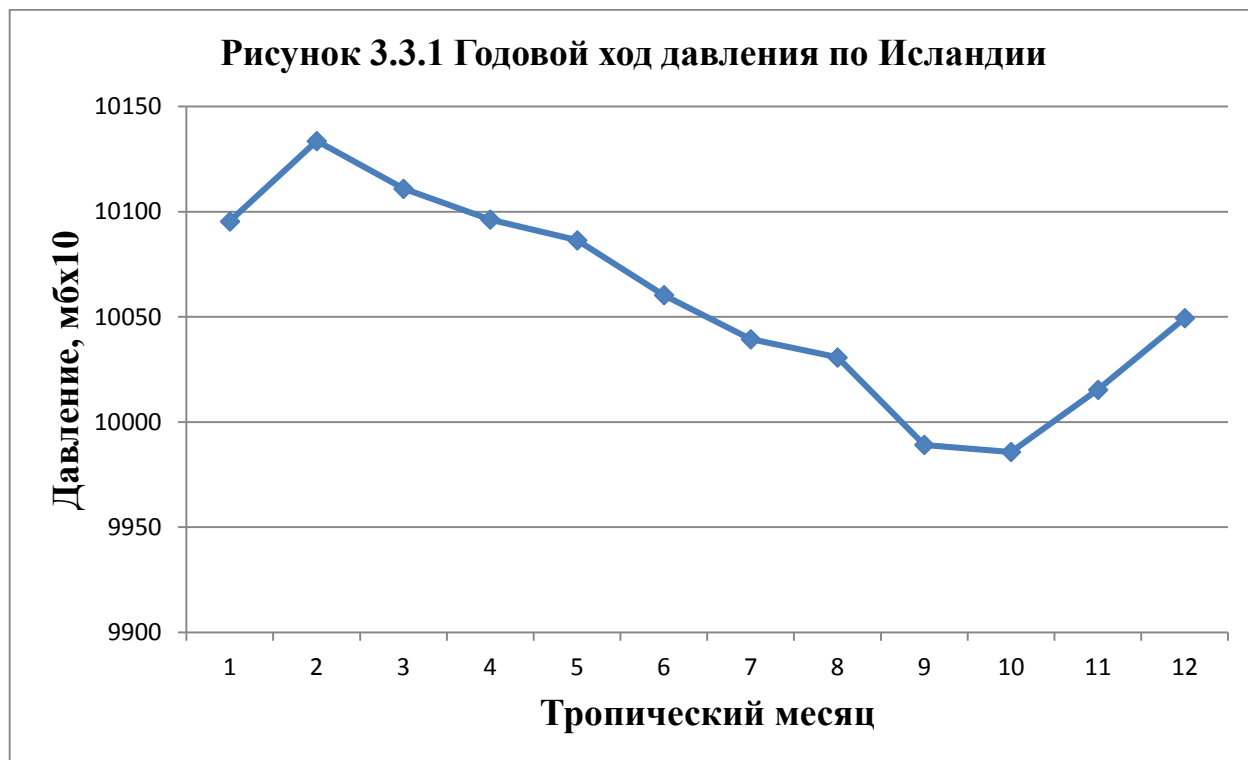
**Таблица 3.2.1**

#### **Коэффициенты ошибок при прогнозировании.**

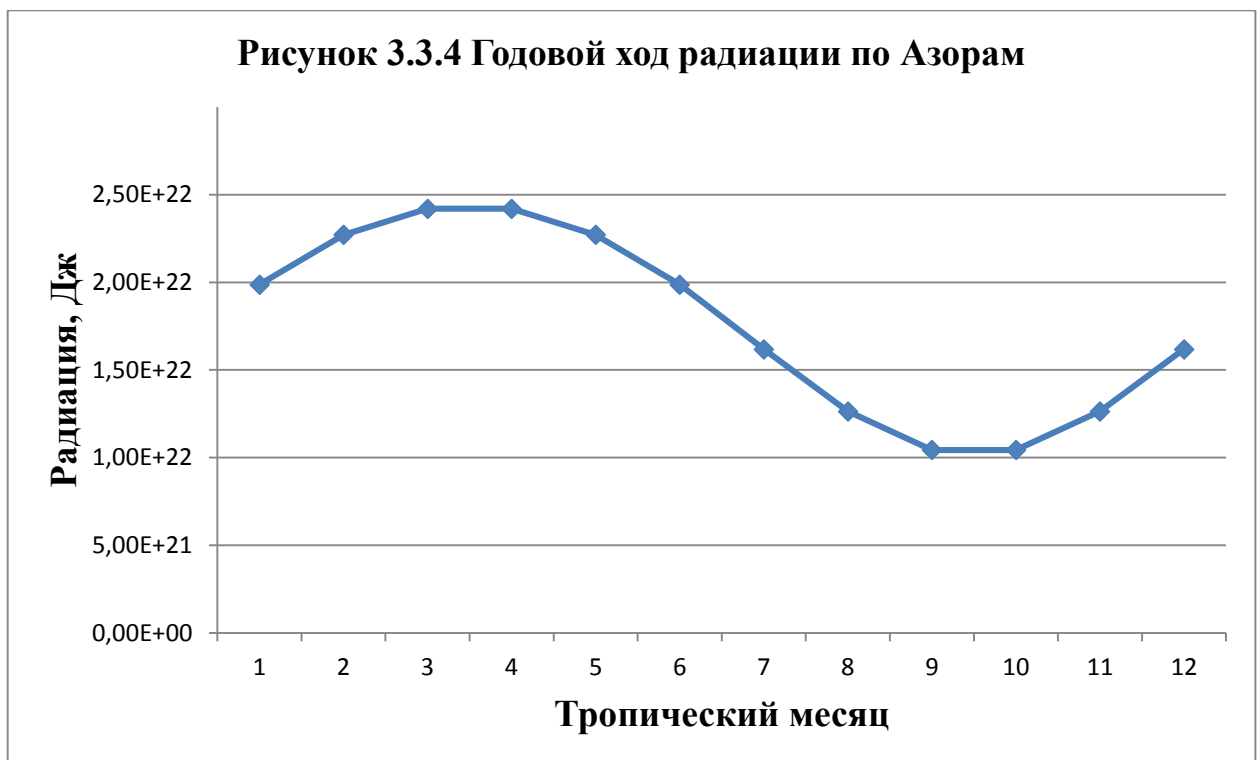
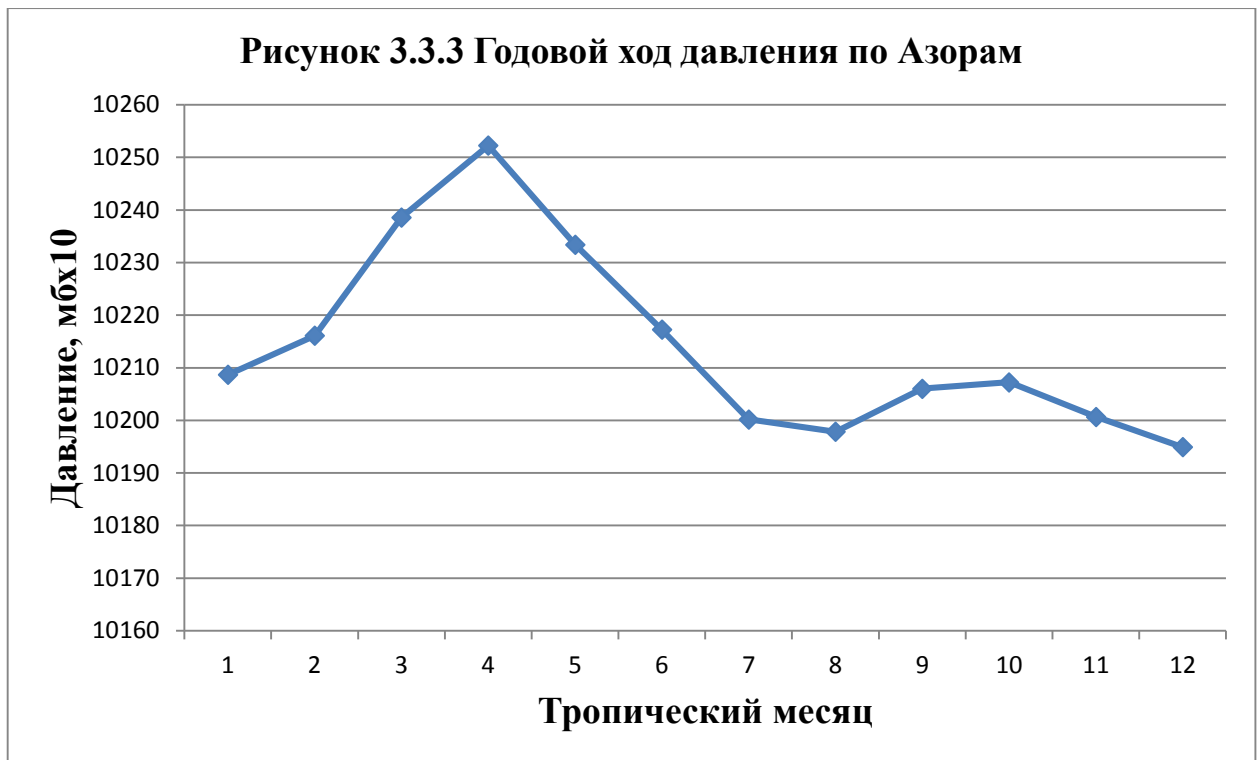
Объект прогноза	Средняя ошибка	Максимальная ошибка
Азоры, общий	0,1184	0,4056
Исландия, общий	0,1663	0,4443
Исландия, месяц 1	0,2149	0,3889
Азоры, месяц 1	0,1813	0,3587
Исландия, месяц 2	0,2058	0,3997
Азоры, месяц 2	0,1762	0,4173
Исландия, месяц 3	0,1816	0,3808
Азоры, месяц 3	0,1700	0,3482
Исландия, месяц 4	0,1099	0,2329
Азоры, месяц 4	0,1817	0,3510
Исландия, месяц 5	0,2233	0,4445
Азоры, месяц 5	0,1768	0,4289
Исландия, месяц 6	0,1891	0,4356
Азоры, месяц 6	0,1732	0,3778

### 3.3 Поиск связи годового хода давления и радиации.

В ходе обработки имеющихся данных были получены следующие результаты по годовому ходу:

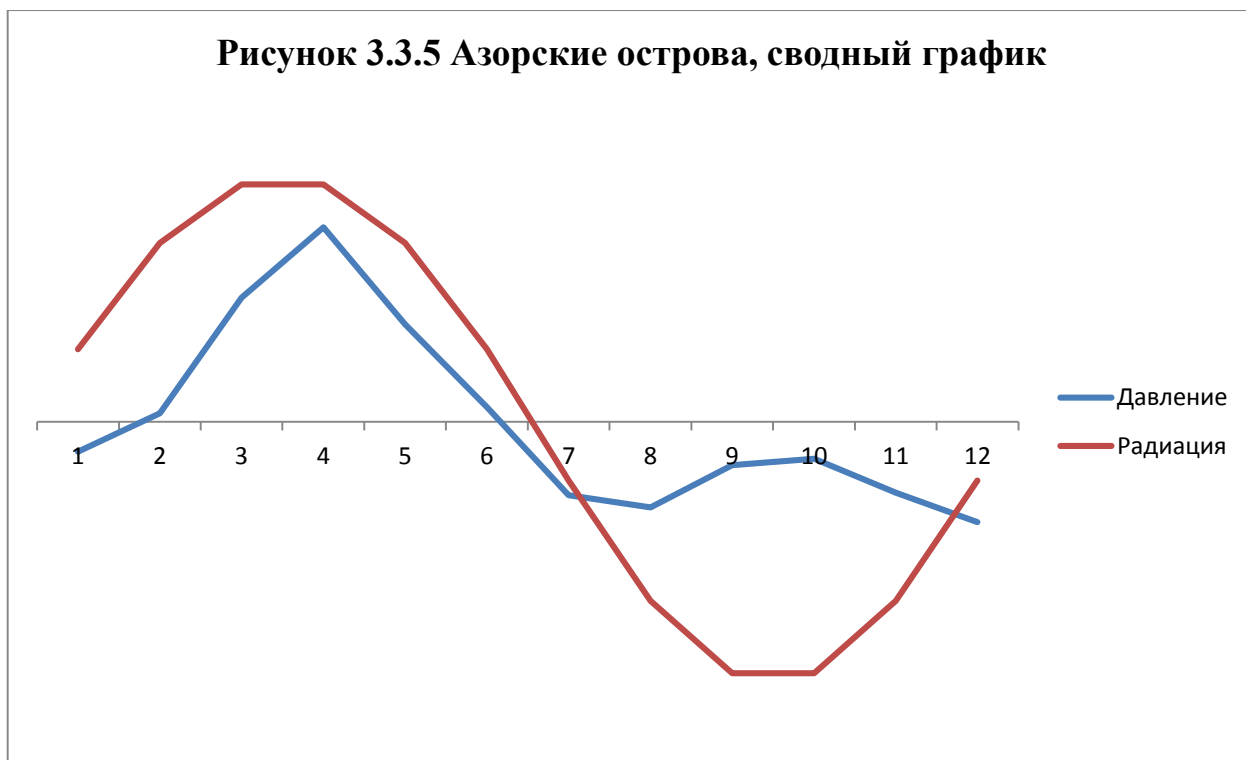




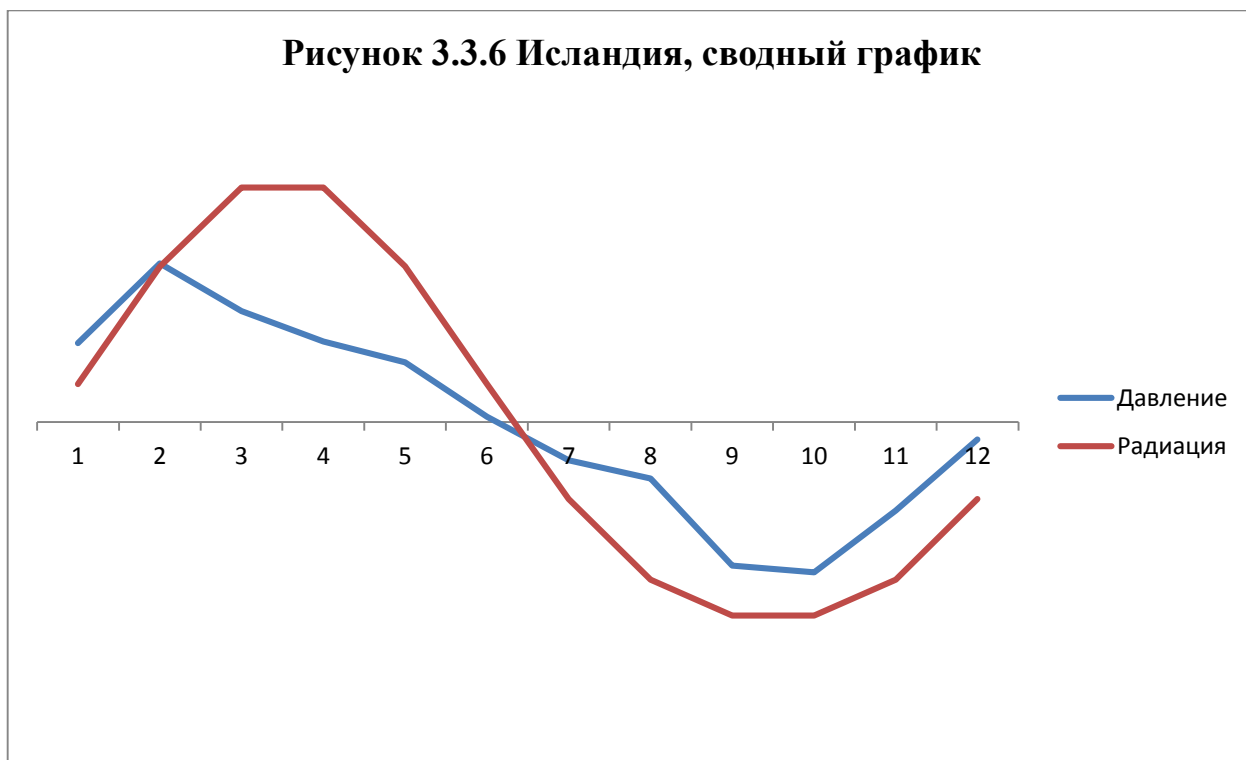


Таким образом, при совмещении данных по давлению и радиации на одном графике, получим:

**Рисунок 3.3.5 Азорские острова, сводный график**



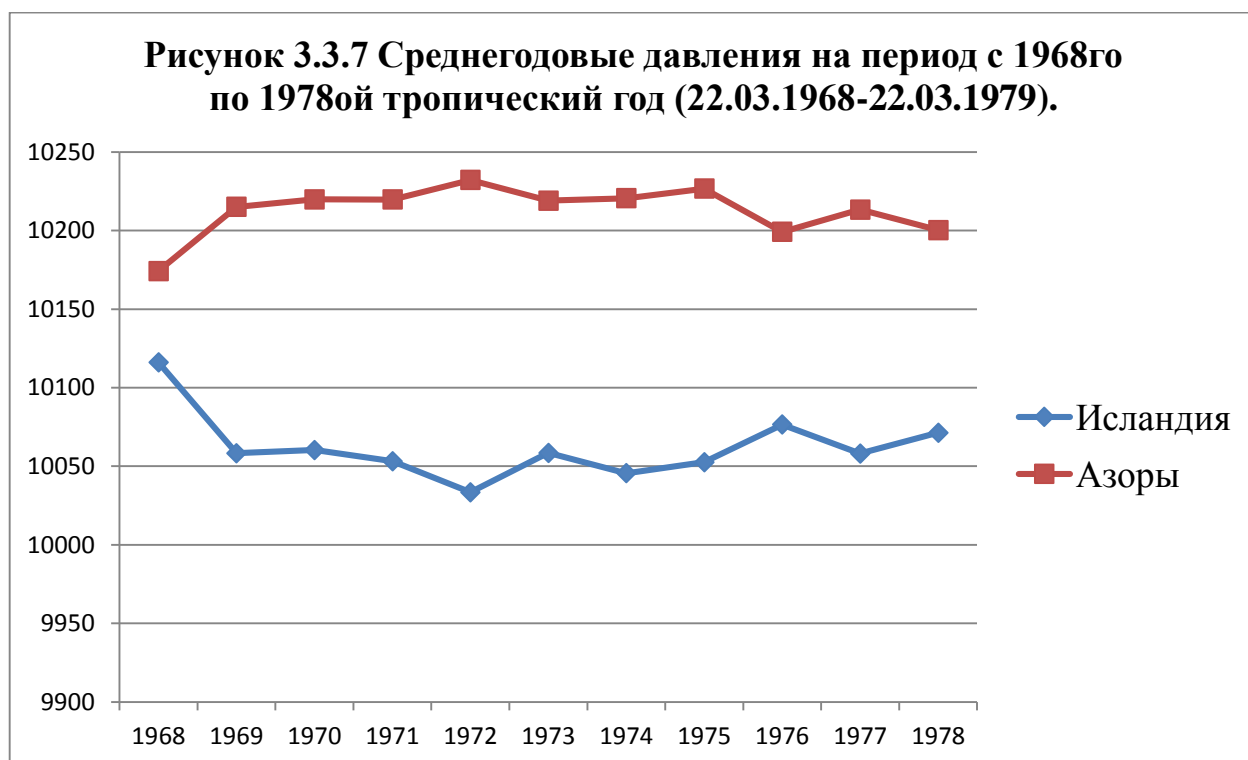
**Рисунок 3.3.6 Исландия, сводный график**

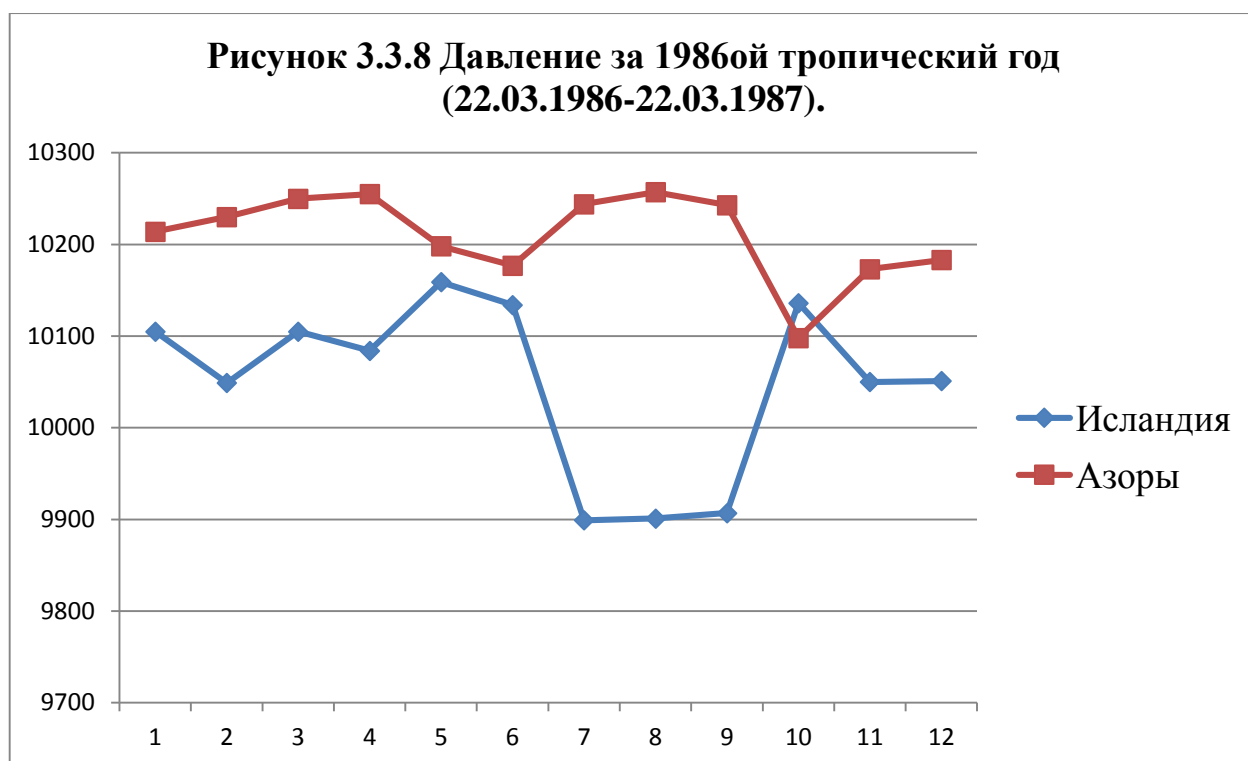


По данным графикам наглядно видно, что давление Азорских островов прямо зависит от радиации на промежутке с 1го по 7ой тропический месяц (примерно соответствует апрелю-октябрю календарного месяца), а давление Исландии, наоборот, прямо зависит от радиации в «зимнее» полугодие. При этом в

противоположных полугодиях («летнего» для Исландии и «зимнего» для Азорских островов) наблюдается тенденция скорее к обратной пропорциональности давления и приходящей радиации.

Вместе с тем, тщательное исследование данных по давлению показало, что как среднегодовые давления в рассматриваемых пунктах, так и отклонения их от средних значений, имеют достаточно высокую отрицательную корреляцию (-0.6133 для среднегодовых давлений и -0.53408 для отклонений от средних значений). Более того, во многих случаях вышеуказанные параметры имеют ярко выраженную обратную зависимость, как, например, на следующих двух графиках.





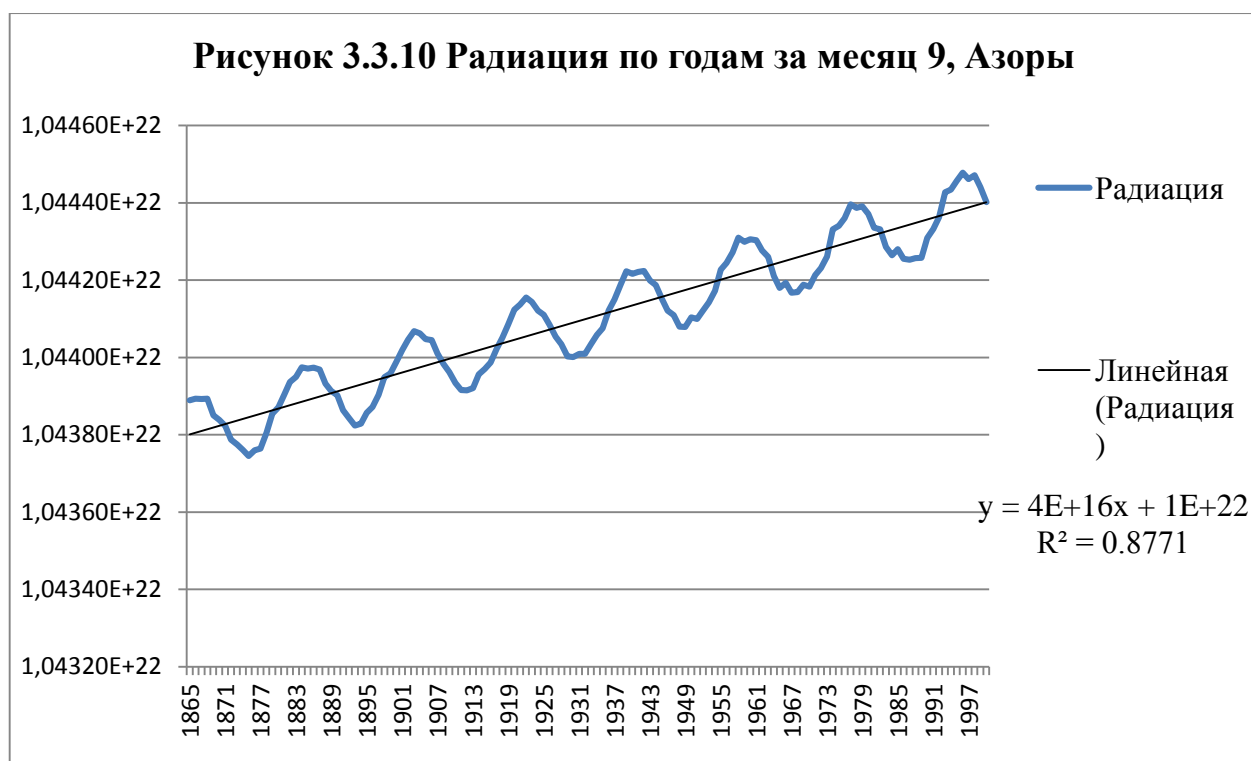
Данное явление объясняется переносом воздушных масс между областями с разными температурами. Таким образом, повышение температуры (и, соответственно, давления) на Азорских островах в летний период приводит к оттоку воздуха в направлении экватора и понижению давления в Исландии, что мы и видим на рисунках 3.3.5 и 3.3.6 на промежутке с 1го по 5ый месяцы. Аналогично, понижение давления в Исландии в 7ом-10ом месяцах приводит к повышению давления на Азорских островах.

В качестве заключения можно сделать следующий вывод: давление в двух исследуемых точках Северо-Атлантического течения прямо зависят от приходящей радиации в рассматриваемую точку радиации и обратно – от давления во второй из рассматриваемых точек, которое, в свою очередь, прямо зависит от приходящей радиации. Таким образом, мы доказали зависимость климатических параметров от орбитальных.

### 3.4 Многолетний ход по среднемесячным значениям.

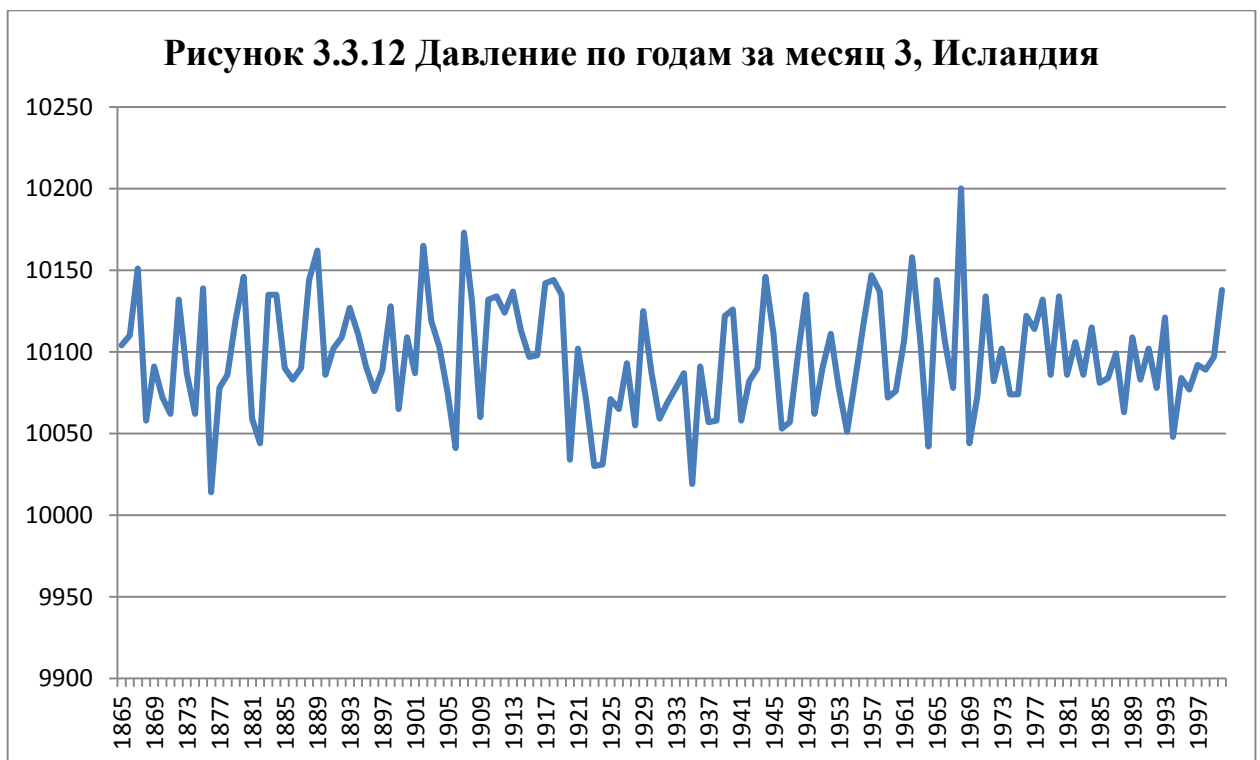
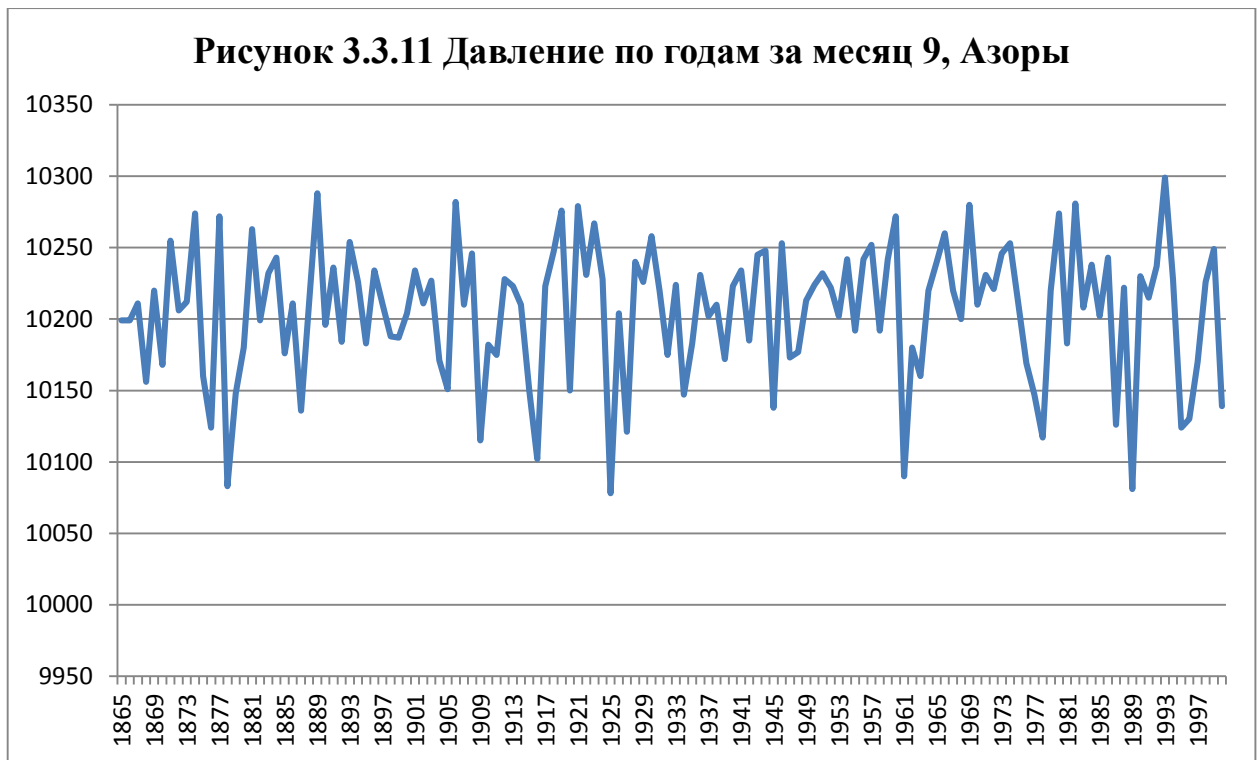
В рассматриваемый период времени (1865-2000гг) приходящая на Исландию и Азорские острова радиация имеет две ярко выраженные тенденции: снижение количества приходящей радиации в каждом месяце «летнего» полугодия и повышение этого количества в «зимнем». Пример показан на рисунках 3.3.9 и 3.3.10.





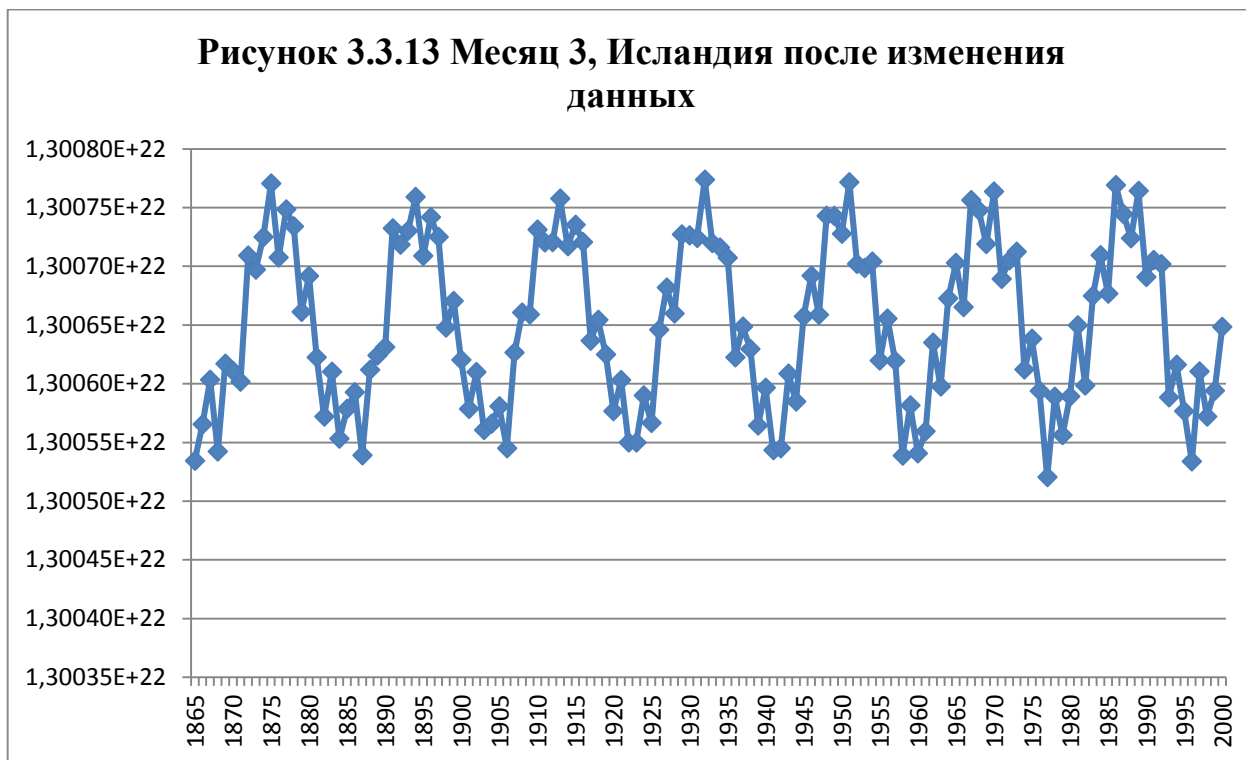
Как видно, приходящая в месяц радиация в то же время подвержена незатухающим колебаниям с периодом в 19 лет, в точности так, как и было сказано в пункте 1.3.2.

Вместе с этим, давление за соответствующие месяцы не подвержено ни одной из данных тенденций, как можно увидеть на рисунках 3.3.11 и 3.3.12.



Наблюдения подтверждаются вычислениями: максимальный модуль корреляции между давлением и радиацией за месяц составляет 0.11699 для Азорских островов и 0.16124 для Исландии, что недостаточно для существования зависимости.

Была проведена попытка изменения данных радиации так, чтобы линией тренда являлась прямая, параллельная оси абсцисс. Проверка проводилась на Зем месяце Исландии (коэффициент корреляции  $-0.026047$ ). В результате исчезла тенденция убывания, но сохранилась тенденция периодичности.



Коэффициент корреляции, однако же, не претерпел значительных изменений, что означает либо отсутствие чёткой зависимости между климатическими и орбитальными параметрами за конкретный месяц на протяжении многих лет, либо сильную зашумлённость данной зависимости и необходимость иных методов подхода к задаче.



## **Заключение.**

Основной задачей работы являлось изучение зависимости климатических параметров Исландии и Азорских островов, представленных давлением, от орбитальных параметров, совокупной характеристикой которых является приходящая на Землю радиация. Задача была преимущественно выполнена применительно к годовому ходу, подтвердив имеющуюся теорию, и выявила недочёты в имеющемся представлении данных, значительно усложняющие поиск возможных закономерностей в прогнозировании месячных климатических параметров на основании месячной же радиации.

Таким образом, несмотря на получение значительных результатов, в данных задачах остаётся немалый потенциал для их дальнейшей разработки и углубления, особенно в части, касающейся прогнозирования.

## Список литературы.

1. Воейков А.И. Климаты земного шара, в особенности России / Собр. соч. – М. – Л.: АН СССР, 1948. – т. 1.– с. 163 – 671.
2. Шулейкин В.В. Физика моря. – М.: АН СССР, 1953. – 990 с.
3. Погосян Х.П. Общая циркуляции атмосферы. – Л.: Гидрометеиздат, 1959. – 260 с.
4. Погосян Х.П. Циклоны. – Л.: Гидрометеиздат, 1976. – 148 с.
5. Кац А.Л. Сезонные изменения общей циркуляции атмосферы и долгосрочные прогнозы – Л.: Гидрометеиздат, 1960. – 270 с.
6. Хромов С.П. Метеорология и климатология. – Л.: Гидрометеиздат, 1968. – 492 с.
7. Лоренц Э.Н. Природа и теория общей циркуляции атмосферы. – Л.: Гидрометеиздат, 1970. – 260 с.
8. Монин А.С. Введение в теорию климата. – Л.: Гидрометеиздат, 1982. – 246 с.
9. Хргиан А.Х. Физика атмосферы. – М.: МГУ, 1986. – 328 с.
10. Монин А.С., Шишков Ю.А. Климат как проблема физики // Успехи физ. Наук, 2000. – т. 170. – № 4. – с. 419 – 445.
11. Хромов С.П., Петросянц М.А. Метеорология и климатология. – М.: МГУ, 2006. – 582 с.
12. Витвицкий Г.Н. Зональность климата Земли. – М.: Мысль, 1979. – 253 с.
13. Кочин Н.Е. Построение модели зональной циркуляции атмосферы // Труды ГГО, 1936. – вып. 10 (2). – с. 3 - 27.
14. Филлипс Н.А. Общая циркуляция атмосферы: численный эксперимент. / Теория климата. Ред. Л.С. Гандин. А.С.Дубов, М.Е.Швец. – Л.: Гидрометеиздат, 1967. – с. 6 – 60.
15. Смагоринский Дж., Манабе С., Холлоуэй Дж. Л. Результаты численных опытов с девятиуровневой моделью общей циркуляции атмосферы / Теория климата. Ред. Л.С. Гандин. А.С.Дубов, М.Е.Швец. – Л.: Гидрометеиздат, 1967. – с. 117 – 184.

16. Манабе С., Смагоринский Дж., Стриклер Р.Ф. Численное моделирование средней картины общей циркуляции атмосферы с учетом процессов влагообмена / Теория климата. Ред. Л.С. Гандин. А.С.Дубов, М.Е.Швец. – Л.: Гидрометеиздат, 1967. – с. 185 – 229.
17. Будыко М.И. Изменение климата. – Л.: Гидрометеиздат, 1974. – 280 с.
18. <http://ssd.jpl.nasa.gov>. – NASA, Jet Propulsion Laboratory California Institute of Technology (JPL Solar System Dynamics). Электронный ресурс национального аэрокосмического агентства США.
19. Федоров В.М. Теоретический расчет межгодовой изменчивости солнечной постоянной // *Астрономический вестник*, 2012. – т. 46. – № 2. – с. 184 – 189.
20. Федоров В.М. Межгодовые вариации продолжительности тропического года // *Доклады РАН*, 2013. – т. 451. – № 1. – с. 95 – 97. DOI: 10.7868/S086956521319016X
21. Fedorov V.M. Interannual Variability of the Solar Constant // *Solar System Research*, 2012. – v. 46. – №. 2. – p. 170 – 176.
22. Fedorov V.M. Interannual Variations in the Duration of the Tropical Year // *Doklady Earth Sciences*, 2013. – Vol. 451. – Part 1. – p. 750–753.
23. Пальмен Э. Ньютон Ч. Циркуляционные системы атмосферы. – Л.: Гидрометеиздат, 1973. – 616 с.
24. Оорт А.Х. Балансовые соотношения в земной климатической системе / *Динамика климата*. Ред. С. Манабе. –Л.: Гидрометеиздат, 1988. – с. 91 – 113.
25. Гилл А. Динамика атмосферы и океана. – М.: Мир, 1986. – т.1. – 400 с. – т.2. – 415 с.
26. Дроздов О.А., Васильев Н.В., Раевский А.Н., Смекалова Л.К., Школьный В.П. Климатология. – Л.: Гидрометеиздат, 1989. – 568 с.
27. Rossby C.G. Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacement of the semi-permanent centers of action // *J. Marine Res.*, 1939. – v. 2. – № 1. – p. 38 – 55.
28. Петерсен С. Анализ и прогноз погоды. – Л.: Гидрометеиздат, 1961. – 652 с.

29. Усманов Р.Ф. О причинах формирования планетарной фронтальной зоны и субтропической полосы высокого давления // *Метеорология и гидрология*, 1953. – № 3. – с. 41 - 45.
30. Уоллос Дж. М., Лау Н.К. Роль превращения баротропной энергии в общей циркуляции атмосферы / *Динамика климата*. Ред. С. Манабе. – Л.: Гидрометеоиздат, 1988. – с. 50 – 90.
31. Wells N. *The atmosphere and ocean: A physical introduction*. – N.Y.: John & Sons, 1977. – 394 p.
32. Peixoto J.P., Oort A.H. *Physics of climate* // *Rev. Modern. Phys.*, 1984. – v. 56. – p. 365 – 429.
33. Анищенко В.С., Нейман А.Б., Мосс Ф., Шиманский – Гайер Л. Стохастический резонанс как индуцированный шумом эффект увеличения порядка // *Успехи физических наук*, 1999. – т. 169. – № 1. – с. 7 – 38.
34. Anishchenko V., Astakhov V., Neiman A., Vadivasova T., Schimansky-Geier L. *Nonlinear Dynamics of Chaotic and Stochastic Systems / Tutorial and Modern Developments*. – Berlin: Springer Publ., 2002. – 374 p.
35. Бухаров, О. Е. Разработка оболочки системы прогнозирования временных рядов с использованием эволюционных алгоритмов / О. Е. Бухаров, Д. П. Боголюбов // *Новые информационные технологии: Сборник трудов XVI Всероссийской научно-технической конференции*. - М.: МГУПИ, 2013. - С. 63-68.
36. Бухаров, О. Е. Разработка оболочки системы поддержки принятия решений с использованием эволюционных алгоритмов / О. Е. Бухаров, А. А. Мизикин, Д. П. Боголюбов // *Промышленные АСУ и Контроллеры*. - М.: НАУЧТЕХЛИТИЗДАТ, №7 2013. - С. 37-45.