

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ — ПРОЦЕССОВ УПРАВЛЕНИЯ

Иванов Данил Андреевич

Магистерская диссертация

**Долгосрочное прогнозирование ледовой
обстановки в Арктике**

Направление 010300

Фундаментальные информатика и информационные технологии

Магистерская программа «Технологии баз данных»

Руководитель магистерской программы,
кандидат физ.-мат. наук,
доцент, зав. каф. технологии программирования

Сергеев С.Л.

Научный руководитель,
ст. преподаватель

Севрюков С.Ю.

Рецензент,
доктор физ.-мат. наук,
исп. директор Лаборатории спутниковой океанографии

Кудрявцев В.Н.

САНКТ-ПЕТЕРБУРГ
2013

СОДЕРЖАНИЕ

Введение	3
Цель	5
Задачи	5
Глава 1. Литературный обзор	8
1.1. Ледовая обстановка в Арктике	8
1.2. Методы дистанционного зондирования Земли	10
1.3. Прогнозирование временных рядов	24
1.4. Геоинформационные системы	35
Глава 2. Экспериментальная часть	47
2.1. Выбор спутниковых данных	47
2.1.1. Продукт Special Sensor Microwave/Imager	47
2.1.2. Продукты National Snow & Ice Data Center	49
2.2. Модель данных	52
2.3. Подготовка данных	56
2.4. Анализ данных	58
2.4.1. Экстент ледяного покрова	58
2.4.2. Концентрация льда	60
2.4.3. Пространственные характеристики	65
2.5. Выбор модели и метода прогнозирования	67
2.6. Определение глубины прогноза	76
2.7. Прогнозирование	76
2.8. Система мониторинга	79
Заключение	87
Дальнейшая работа	89
Список литературы	90

Введение

В последние годы, как в научном сообществе [22, 20, 28, 37], так и в средствах массовой информации увеличивается интерес к Арктическому региону. Повышение внимания вызвано резким уменьшением площади поверхности ледяного покрова.

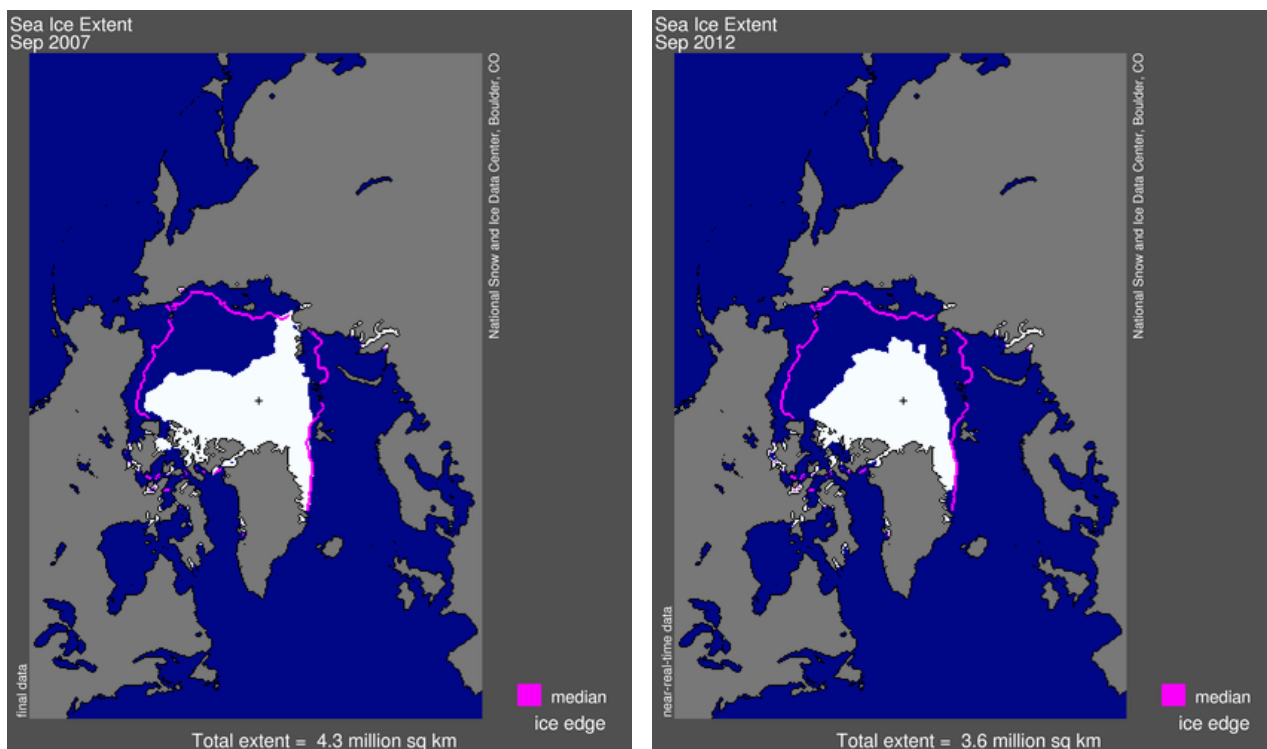
По данным National Snow and Ice Data Center в сентябре 2012 года был зарегистрирован самый низкий показатель с 1979 года, площадь оказалась вдвое меньше средней за 1979 и 2000 годы [23].

С одной стороны, с уменьшением площади открывается доступ к новым ресурсам:

- Открываются новые зоны рыбной ловли.
- По данным Геологической службы США [7] в арктическом регионе скрыты до 30% мировых запасов газа и 13% нефти.
- Появляется возможность круглогодичного доступа к северному морскому пути;
- Появляется возможность создания новых морских путей и туристических маршрутов.

Однако, с другой, уменьшение площади ледового покрова влечет за собой и другие климатические изменения: нарушение глобальной циркуляции воздушных масс, изменения течений и другие масштабные процессы.

Точные последствия изменений такого характера на данный момент до конца не известны. Существует много проектов, имеющих цель научно описать возможные климатические изменения. Одним из таких про-



(а) Сентябрь 2007

(б) Сентябрь 2012

Рисунок 1 – Карта ледового покрытия, отображающая минимумы за 2007 и 2012 годы по данным [32]

ектов является проект «SWIPA». Важные выводы проекта приведены в Разделе 1.1 **Ледовая обстановка в Арктике**.

Стоит также отметить что таяние льда приобретает аномальный характер. Совсем недавно было принято решение о проведении экстренной эвакуации дрейфующей полярной станции СП-40 [19]. Изначально база располагалась на большой льдине, площадью 9 км². В настоящее время из-за произошедшего раскола льда, который связывают с резким весенним потеплением, СП-40 размещается на льдине 100 на 150 метров.

На сегодняшний день вопрос прогнозирования конфигурации льда в Арктическом регионе остается открытым. Данный прогноз необходим для дальнейшего анализа обстановки в Арктическом регионе.

К примеру, объединяя полученные результаты с моделями ветра, можно рассмотреть как изменятся характеристики волн и волновая обстановка в регионе в целом. Такая информация необходима при расположении новых и обеспечения безопасности на уже существующих

буровых вышках, прокладывание новых и изменения существующих торговых и туристических маршрутов.

При построения ледовых прогнозов на основе физических моделей, долгосрочным считается прогноз на срок от полугода [4]. Прогноз, рассматриваемый в работе имеет скорее отношение к экологии, где долгосрочным принято считать прогноз на срок от 8 до 15 лет.

Цель

В связи с этим, **целью данной работы** является исследование методов и моделей прогнозирования для создание долгосрочного прогноза конфигурации льда в Арктике на основе спутниковых данных.

Данная цель является частью работы по созданию системы мониторинга Арктики. Базовая часть функциональности системы мониторинга разработана на основе модулей и программного обеспечения, написанного в ходе работы над магистерской диссертацией.

Задачи

Для достижения поставленной цели необходимо **решить следующие задачи:**

- Рассмотреть существующие глобальные прогнозы обстановки в Арктике;
- Рассмотреть область спутниковой океанографии и инструментария обработки спутниковых данных;
- Создать фреймворк, позволяющий работать с коллекциями спутниковых данных;
- Рассмотреть методы и модели прогнозирования, которые можно использовать для анализа и прогнозирования полей параметров представленных в спутниковых данных;

- Построить долгосрочный прогноз конфигурации льда;
- Предложить метод оценки адекватности, полученной модели;
- Рассмотреть инструментарий создания геоинформационных систем с целью выбора наиболее подходящего;
- Создать прототип системы визуализации данных во времени с возможностью коллокации различных слоев;

Структура работы

Данная работа выполнена на пересечение трех областей:

- Спутниковая океанография и методы дистанционного зондирования;
- Прогнозирование временных рядов;
- Геоинформационные системы.

Как следствие этого структура следующих глав имеет три раздела, связанных с обсуждением вышеперечисленных областей и дополнительные разделы, описывающие их пересечения.

В первой главе проведен обзор литературы, в который вошли:

- обзор ледовой обстановки в Арктике;
- введение в спутниковую океанографию и методы дистанционного зондирования;
- обзор методов прогнозирования временных рядов;
- обзор инструментария для создания геоинформационных систем.

Во второй главе приведено:

- описание модели представления данных;
- описание выбора спутникового продукта, используемого для построения прогноза;

- описание процесса подготовки данных и их анализ;
- описание выбора метода для построения прогноза;
- описание функциональных требований, архитектуры и особенностей реализации системы мониторинга Арктики.

В заключение приведены результаты, которые были получены при выполнении поставленных задач.

В разделе **Дальнейшая работа** описаны предполагаемые направления для улучшения: прогноза, полученного фреймворка для анализа, а так же системы мониторинга Арктики.

Глава 1

Литературный обзор

1.1 Ледовая обстановка в Арктике

Одной из самых всеохватывающих научных подборок о влиянии климатических изменений на замерзающие регионы Арктики является криосферный¹ проект «Snow, Water, Ice, Permafrost in the Arctic» (SWIPA, перевод «Снег, Вода, Лед и Вечная мерзлота в Арктике»). По текущим выводам проекта SWIPA, в Арктике происходят беспрецедентные изменения. Далее следуют основные выводы проекта SWIPA, приведенные в [16]:

- В последнее время происходят значительные изменения морского льда, в массе Гренландского ледового покрытия и Арктических ледовых шапках и ледниках. В прогнозах Международной комиссии ООН по изменениям климата от 2007 года была недооценена степень изменений, наблюдаемых в настоящее время.
- Уровень потепления Арктики из-за климатических изменений в два раза выше, чем в среднем по миру, начиная с 1980 г. Летняя температура воздуха в Арктике, наблюдаемая в течение последних десятилетий, выше, чем в любой заданный период в течение последних 2 000 лет.

¹Криосфера (от др.-греч. *κρυος* — холод, *σφαίρα* — шар)

- Потепление имеет место практически во всех замерзающих частях Арктики. Температура слоя вечной мерзлоты повысилась в среднем на 2 градуса, и площадь почти всех ледников и ледовых шапок Арктики начиная с 2000 г. сокращается в два раза быстрее, чем в предшествующее этому десятилетие.
- Это приводит к фундаментальным изменениям арктических экосистем и, в некоторых случаях, к уничтожению конкретных сред обитания.
- Доступ к ресурсам и транспортные опции радикально изменились. Арктическая инфраструктура может понести ущерб из-за этих изменений.

В своей оценке SWIPA описывает возможные будущие изменения. Ожидается, что средние температуры осенне-зимнего сезона в Арктическом регионе повысятся от 3 до 7 градусов к концу 21-го века (2080). Предполагается, что горные ледники и ледовые шапки потеряют от 10% до 30% от своей суммарной массы к 2100 г. Согласно прогнозам, Северный Ледовитый Океан полностью освободится ото льда в летний период в течение этого столетия, возможно в ближайшие 30-40 лет.

Так же имеются доказательства существования механизмов ответной реакции, которые способны ускорить процесс таяния льдов. В качестве примера приводится взаимодействие снега и морского льда с климатической системой. По мере того, как сокращаются снеговые и ледяные отражательные поверхности, более темные участки поглощают большее количество солнечной энергии и увеличивают потепление.

В основу магистерской работы легла одна из рекомендаций, приведенная учеными участвующими в проекте SWIPA, в качестве предлагаемых действий. Этой рекомендацией является *улучшение систем наблюдения и прогнозирования, а так же проведение дополнительных оценок криосферных изменений.*

1.2 Методы дистанционного зондирования Земли

«Дистанционное зондирование» — это получение информации об объекте без вступления с ним в прямой контакт. Таким образом, основной идеей, заложенной в дистанционном зондировании, является возможность удаленно по внутренним и/или внешним признакам определять характеристики объектов или феноменов. Первыми из таких попыток по изучении скорости ветра над океаном были работы английского адмирала Френсиса Бофорта. В 1806 году им была разработана шкала, позволяющая производить приближённую оценку ветра по его воздействию на наземные предметы или по волнению в открытом море [27]. Расширенная шкала Бофорта используется и сегодня.

Ниже приведены события и даты, сыгравшие важную роль в истории становления спутниковой океанографии:

- 1826 год — Жозефом Нисефором Ньепсом была получена первая фотография;
- 1858 год — Гаспар Феликс Турнашон делает первый аэрофотоснимок с воздушного шара;
- 1913 год — Первые аэрофотоснимки сделанные с аэроплана;
- 1935 год — Изобретение радара;
- 1942 год — Kodak патентует фотопленку, чувствительную к инфракрасному излучению.
- 1950е годы — Первый бортовой тепловой сканер;
- 1957 год — СССР запускает первый искусственный спутник — «Спутник-1»;
- 1957 год — Изобретение радара с синтезированной апертурой;
- 1960 год — Запуск первого метеорологической спутника США — «TIROS-7»;
- 1962 год — Первый бортовой многоспектральный сканер.

В последние годы инновацией стали спутники малых размеров «CubeSat». Объем спутника составляет 1 литр, размер соответственно 10x10x10 см³ [26], а вес 1.33 кг. На борту такого спутника можно разместить один или два сенсора.

На сегодняшний день благодаря методам спутниковой океанографии ведется наблюдение за температурой морской поверхности, исследуется взаимодействие атмосферы и океана, осуществляется мониторинг океанических волн, производится анализ океанических течений и завихрений, измеряется биологическая продуктивность, строятся прогнозы погоды, анализируется ледовая обстановка.

Основы спутниковой океанографии, её приложения и методы отлично изложены в книге «Measuring the oceans from space: the principles and methods of satellite oceanography» [35]. При написании данного раздела книга являлась опорным материалом.

Далее в разделе *a)* вводится базовая терминология; *b)* описывается процесс получения спутниковых продуктов; *c)* рассматриваются характеристики, определяющие качество продуктов; *d)* определяются уровни обработки данных; *e)* описываются основные форматы данных, используемые для удобного хранения и обмена; *f)* рассматривается программное обеспечение, которое можно использовать для анализа данных.

В заключительной части раздела будут рассмотрены существующие продукты и способы работы с данными для изучения ледовой обстановки в Арктике.

Терминология

Под термином «дистанционное зондирование» принято понимать зондирование поверхности Земли авиационными и космическими средствами с использованием свойств электромагнитных волн, излучаемых, отражаемых или рассеиваемых зондируемыми объектами, с целью лучшего

распоряжения природными ресурсами, совершенствования землепользования и охраны окружающей среды [5].

Также на основе документа «Принципы, касающиеся дистанционного зондирования Земли из космического пространства», введем другие термины.

Первичные данные — необработанные данные, которые получаются с помощью аппаратуры дистанционного зондирования, установленной на борту космического объекта, и которые передаются или доставляются на Землю из космоса посредством телеметрии в виде электромагнитных сигналов, фотопленки, магнитной ленты или какими-либо другими способами.

Обработанные данные (или Продукт) — материалы, полученные в результате такой обработки первичных данных, которая необходима для обеспечения возможности пользоваться этими данными. Гранулами будем называть файлы отдельного продукта.

Проанализированная информация — означает информацию, полученную в результате интерпретации обработанных данных, дополнительно введенных данных и сведений из других источников.

Параметр — измеряемое значение, полученное в результате обработки сырых данных, например концентрация льда или температура морской поверхности.

Сенсор (или прибор) — техническое устройство, используемое для зондирования поверхности.

Платформа — носитель для сенсоров. В качестве платформы могут выступать самолеты или вертолеты (такая платформа считается суборбитальной) или искусственные спутники земли (платформа считается орбитальной).

Процесс создания продуктов

На Рис. 1.1 приведена схема информационных потоков в спутниковой океанографии, этой схемой можно описать процесс реализации продукта [35, глава 2 стр. 13]. При помощи сенсоров, находящихся на спутнике, производится дистанционное зондирование океана, данные записываются на внутренние устройства хранения. После этого данные передаются на наземную принимающую станцию и распространяются между центрами обработки.

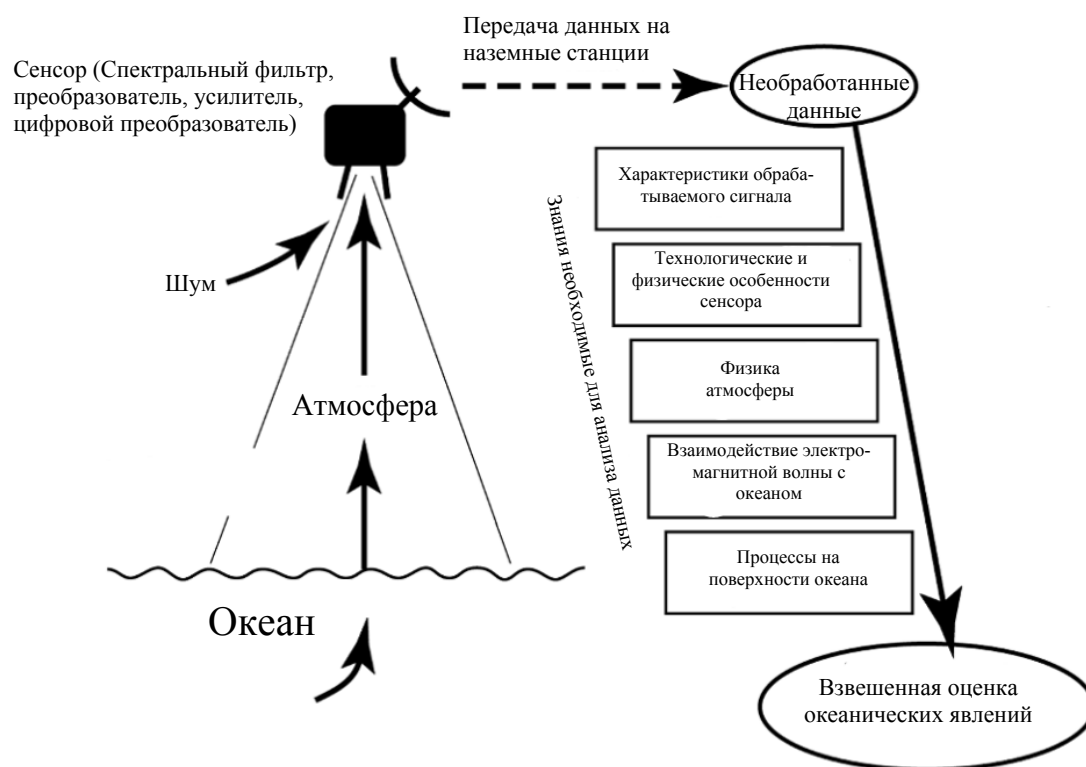


Рисунок 1.1 – Информационные потоки в спутниковой океанографии.

Справа расположены области знания, в которых необходимо разбираться для понимания методов спутниковой океанографии.

Для того чтобы получить готовый продукт из необработанных данных необходимо обладать комплексом знаний:

- о характеристиках обрабатываемого сигнала;
- о технологических и физических особенностях сенсора;

- о физике атмосферы;
- о взаимодействиях электро-магнитных волн с океаном;
- о процессах на поверхности океана.

В результате будет получен продукт со взвешенной оценкой океанических явлений. На сегодняшний день, есть возможность получать информацию о следующих полях и явлениях:

- температура поверхности океана;
- соленость на поверхности океана;
- морские течения, динамика водных масс;
- состояние поверхности моря, волнение;
- уровень моря;
- внутренние волны;
- приводный ветер;
- цвет воды, биопродуктивность;
- морские льды;
- мезо/мелкомасштабные явления на морской поверхности;
- рельеф дна.

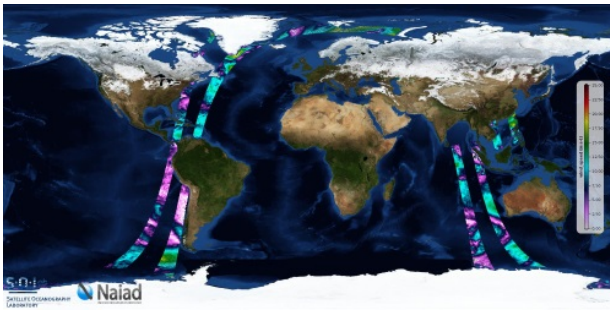
Характеристики спутниковых данных

Уровни обработки данных

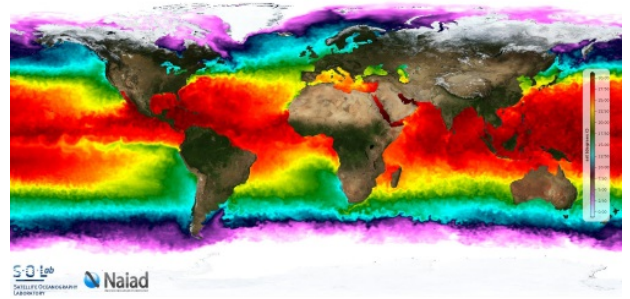
Для определения степени обработки продукта сформулирована классификация по уровням обработки данных дистанционного зондирования:

Данные **0**-го уровня (англ. Raw data) — данные, поступающие непосредственно от устройства, без служебных данных.

Данные уровня **1a** — реконструированные данные устройства, снабженные маркерами времени, радиометрическими коэффициентами, орбитальными координатами спутника.



(а) Продукт второго уровня



(б) Продукт третьего уровня

Рисунок 1.2 – Примеры изображений продуктов.

Данные уровня **1b** — данные уровня 1a, преобразованные в физические единицы измерения.

Данные **2**-го уровня (англ. Swath data, дословно: полосовые данные) — производные геофизические переменные или параметры (высота океанических волн, влажность почвы, концентрация льда) с тем же разрешением, как у данных уровня 1. В файлах таких продуктов должны храниться значения параметра и их геопривязка. На Рис. 1.2, а изображен снимок, сделанный за время одного витка спутника вокруг Земли. На глобальной карте представлены в виде двух полос в силу особенности сенсора.

Данные **3**-го уровня (англ. Gridded data, дословно: гридированные данные) — переменные, отображенные в универсальной пространственно-временной шкале, возможно усредненные либо дополненные интерполяцией. Пример продукта третьего уровня приведен на Рис. 1.2, б.

Данные **4**-го уровня (англ. Blended data, дословно: Смешанные данные) — данные, полученные в результате расчетов на основе предыдущих уровней.

Данные, на которых будет производиться дальнейший анализ, являются данными третьего и четвертого уровней.

Характеристики, определяющие качество данных

Характеристиками определяющими качество данных дистанционного зондирования являются:

1. Пространственное разрешение. Определяется размером пикселя на поверхности Земли, записываемого в растровое изображение.
2. Временное разрешение. Определяется частотой пролета спутника над областью интереса.
3. Спектральное разрешение. Определяется количеством спектральных полос.
4. Радиометрическое разрешение. Определяется количеством уровней сигнала, которые сенсор способен регистрировать.

Пространственное покрытие

Пространственное покрытие — характеристика, определяющая область, для которой производится конкретный продукт. Пространственное покрытие бывает двух типов: глобальное и локальное. Глобальные продукты охватывают весь земной шар, примером таких продуктов может являться продукт под названием «Global Odyssey Sea Surface Temperature Analysis», производимый во французском научно-исследовательском институте по эксплуатации морских ресурсов (IFREMER). Данный продукт предоставляет информацию о температуре морской поверхности.

К локальным продуктам относятся продукты региональной спутниковой океанография, такие продукты обычно отличаются более высоким пространственным разрешением. Примером такого продукта с локальным пространственным покрытием является «Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data

(Arctic, Monthly)», благодаря данному продукту можно получить концентрацию льда по Арктическому региону. Данный продукт получают из глобального продукта SSM/I. Существуют изначально локальные продукты, полученные например на геостационарных спутниках. Особенность геостационарных спутников заключается в том, что они вращаются вместе с Землей и за счет этого их положение относительно некоторого региона на земле остается стационарным. Это позволяет производить постоянный мониторинг интересующего региона.

Форматы данных

Для хранения и удобства обмена данными принято использовать следующие форматы [30]:

1. **HDF** (Hierarchical Data Format) — Иерархический формат данных предназначен для эффективного хранения и доступа к большому объему сложных данных. Кроме описания в формат также входят и программное обеспечение для доступа и работы с HDF.
2. **NetCDF** (network Common Data Form) — формат представляющий собой интерфейс для доступа к данным, хранящихся в виде массивов. С форматом свободно распространяются библиотеки для языков C, FORTRAN, C++, JAVA и PYTHON, которые реализуют интерфейс доступа к данным.
3. **ASCII** (An American Standard Code for Information Interchange) — Американский стандартный код для обмена информацией, текстовый файл, в котором каждый байт представляет собой один символ в соответствии с кодом ASCII.
4. **TIFF** (Tagged Image File Format) — Растровый формат для хранения, передачи, отображения и печати растровых изображений. Формат может быть использован для хранения и передачи цифровых спутниковых снимков, отсканированных аэрофотоснимков,

моделей рельефа и т.п. TIFF поддерживает поддерживает сжатие данных, включение тайлов и возможность включению географических метаданных.

5. **GeoTIFF** (Geographical Tagged Image File Format) — данный формат является расширением формата TIFF. В качестве дополнительного слоя информации используются теги географических (или картографических) данных. Географические данные могут быть использованы для размещения изображения на карте.
6. **Binary** — файлы в данном формате являются двоичными файлами для чтения. К такому типу данных относятся гранулы продукта SSMI, о котором в будущем пойдет речь.
7. **Shapefile** - векторный формат географических файлов. Все объекты описываются с помощью следующих примитивов: точки, линии, и полигоны.

Инструменты для обработки

Для работы со спутниковыми данными можно использовать специализированное программное обеспечение, предназначенное для визуализации и анализа, так же можно использовать геоинформационные системы и инструменты для обработки изображений:

- **BEAM Software** — Настольное приложение, позволяющее визуализировать и анализировать данные начиная со второго уровня представленные в различных форматах.
- **ODESA** — предоставляет пользователям возможность работать с данными 2-го уровня сенсора MERIS.
- **ArcGIS and satellite data** — расширение для ArcGIS, упрощающее доступ к спутниковым данным. Доступ к данным осуществляется посредством взаимодействия с THREDDS каталогом.

- **WIM** (Windows Image Manager) — инструмент для визуализации и обработки спутниковых данных общего назначения.

Для более гибкой обработки можно использовать **MATLAB**² или любой язык программирования, для которого существуют модули для обработки изображений. В научном сообществе в последнее время большое распространение получил язык программирования **Python**.

Язык программирования Python

Язык Python приобретает популярность в научном сообществе за счет своей простоты и выразительности. Для Python есть хорошие модули, позволяющие проводить все необходимые вычисления.

IPython — это интерактивная оболочка программирования Python, предоставляющая расширенную интроспекцию и много других расширений языка. Не так давно на свет появился проект «The IPython Notebook», представляющий из себя интерактивную вычислительную среду, реализованную в виде веб-приложения. IPython Notebook, позволяет создавать «блокноты», в которых можно совмещать, как математические вычисления, так и текст, видео и tex-формулы. Сайт проекта: ipython.org/notebook.html.

Для удобства расчетов можно использовать следующие модули:

- **NumPy** (англ. Numerical Python) — расширение, добавляющее поддержку больших многомерных массивов и матриц. Сайт проекта: www.numpy.org.
- **SciPy** (англ. Science Python) — это открытая библиотека высококачественных научных инструментов. В SciPy реализованы модули для обработки сигналов, оптимизации, интегрирования, обработки изображений, решения обычных дифференциальных уравнений. В

²MATLAB (сокращение от Matrix Laboratory) — пакет прикладных программ для решения задач технических вычислений и одноименный язык программирования

SciPy реализованы модули для большинства повседневных задач, с которыми приходится сталкиваться при исследованиях. Сайт проекта: www.scipy.org.

- **Matplotlib** — библиотека для визуализации данных двумерной и трехмерной графикой. Принцип работы очень похож на процесс построения графиков в MATLAB. Сайт проекта: matplotlib.org.

Для доступа к спутниковым данным можно использовать:

- **pydap.client** — пакет, обеспечивающий доступ к данным посредством OPeNDAP сервера.
- **Scientific.IO.NetCDF** — пакет, обеспечивающий доступ к NetCDF файлам и работающий с OPeNDAP сервером. Не поддерживается 4-ая версия стандарта NetCDF.
- **netcdf4-python** — пакет, обеспечивающий доступ, как к данным хранящимся в NetCDF формате, так и к серверу OPeNDAP.

Для облегчения процесса установки можно воспользоваться собранным дистрибутивом под названием «Anaconda». Сайт проекта: docs.continuum.io/anaconda.

Изучение ледовой обстановки

Существует множество продуктов для изучения ледовой обстановки. Ознакомиться с этими продуктами можно в интернет каталогах спутниковых данных. Одним из передовых каталогов для поиска спутниковых данных на сегодня является проект «PO.DAAC». Данный проект является распределенным архивом спутниковых данных, имеющим в наличии 36 различных продуктов для анализа ледовой обстановки. Тем не менее, среди них нет продуктов, которые бы позволили сделать долгосрочный прогноз ледовой обстановки в Арктике.

Существуют и другие проекты подобного рода, например проект «SATIN» — SATellite Data Search and Manage INformation Portal [36], разрабатываемый в Лаборатории спутниковой океанографии (SO_{Lab}.rshu.ru). Базовый функционал данного каталога изложен в докладе [17] и на самом сайте каталога (satin.rshu.ru).

В каталоге SATIN присутствуют несколько продуктов, на базе которых можно производить прогнозирование.

SSMI

Полное название: Special Sensor Microwave/Imager NASA's Measures Program (Daily). SSMI является глобальным гридированным³ продуктом с шагом сетки в 0.25 градуса. Данный продукт предназначен для исследования полей ветра, водяного пара в атмосфере и водозапаса облаков. Поле льда представлено здесь в качестве битовой маски, то есть для каждого пикселя доступна информация о том, покрыта ли соответствующая область льдом или нет.

Данные SSMI доступны с 1987 года по текущий момент. Это позволяет учесть при прогнозировании резкое уменьшение площади ледяного покрова в Арктике в 2012 году, в отличие от данных продукта ASI-AMSRE, описанного далее.

Гранулы SSMI распространяются в бинарном формате. Существует однодневные, трехдневные, недельные и месячные изображения.

ASI-AMSRE

Полное название: Sea-Ice concentration for Arctic & Antarctic. Гридированный продукт 4-го уровня с пространственным разрешением равным 6.25 км, хранящийся в формате NetCDF.

³Гридированный продукт — продукт, в котором поле параметра (льда, ветра и т.п.) представляется в виде матрицы. Каждая ячейка матрицы соответствует фиксированной области на земле с размерами заданными либо в пикселях, либо в километрах. Значение ячейки получается осреднением значений измерений в этой ячейке

Продукт доступен начиная с 19 июня 2002 года по 30 сентября 2011. ASI-AMSRE перестали производить по причине отказа антенны на сенсоре AMSR-E, установленном на спутнике Aqua.

Доступ к данным можно получить посредством модуля netcdf4-python, либо через API к OPeNDAP⁴ серверу.

NSIDC-SSMI-SEAICE-GSFC

Полное название: Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data.

Гридированный продукт 4-го уровня с пространственным разрешением 25 км. Уникальный в своем роде продукт, являющийся примером того как можно объединять в один архив и приводить к единому стандарту данные полученные с разных спутников и сенсоров [34]. Данные доступны начиная с 26 октября 1978 года по 31 декабря 2011 года, начиная с 2012 года создан новый продукт NSIDC-SSMI-NRT-SEAICE.

NSIDC-SSMI-NRT-SEAICE

Полное название: Near-Real-Time DMSP SSM/I-SSMIS Daily Polar Gridded Sea Ice Concentrations (Arctic).

Данный продукт имеет аналогичный формат что и NSIDC-SSMI-SEAICE-GSFC, предоставляет NRT (near-real-time — близкую к реальному времени) карту концентрации льда для обоих полушарий, на основе данных полученных сенсором Special Sensor Microwave Imager/Sounder (SSMIS), находящегося на борту спутника F17, запущенного по программе Defense Meteorological Satellite Program (DMSP) [33].

Данные доступны с января 2012 года по настоящее время. Более подробно продукты от NSIDC будут рассмотрены во второй главе.

⁴OPeNDAP (англ. Open-source Project for a Network Data Access Protocol) — это архитектура и протокол передачи данных, созданный специально для доступа к данным дистанционного зондирования.

Другие продукты

Стоит отметить, что существуют продукты с информацией о толщине и о возрасте льда. В работе [25] используются данные со спутника CryoSat-2 Европейского Космического Агентства (The European Space Agency). Такие данные позволили бы улучшить прогноз, однако доступ к ним получить не удалось.

1.3 Прогнозирование временных рядов

В основу обзора по данной части легли книги, в которых изложена теория временных рядов, их анализа и прогнозирования [1, 18], с некоторыми дополнениями из работ [9, 10, 11, 13].

Временной ряд — это совокупность значений какого-либо показателя за несколько последовательных моментов или периодов времени. Временные ряды обычно используются в сложных системах, когда нет возможности построить уравнение движения. На Рис. 1.3 представлен пример временного ряда концентрации льда в пикселе с координатами (80, 137).

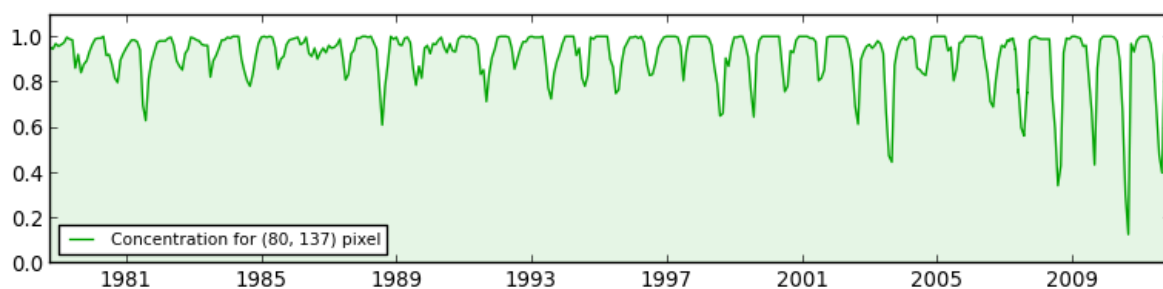


Рисунок 1.3 – Временной ряд концентрации льда, осредненных за месяц, в пикселе с координатами (80, 137)

В зависимости от способа определения значений, временные ряды делятся на:

- интервальные;
- моментальные.

Моментальными временными рядами считаются ряды значения в которых получены в результате измерения в конкретный момент времени. В отличие от значений моментального временного, значения интервального являются накоплением за определенный интервал времени. На Рис 1.3 значениями временного ряда являются значения концентрации льда в конкретном пикселе.

Так же временные ряды принято классифицировать:

- по форме представления уровней:
 - абсолютные показатели;
 - относительные показатели;
 - средние величины.
- по количеству показателей, для которых необходимо определить уровень в каждый момент времени: одномерные и многомерные временные ряды;
- по расстоянию m/u временными отметками: равностоящие и неравностоящие;
- по наличию пропущенных значений: полные и неполные временные ряды.

Значения (или уровни) временного ряда складываются из двух составляющих:

- регулярной (так же в литературе встречается: детерминированной, систематической);
- нерегулярной (случайной, не зависящей от времени).

В общем случае регулярная составляющая может состоять из: тренда, сезонной и циклической компонент. Нерегулярная составляющая формируется обычно в результате:

- внезапных возмущения, например, стихийных бедствий, эпидемий и т.п., вызывает значительные отклонения со случайными колебаниями;
- текущие возмущения, формируются в результате действия большого числа неучтенных, но при этом не значительных причин.

Таким образом, временной ряд состоит из четырех компонент:

- тренда T_t ;
- сезонной компоненты S_t ;

- циклической компоненты C_t ;
- случайной компоненты E_t .

Нижним индексом t обозначается зависимость от времени. Циклическая компонента отличается от сезонной более длительным периодом, обычно циклическая компонента объединяется с трендом и обозначается как T_t . Взаимодействие этих четырех компонент определяет тип модели:

- аддитивная модель: $Y_t = T_t + S_t + E_t$;
- мультипликативная модель: $Y_t = T_t * S_t * E_t$;
- модель смешанного типа: $Y_t = T_t * S_t + E_t$.

Выбор модели представления временного ряда основывается на анализе структуры сезонных колебаний. Если амплитуда колебаний приблизительно постоянна, выбирают аддитивную модель, в противном случае выбирают мультипликативную. Построение моделей сводится к расчету значений T_t , S_t и E_t .

Прогноз (от греч. *προγινωσις* — предвидение, предсказание) — предсказание будущего с помощью научных методов, а также сам результат предсказания. Согласно, [1], процессы, для которых необходимо строить прогнозы, чаще всего можно представить в виде временных рядов.

Временем упреждения ряда, называет горизонт времени, ка которые необходимо определить значения временного ряда [1]. В зависимости от времени упреждения задачи прогнозирования делятся на следующие категории срочности:

- краткосрочное прогнозирование;
- среднесрочное прогнозирование;
- долгосрочное прогнозирование.

Стоит отметить, что для различных временных рядов, с различными временным разрешением классификация срочности задач прогнозирования индивидуальна. При построения ледовых прогнозов на основе физических моделей, долгосрочным считается прогноз на срок от полугода

[4]. Прогноз, рассматриваемый в работе имеет скорее отношение к экологии, где долгосрочным принято считать прогноз на срок от 8 до 15 лет.

В теории прогнозирования используется два взаимосвязанных понятия: метод и модель прогнозирования.

Метод прогнозирования описывает последовательность действий, для получения модели прогнозирования временного ряда. Так же метод прогнозирования, содержит критерии по оценке качества прогнозируемых значений.

Под **моделью** прогнозирования принято понимать функциональное представление, адекватно описывающее временной ряд. Модель является основой для получения прогноза.

Обзор моделей прогнозирования

В настоящее время существует более 100 классов моделей для прогнозирования временных рядов. В работе [12] все методы прогнозирования делятся на две группы: интуитивные и формализованные.

К интуитивным методам относят класс методов экспертных оценок: метод интервью, аналитический метод, сценарный метод, метод комиссии, метод мозгового штурма и метод Дельфи. Такие методы используются либо когда объект прогнозирования слишком прост, либо напротив, когда он слишком сложен, и частично или полностью не поддается математической формализации. В настоящее время экспертные системы широко распространены, в частности активно экспертные системы основанные на нечеткой логике.

К формализованным методам относятся методы связанные с построением математической модели процесса. К таким методам можно отнести модели предметной области и модели временных рядов.

Модели предметной области — математические модели, для построения которых используют законы предметной области. Например,

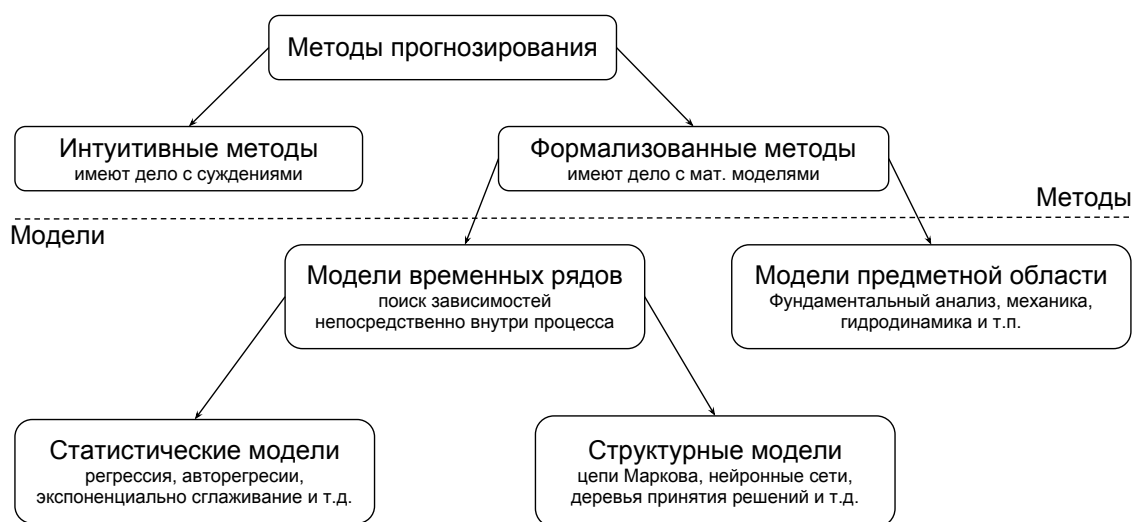


Рисунок 1.4 – Классификация методов и моделей прогнозирования [15].

при построения модели для прогнозирования конфигурации льда, такая модель бы содержала уравнения гидродинамики, уравнения взаимодействия атмосферы и океана и др.

Модели временных рядов — математические модели прогнозирования, стремящиеся найти зависимости будущего значения от прошлого внутри самого процесса. На основе такой зависимости можно производить прогноз. Такие модели являются универсальными для различных предметных областей.

Модели временных рядов могут быть разделены на статистические и структурные. В статистических моделях зависимость будущего от прошлого задается в виде уравнения.

К моделям временных рядов относятся:

- регрессионные модели;
- авторегрессионные модели (AR, ARMA, ARIMA);
- модель экспоненциального сглаживания;
- и т.д.

В структурных моделях зависимость будущего значения от прошлого описывается определенной структурой и правилами перехода по этой структуре. К структурным моделям относятся:

- нейронные сети;
- модели на базе Марковских цепей;
- и т.д.

Далее в главе будут рассмотрены основные модели временных рядов, на основе которых можно произвести прогноз конфигурации ледового покрова.

Трендовые модели

Если в модели отсутствует сезонная составляющая, то такую модель можно прогнозировать с помощью нахождения тренда с последующей его экстраполяцией. Тренд может быть описан различными функциями:

1. Линейная функция: $Y_t = \alpha_0 + \alpha_1 X_t + \varepsilon_t$;
2. Гипербола: $Y_t = \alpha_0 + \frac{\alpha_1}{X_t} + \varepsilon_t$;
3. Степенная: $Y_t = \alpha_0 X_t^{\alpha_1} + \varepsilon_t$;
4. Логарифмическая: $Y_t = \alpha_0 + \alpha_1 \ln X_t + \varepsilon_t$;
5. S-образная: $Y_t = \exp(\alpha_0 + \frac{\alpha_1}{X_t} + \varepsilon_t)$;
6. Полиномиальный тренд n-го порядка: $Y_t = \alpha_0 * X_t^{(n-1)} + \alpha_1 * X_t^{(n-2)} + \dots + \alpha_n + \varepsilon_t$.

В случае полиномиального тренда второго порядка возможен учет и циклической компоненты. Так же важным достоинством является устоявшиеся процессы анализа и проектирования. К *недостаткам* данного класса моделей относятся низкая адаптивность, отсутствие способности моделирования нелинейных процессов. К *достоинствам* простота и скорость их создания.

Регрессионные модели

Регрессионный анализ используется для решения задач, требующих изучения взаимосвязей между переменными. Регрессионный анализ для задач физики хорошо описан в книге [8]. Целью регрессионного анализа является определение зависимости между исходной прогнозируемой переменной и множеством внешних факторов.

Модель простой линейной регрессии. В основе данной модели лежит предположение о наличии линейной зависимости между процессом и внешним фактором, оказывающим на него влияние. Модель прогнозирования на основе линейной регрессии описывается уравнением:

$$Y_t = \alpha_0 + \alpha_1 X_t + \epsilon_t, \quad (1.1)$$

где α_0 и α_1 — коэффициенты регрессии, ϵ_t — ошибка модели. Из уравнения видно, что для получения прогнозных значений Y_t необходимо знать значение X_t , что на практике бывает редко допустимым.

Модель множественной регрессии. На процесс Y_t влияние могут оказывать целый ряд внешних факторов. Модель прогнозирования в данном случае представляется следующим уравнением:

$$Y_t = \alpha_0 + \alpha_1 X_{1t} + \alpha_2 X_{2t} + \dots + \alpha_n X_{nt} + \epsilon_t \quad (1.2)$$

Модель множественной регрессии имеет тот же самый недостаток, что и модель линейной регрессии — для получения прогнозных значений необходимо знать значения $X_{1t}, X_{2t}, \dots, X_{nt}$.

Модель нелинейной регрессии. В основе данной модели лежит предположение о наличии связи между процессом и внешним фактором, которую можно описать некоторой известной функцией:

$$Y_t = F(X_t, A) \quad (1.3)$$

На практике модели нелинейной регрессии встречаются редко, это связано с тем что встречается мало процессов для которых известна функциональная зависимость с их внешним фактором.

Достоинство регрессионных моделей относится простота, гибкость, скорость создания. Так же важным достоинством является устоявшиеся процессы анализа и проектирования. К *недостаткам* данного класса моделей относятся низкая адаптивность, отсутствие способности моделирования нелинейных процессов и трудоемкость определения параметров модели.

Авторегрессионные модели

В основе авторегрессионных моделей лежит предположение о наличии зависимости между текущим значением процесса и некоторым количеством предшествующих ему значений этого же процесса. Наиболее часто используется комбинация двух моделей: модели авторегрессии (AR, autoregressive) и модели скользящего среднего (MA, moving average).

Модель авторегрессии. В данной модели текущее значение процесса выражается как конечная линейная совокупность предыдущих значений и некоторого импульса называемого «белым шумом»

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + \varepsilon_t, \quad (1.4)$$

где α_0 — вещественная константа, $\alpha_1, \dots, \alpha_p$ — коэффициента, авторегрессии, ε_t — ошибка модели, p — порядок авторегрессии. Обычно модель (1.4) обозначается как $AR(p)$. Для нахождения коэффициентов и вещественной константы используется метод наименьших квадратов [18].

Модель 1.5 называется моделью скользящего среднего порядка q , обычно обозначаемая как $MA(q)$. Данная модель часто используется с моделью AR.

$$Y_t = \frac{(Y_{t-1} + Y_{t-2} + \dots + Y_{t-q})}{q} + \varepsilon_t, \quad (1.5)$$

где q — порядок скользящего среднего, ε_t — ошибка прогнозирования.

В работе [1] описывается, что для получения лучшего результата может быть целесообразно объединение моделей AR и MA.

Векторная авторегрессия. Модель динамики нескольких временных рядов, в которой текущие значения этих рядов зависят от прошлых значений этих же временных рядов. В литературе, модель обозначается как VAR. Модель представляет собой систему уравнение, каждое из которых является комбинацией модели авторегрессии и модели распределенного лага.

Достоинствами моделей данного класса является простота, прозрачность моделирования, а также хорошо развитый математический аппарат. *Недостатком* авторегрессионных моделей является большое число параметров модели, определение которых весьма затруднено. Большинство правил подбора коэффициентов не поддаются формализации и приходят с опытом, что накладывает ограничения на использование данной модели при построение прогноза для большого количества временных рядов. Проблема VAR-моделей заключается в резком росте количества параметров с увеличением количества анализируемых временных рядов и количества лагов.

Модели экспоненциального сглаживания

Модели экспоненциального сглаживания распространены в моделировании финансовых и экономических процессов. Модель экспоненциального сглаживания представляется в виде формулы:

$$\begin{cases} Y_t = S_t + \varepsilon_t \\ S_t = \alpha Y_{t-1} + (1 - \alpha)S_{t-1}, \end{cases} \quad (1.6)$$

где α — коэффициент сглаживания в интервале $(0, 1)$. Начальные условия: $S_1 = Y_0$.

В основе экспоненциального сглаживания заложена идея постоянного пересмотра прогнозных значений по мере поступления фактических. Последние доступные наблюдения имеют большее влияние на прогнозируемое значение в сравнение с предшествующими.

К классу моделей экспоненциального сглаживания так же относятся модель Хольта и модель Хольта-Винтерса. Модель Хольта применяется для моделирования процессов, имеющих тренд. В модели Хольта-Винтерса добавляется возможность учитывать сезонную составляющую.

К *достоинствам* моделей экспоненциального сглаживания относят простоту моделирования, единообразие анализа и проектирования. Недостатками моделей данного класса является, недостаточная гибкость и узкая применимость.

Другие методы и модели прогнозирования

Выше описана только малая часть методов и моделей благодаря которым можно прогнозировать процессы. В настоящее время распространены структурные модели на базе нейронных сетей, цепей Маркова. Так же для прогнозирования применяются: Fuzzy Logic (Нечеткая логика), Vector Support Machine (Метод опорных векторов) и др.

В области прогнозирования временных рядов появилась тенденция создания комбинированных моделей, позволяющих при правильном использовании заимствовать хорошие стороны моделей.

Построение модели

Построение моделей сводится к расчету значений T_t , S_t и E_t . Ниже приведен процесс построения модели, предлагаемый в работе [10]:

1. Выравнивание исходного ряда Y_t методом скользящего среднего;
2. Расчет значений сезонной компоненты S_t ;
3. Вычитание сезонной компоненты из исходных уровней ряда: $Y_t - S_t = T_t + E_t$;
4. Аналитическое выравнивание уровней $(T_t + E_t)$ и расчет значений T_t с использованием полученного уравнения тренда;

5. Расчет полученных по модели значений $(T_t + E_t)$;
6. Расчет абсолютных и/или относительных ошибок. Если из временного ряда удалить тренд (T_t) и периодические составляющие $(C_t$ и $S_t)$, то останется нерегулярная компонента (E_t) , так называемая, ошибка. Если полученные значения ошибок не содержат автокорреляции, ими можно заменить исходные уровни ряда и в дальнейшем использовать временной ряд ошибок (E_t) для анализа взаимосвязи исходного ряда и других временных рядов.

Оценка модели

Ранее было описано, что под моделью прогнозирования принято понимать функциональное представление, *адекватно* описывающее временной ряд. Для того чтобы оценить адекватность описания, необходимо исследовать остатки ε_i . Исследование остатков ε_i предполагает проверку наличия следующих пяти предпосылок МНК [18]:

1. случайный характер остатков;
2. нулевая средняя величина остатков, не зависящая от t ;
3. гомоскедастичность — дисперсия каждого отклонения ε_i одинакова для всех значений x ;
4. отсутствие автокорреляции остатков. Значение остатков ε_i распределены независимо друг от друга;
5. остатки подчиняются нормальному закону распределения.

Каждая предпосылка МНК играет свою роль в построение прогноза. Критерии на основе, которых можно проверить вышеперечисленные предпосылки будут рассмотрены в процессе построения модели.

1.4 Геоинформационные системы

Под геоинформационной системой (ГИС) подразумевается система, предназначенная для сбора, хранения, анализа и отображения пространственных данных и связанных с ними не пространственных, а также получение на их основе информации и знаний о географическом пространстве [3].

ГИС позволяет производить обработку информации, благодаря использованию следующих функций [2]:

- Ввод и редактирование данных;
- Поддержка моделей пространственных данных;
- Хранение информации;
- Преобразование систем координат и трансформация картографических проекций;
- Растрово-векторные операции;
- Измерительные операции;
- Полигональные операции;
- Операции пространственного анализа;
- Различные виды пространственного моделирования;
- Цифровое моделирование рельефа и анализ поверхностей;
- Вывод результатов в разных формах.

Классификации ГИС определяются решаемыми задачами, среди которых: учет ресурсов, оценка, анализ, управление, мониторинг и поддержка принятия решений. В [2] приведена классификация по функциональным возможностям, архитектурным принципам построения, пространственному охвату, проблемно-тематической ориентации и по способу организации географических данных.

Обзор инструментария для разработки

ГИС являются сложными комплексными системами, в которых необходимо учесть много различных факторов. В эти факторы входят способы и стандарты передачи, хранения анализа и визуализации данных. В данном разделе приведен обзор инструментария, который можно использовать для построения полноценной ГИС. В разделе [2.8 Система мониторинга](#) приведен список выбранного программного обеспечения, с обоснованием сделанного выбора.

В последнее время все больше и больше привычных пользователю приложений реализуются на веб-платформе. Такую тенденцию можно обосновать следующими преимуществами веб-приложений:

- нет дополнительных затрат на установку на пользовательской машине;
- нет затрат на обновление версии на пользовательской машине;
- возможность предоставления доступа в любую точку мира, при наличии в ней интернета;
- кроссплатформенность;
- более простая адаптация для мобильных платформ.

Естественно, есть ниша, которую настольные приложения будут занимать еще долго, однако функционал ГИС можно реализовывать в качестве веб-приложения.

Протоколы и стандарты

Базовым протоколом для обмена информацией в веб-ориентированных ГИС является HTTP⁵. В качестве форматов для

⁵HTTP (англ. Hypertext Transfer Protocol — «протокол передачи гипертекста») — протокол прикладного уровня в стеке TCP/IP.

обмена информацией клиентом и сервером распространены форматы: XML⁶, JSON⁷, HTML⁸.

В настоящее время стандарты и общие принципы в области разработки ГИС разрабатываются и декларируются международной некоммерческой организацией OPEN GIS Consortium (OGC). Одним из первых форматов для описание геопространственных данных, разработанных OGC, является формат GML⁹. Для описание трехмерных геопространственных объектов используется формат KML¹⁰.

В OGC были созданы спецификации, определяющие несколько типов картографических сервисов [14]:

1. **WMS** (Web Map Service):

- определяет параметры запроса и предоставления пространственной информации в виде графического изображения или набора объектов;
- описывает условия получения и обеспечивает доступ к информации о содержимом карты.

2. **WFS** (Web Feature Service):

- определяет условия получения и обновления пространственно привязанной информации клиентской частью приложения с использованием Geography Markup Language (GML);
- описывает стандартный интерфейс доступа к и манипуляции с географическими объектами с помощью HTTP-протокола.

⁶XML (англ. eXtensible Markup Language — «расширяемый язык разметки») — рекомендованный Консорциумом Всемирной паутины (W3C) язык разметки.

⁷JSON (англ. JavaScript Object Notation) — текстовый формат обмена данными, основанный на JavaScript. Более компактен чем XML

⁸HTML (англ. HyperText Markup Language — «язык разметки гипертекста») — стандартный язык разметки документов в Вебе

⁹GML (англ. Geography Markup Language — «географический язык разметки») — язык разметки, основанный на XML и предназначенный для описания географических примитивов

¹⁰KML (англ. Keyhole Markup Language) — язык разметки на основе XML для описания трехмерных геопространственных данных в программе «Google Планета Земля»

3. **WFS-T** (Web Feature Service Transactions):

- расширение спецификации WFS, добавляющее CRUD¹¹ функции.

4. **WCS** (Web Coverage Service):

- в отличие от WMS, данная спецификация разрабатывается для предоставления свойств и значений в каждой конкретной точке географического пространства, а не на создание готовых картинок, что позволяет проводить интерпретацию данных на клиентской стороне.

Данные спецификации являются основными для разработки геоинформационных веб-сервисов. В дополнение к этим спецификациям существуют стандарты:

- **SLD** Style Layer Descriptor — стандарт предназначенный для описания символики пространственных данных. Применяется в сервисах WMS и WFS
- **SOS** Sensor Observation Service — сервис для отображения показателей с датчиков. Данный стандарт хорошо подходит для отображения *In situ*¹² данных. Каждое наблюдение определяется четырьмя параметрами:
 - `eventTime` — время получения измерения;
 - `featureOfInterest` — объект, который был измерен;
 - `observedProperty` — измеряемая характеристика измерения;
 - `procedure` — каким образом было проведено измерение.

¹¹CRUD (англ. Create Read Update Delete) — сокращенное именование базовых функций: создание, чтение, обновление и удаление

¹²*In situ* (лат. на месте) — В океанографии обозначает данные измеренные контактным способом.

Более подробно ознакомиться со стандартами можно на официальном сайте OGC (www.opengeospatial.org/standards) или в документации, разработанной сообществом OSGeo (live.osgeo.org).

Серверная сторона

Основной частью любого приложения, реализующего функционал ГИС, является картографический сервер. Далее приведены краткое описание трех наиболее распространенных серверов с открытым исходным кодом.

Mapnik — проект с открытым исходным кодом, реализующий инструментарий для рендеринга карт, написанный на языках C++ и Python. Поддерживает все необходимые форматы, описанные на стр. 17 в описании используемых форматов для хранения и передачи данных дистанционного зондирования. Сайт проекта: mapnik.org.

MapServer — кроссплатформенный сервер с открытым исходным кодом для работы с географической информацией. Поддерживает все необходимые стандарты и спецификации OGC, а также необходимые форматы. Начиная с 6-ой версии, MapServer включает в себя систему кеширования. Присутствует поддержка написания плагинов на языках: PHP, Python, Perl, Ruby, .NET, и Java. Сайт проекта: mapserver.org.

FeatureServer — геоинформационный сервер поддерживающий RESTful интерфейс, реализует спецификацию WFS-Transaction. Особенностью данного сервера является наличие большого числа входных источников: DBM, GeoAlchemy (Расширение для SQLAlchemy, являющееся реализацией ORM для языка Python), OGR, SpatialLite (Позволяет хранить геопрозрачные данные в SQLite), Flickr (Сервис для публикации изображений с геопривязкой), WFS. Данные на выходе могут быть представлены в следующих форматах: CSV, DXF (AutoCAD), GeoJSON, GeoRSS Atom, GML, HTML, KML, Shapefile, SQLite БД. Сайт проекта: featureserver.org.

GeoServer — картографический сервер с открытым исходным кодом, поддерживает спецификации WMS, WFS, WCS. В отличие от MapServer'a есть поддержка спецификации WFS-Transaction, позволяющей не только получать данные для построения карт, но и редактировать их, с последующим сохранением на сервере. GeoServer, включает в себя систему кеширования - GeoCache, а так же систему управления файлами настроек и описания данных для проектов. Сайт проекта: geoserver.org.

Также существует ряд решений для предоставления удобного доступа к форматам, в которых принято хранить и распространять спутниковые продукты.

OPeNDAP (англ. Open-source Project for a Network Data Access Protocol) — это архитектура и протокол передачи данных, созданный специально для доступа к данным дистанционного зондирования. Сайт проекта: www.opendap.org.

ncWMS — Сервис веб-карт (Web Map Service), работающий с файлами в формате NetCDF. Так же ncWMS поддерживает доступ к OPeNDAP. В декабре 2012 года вышла первая версия данного проекта, с очень полезной функциональностью. Есть возможность, произведя минимальные настройки, отображать большинство необходимых данных в проекции Меркатора, с возможностью изменения градиента. Сайт проекта: www.resc.rdg.ac.uk/trac/ncWMS.

Кроме готовых серверных решений существуют много различных инструментов и фреймворков для создания серверов для реализации функциональности ГИС, например GeOrganizer, p.mapper и GeoDjango.

Django — свободный фреймворк, написанный на языке Python. Django дает возможность разрабатывать полноценные веб-приложения в краткие сроки. Из функциональности можно отметить следующее: ORM¹³, встроенный интерфейс администратора (обеспечивает интер-

¹³ORM (англ. Object-relational mapping, рус. Объектно-реляционное отображение) — технология программирования, связывающая реляционные БД и концепцию ООП. Позволяет предоставлять прозрачный доступ к БД, посредством операций с объектами (экземплярами классов)

фейс к CRUD операциям над описанными объектами), диспетчер URL, система кеширования, модульная архитектура приложений. Так же стоит отметить наличие хорошей документации, качественного документированного кода и доброжелательного сообщества. Сайт проекта: www.djangoproject.com.

GeoDjango — расширение фреймворка Django, позволяющего оперировать в нем пространственными данными. Использование данного фреймворка осложнено из-за ряда его зависимостей. Для установки фреймворка понадобится собрать или установить следующие библиотеки:

- **GEOS** (англ. Geometry Engine Open Source) — библиотека для работы с геометрическими объектами.
- **PROJ.4** — библиотека для работы с географическими проекциями.
- **GDAL** (англ. Geospatial Data Abstraction Library — библиотека абстракции гео-пространственных данных) — библиотека для чтения, записи и преобразования растровых геопространственных форматов данных.
- **PostGIS** — расширение для СУБД PostgreSQL, позволяющее работать с геопространственными данными.

Хранение пространственных данных

Хранение пространственных данных в ГИС можно реализовать двумя способами:

- Пространственные данные можно хранить в виде файлов в форматах перечисленных в разделе 1.2 на стр. 17.
- Базы данных, оптимизированные для хранения и выполнения запросов к данным о пространственных объектах, которые представлены некоторыми абстракциями: точка, линия, полигон и др.

На сегодняшний день большинство наиболее распространенных систем имеют изначальную поддержку пространственных данных, либо имеют расширения для их поддержки. Рассмотрим некоторые их пространственных баз данных:

- **Oracle Database.** Для работы с пространственными данными в Oracle Database существует Oracle Spatial — это компонент базы данных, состоящий из типов данных, набора функций и процедур, а также внешних по отношению к базе данных веб-сервисов, которые позволяют эффективно хранить, быстро получать доступ и анализировать пространственные данные.
- **Microsoft SQL Server.** С 2008 версии поддерживает пространственные, запросы к данным обеспечиваются за счет расширения синтаксиса SQL. Поддерживается индексация.
- **PostgreSQL.** Для работы с пространственными данными в PostgreSQL необходимо произвести установку расширения PostGIS. PostGIS реализует поддержку типа данных *geometry*, обеспечивает поддержку запросов и индексации пространственных данных.
- **MySQL.** В MySQL поддержка типа данных *geometry* реализована не в соответствии со спецификациями OGC, следовательно при работе с пространственными данными в БД MySQL необходимо учитывать особенности реализации.
- **SQLite.** SQLite является легковесной встраиваемой реляционной БД. Для работы SQLite не требует отдельно запущенного сервера, что может быть удобно при реализации настольных приложений. Для поддержки пространственных данных необходимо использовать SpatialLite.

Это далеко не весь перечень существующих баз данных с поддержкой пространственных данных, здесь были приведены только наиболее распространенные реляционные БД. Стоит отметить, что существуют расширения и для баз данных относящихся к классу NoSQL. Например существуют расширения для key-value базы данных MongoDB.

В силу особенностей объектно-ориентированных баз данных, поддержка в них пространственных данных присутствует, если реализованы классы в языке на которых данная база данных используется. Например, для Python существует объектно-ориентированная база данных ZODB, позволяющая хранить любые объекты, которые можно сериализовать.

Клиентская сторона

Для отображение карт, данных и сопутствующей информации существуют различные JavaScript библиотеки:

OpenLayers — Наиболее распространенная библиотека для отображения карт на стороне пользователя. Далее перечислены некоторые функциональные возможности:

- поддержка спецификаций WMS, WFS и WFS-T;
- поддержка любых проекций при помощи PROJ4.JS;
- поддержка стандарта SOS;
- поддержка кластеризации объектов.

Большим плюсом OpenLayers является хорошая документация и множество примеров. С полным списком функциональных возможностей OpenLayers можно ознакомиться на сайте проекта: openlayers.org.

GeoExt — Расширение для библиотеки OpenLayers, объединяющая её с библиотекой для реализации пользовательского интерфейса ExtJS. Так же как и у OpenLayers имеется хорошая документация и много примеров. Сайт проекта: www.geoext.org.

LeafLet — современная библиотека ориентированная на браузеры с поддержкой HTML5. Для расширения функциональности существу-

ет большое количество плагинов. Есть поддержка работы с различными проекциями. Сайт проекта: leafletjs.com.

MapQuery — jQuery плагин для OpenLayers. С плагином поставляются несколько функциональных виджетов для манипуляции со слоями. Сайт проекта: mapquery.org.

Polymaps — Библиотека ориентированная в первую очередь на работу с векторными форматами. Сайт проекта: polymaps.org.

Так же для отображения карт на стороне клиента существует ряд сервисов с открытым API: **GoogleMaps**, **BingMaps** и **Яндекс карты**. У данных сервисов есть один недостаток, который не позволяет использовать их в проекте по мониторингу Арктики — отсутствие возможность перепроецирования данных.

Библиотеки и фреймворки для разработки веб-приложений

Для того чтобы ускорить процесс реализации полноценного веб-приложения существует много различных решений.

jQuery — JavaScript библиотека, позволяющая легко получать доступ к элементам DOM¹⁴, обращаться к их атрибутам и содержимому. Так же библиотека предоставляет удобный API по работе с AJAX. Сайт проекта: jquery.org.

CoffeeScript — язык программирования, транслируемый в JavaScript. В CoffeeScript реализована функциональность схожая с языками Ruby, Python и Haskell, что позволяет улучшить читаемость кода и уменьшить его объем. Можно использовать любые JavaScript библиотеки. Сайт проекта: coffeescript.org.

Sass (англ. Syntactically Awesome Stylesheets) — метаязык на основе CSS3 (англ. Cascading Style Sheets — каскадные таблицы стилей). Добавляет в CSS вложенные правила, переменные, примеси наследова-

¹⁴DOM (англ. Document Object Model — объектная модель документа) — программный интерфейс, позволяющий программам и скриптам получить доступ к содержимому HTML, XHTML и XML

ние селекторов и многое другое. Позволяет создавать кроссбраузерные правила CSS3. Сайт проекта: sass-lang.com.

Zurb Foundation Framework — свободная коллекция инструментов для создания веб-приложений с адаптивной версткой. Имеет большую часть необходимых стилей и элементов для реализации полноценного пользовательского интерфейса. Функционал аналогичен функциональности фреймворка Bootstrap от создателей Twitter. Сайт проекта: foundation.zurb.com.

Bootstrap Framework — свободно распространяемый веб фреймворк для создания сайтов и веб-приложений. Включает в себя HTML и CSS шаблоны оформления для типографики, веб-форм, кнопок, меток, блоков навигации и прочих компонентов веб-интерфейсов, включая JavaScript расширения. Сайт проекта: twitter.github.io/bootstrap.

jQuery UI — библиотека JavaScript с открытым исходным кодом для создания насыщенного пользовательского интерфейса в веб-приложениях, часть проекта jQuery. Построена поверх главной библиотеки jQuery и предоставляет разработчику упрощенный доступ к её функциям взаимодействия, анимации и эффектов, а также набор виджетов. Недостатком данной библиотеки является перегруженная разметка. Сайт проекта: jquery.com.

ExtJS — библиотека JavaScript для разработки веб-приложений и пользовательских интерфейсов, изначально задуманная как расширенная версия Yahoo! UI Library, однако преобразовавшаяся затем в отдельный фреймворк. Поддерживает технологию AJAX, анимацию, работу с DOM, реализацию таблиц, вкладок и обработку событий. Такой набор позволяет реализовать все необходимые элементы пользовательского интерфейса. Сайт проекта: www.sencha.com.

Backbone.js — MVC фреймворк для создания веб-приложений. В основе данного фреймворка лежат модели, коллекции, представления и роутеры. Фреймворк позволяет связать веб-приложение с существующим

REST API. Позволяет писать хорошо структурированный javascript код.
Сайт проекта: backbonejs.org.

В данном разделе были рассмотрен инструментарий, позволяющий создать полноценное веб-приложение с функциональностью ГИС. Выбор программного обеспечения, которое используется для создания прототипа системы мониторинга Арктики, приведен в разделе [2.8 Система мониторинга](#).

Глава 2

Экспериментальная часть

2.1 Выбор спутниковых данных

В разделе 1.2 приведено краткое описание продуктов, на основе которых можно производить прогноз. Еще раз рассмотрим наиболее подходящие из них и укажем особенности, которые являются важными при прогнозировании.

2.1 Продукт **Special Sensor Microwave/Imager**

В SSM/I доступна только маска льда, то есть содержится информация о наличии льда в конкретном пикселе. SSM/I является гридированным по градусам продуктом в картографической равноугольной цилиндрической проекции Меркатора, с разрешающей способностью в 0.25 градусов. В связи с тем что проекция Меркатора является проекция с искажением длин и площадей, получается что ближе к северному полюсу длина одного пикселя как по вертикали так и по горизонтали может принимать значения от 1 до 10 км. Такой разбор плотности дает различную корреляцию для пикселей находящихся на различных параллелях.

Данные SSM/I доступны начиная с 1979 года по настоящее время, что дает возможность учесть долгосрочный тренд и два значимых минимума в 2007 и 2012 годах. При изучении конфигурации льда в Арктическом регионе можно отметить интересную особенность, этой особенностью является наличие симметрии границ льда относительно северного полю-

са. Это дает возможность прогнозировать удаленность границы льда по меридианам с некоторыми ограничениями.

Преимуществом такого прогноза является возможность применения модели Vector Auto Regression (VAR), с небольшой предварительной подготовкой данных.

Однако, на Рис. 2.1 можно заметить сложность с которой придется столкнуться, если будет выбран такой способ прогнозирования.

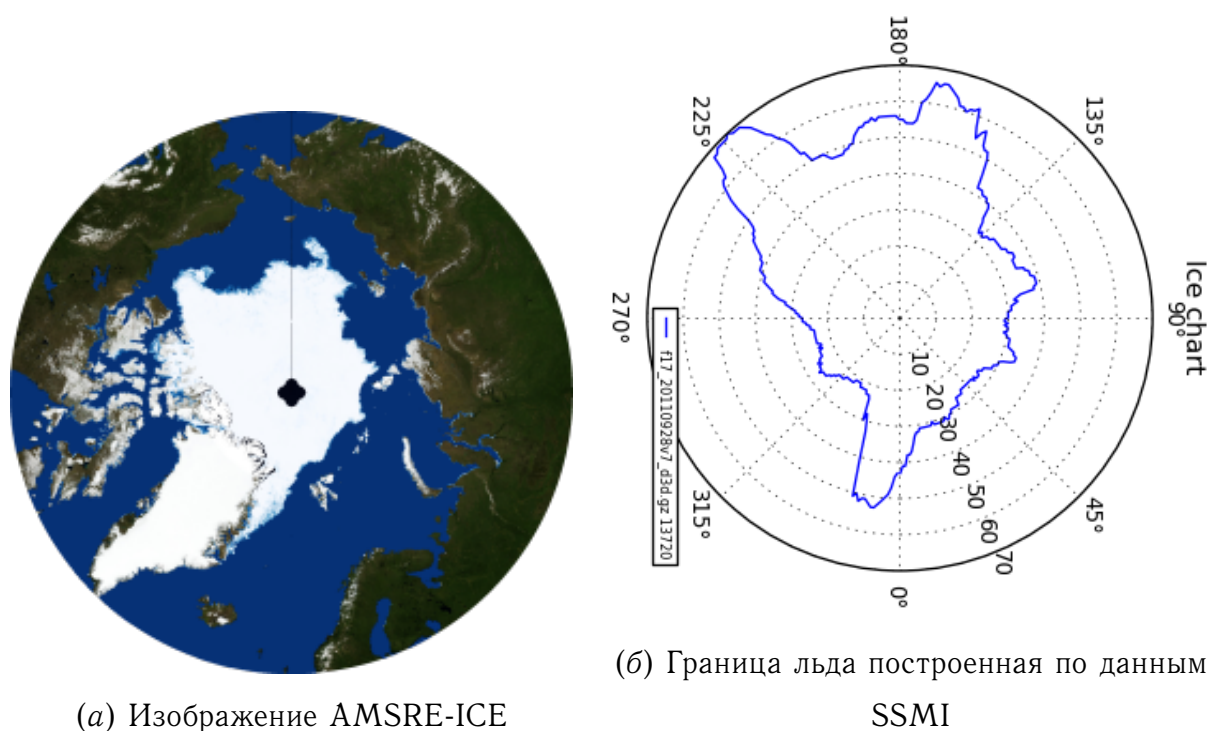


Рисунок 2.1 – Конфигурация и граница льда за 30-ое сентября 2011 года, построенные по продуктам ASI-AMSRE и SSMI соответственно.

Сравнивая Рис. 2.1, а и Рис. 2.1, б можно определить такие проблемные области как восточный берег Гренландии, Северо-Восточная Земля, Шпицберген, Новая Земля и ряд других областей.

Например, рассмотрим восточный берег Гренландии, в этой области граница льда постоянно колеблется вдоль береговой линии. Следовательно, если значение удаленности крайней точки границы льда на меридиане будет резко и не предсказуемо изменяться на большое значение.

Можно ввести ограничение на площадь прогнозирования, но тогда из виду будет упущена большая часть информации. Таким образом, было

принято решение производить прогноз конфигурации льда по данным продуктов, содержащих концентрацию льда.

В результате проведения данного исследования были получены некоторые полезные результаты. Введена оценка *ошибки* для снимка, описанная в формуле 2.1. Такая простая оценка позволяет понять актуальность данного снимка для его дальнейшего анализа.

$$\text{ошибка} = \frac{\text{Количество пикселей со значением 254 (unknown)}}{\text{Количество пикселей в области интереса}} \quad (2.1)$$

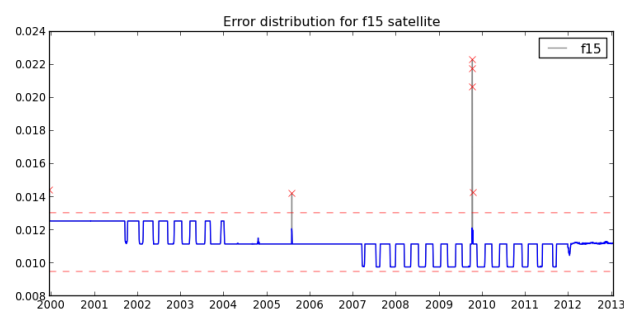
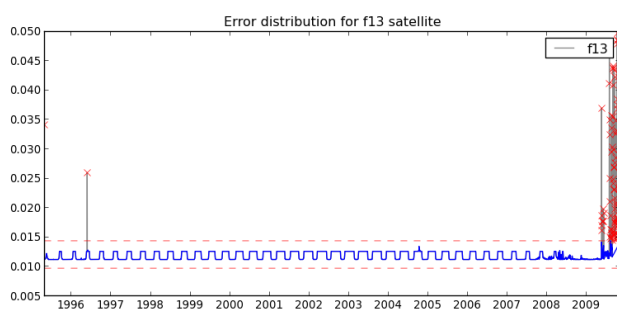
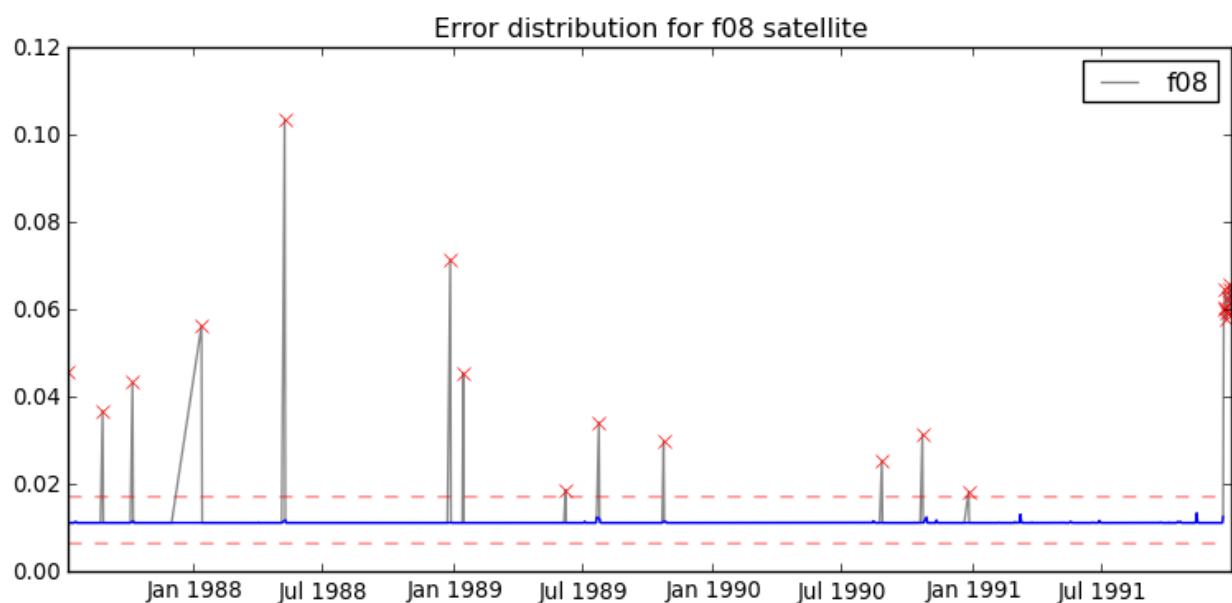
На Рис. 2.2 изображены временные ряды значения ошибок, которые были подсчитаны для снимков, полученных со спутников f08, f15 и f15 в Арктическом регионе. Красными пунктирными линиями обозначен доверительный интервал для ошибок снимков. Если значение ошибки находится в интервале, то снимок может быть использован для построения прогноза. Границы интервала равны: $M[X] \pm D[X]$, где $M[X]$ - математическое ожидание, а $D[X]$ - корень из среднеквадратичного отклонения.

Такая простая оценка дает возможность увидеть, например, на Рис. 2.2, а как приходит в негодность спутник f13. На Рис. 2.2, б представлен временной ряд значений ошибки для спутника f15. Можно отметить интересную особенность: в 2006 году значение ошибки уменьшилось, в связи с этим, доверительный интервал увеличен вдвое, связано такое изменение с тем, что данные были подкорректированы.

2.1 Продукты National Snow & Ice Data Center

В National Snow & Ice Data Center (NSIDC) существуют два продукта при объединении которых, можно получить максимальное время наблюдения [24, 29]. Основные характеристики продуктов NSIDC были рассмотрены в Главе 1.2. Далее приведено более детальное рассмотрение этих продуктов.

NSIDC-SSMI-SEAICE-GSFC — Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data.



(а) Ошибки для спутника f13

(б) Ошибки для спутника f15

Рисунок 2.2 – Временной ряд значения ошибок, которые были подсчитаны для данных, полученных со спутников f08, f13 и f15 в Арктическом регионе.

Этот набор данных имеет временные границы с 26 октября 1978 по 31 декабря 2011 года, он был получен с нескольких сенсоров, которые находились на разных спутниках: *a*) сенсор Scanning Multichannel Microwave Radiometer (SSMR), находившийся на спутнике Nimbus-7; *b*) сенсор Special Sensor Microwave/Imagers (SSM/Is) был расположен на спутниках f8, f11 и f13, запущенных в рамках программы DMSP (англ. Defense Meteorological Satellite Program, перевод Оборонная метеорологическая спутниковая программа); *c*) сенсор Special Sensor Microwave Imager/Sounder (SSMIS), находящийся и продолжающий работать на спутнике DMSP-f17.

Данные получены с помощью алгоритма NASA, разработанного в Oceans and Ice Branch, Laboratory for Hydrospheric Processes, находящейся в NASA Goddard Space Flight Center (GSFC).

Данные включают в себя дневную, а так же осредненную месячную концентрации льда для северного и южного полярных регионов.

NSIDC-SSMI-NRT-SEAICE — Near-Real-Time DMSP SSM/I-SSMIS Daily Polar Gridded Sea Ice Concentrations.

Данные получены с сенсора Special Sensor Microwave Imager/Sounder (SSMIS), находящегося на спутнике DMSP-f17. Отличия от данных первого продукта, заключаются в отсутствии месячного осреднения и улучшенном алгоритме обработки.

Данные доступны для скачивания с FTP сервера NSIDC. Так же доступ можно получить через OPeNDAP сервер Лаборатории Спутниковой Океанографии (opendap.solab.rshu.ru).

Оба продукта имеют пространственное разрешение 25 км. Концентрации льда представлена матрицей 304 x 448, имеющей значения:

0-250 — концентрация льда;

251 — маска, отображающая недоступную область, связанную с орбитой спутника;

252 — неиспользованный бит;

253 — береговая линия;

254 — наложенная маска земли;

255 — пропущенные данные.

Для анализа можно использовать только центральную часть матрицы размером 280 x 280, так как областью интереса является акватория Арктического региона. На Рис. 2.3 приведено изображение концентрации льда для Арктического региона с указанием значения в точке. Данное изображение создано с помощью написанной, при выполнении магистерской диссертации, программы.

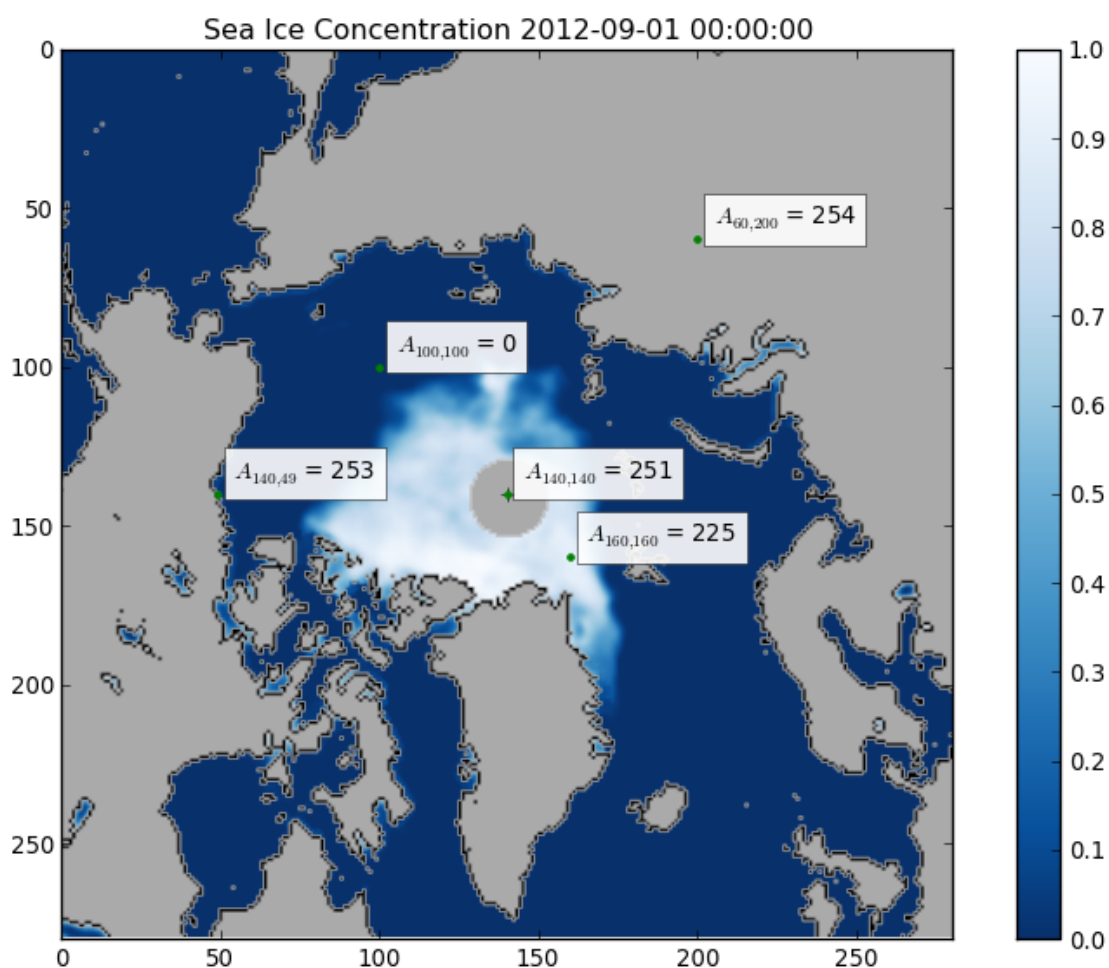


Рисунок 2.3 – Карта ледового покрытия, отображающая концентрацию льда в сентябре 2012 года, с указанием значений в точках.

2.2 Модель данных

Для анализа и прогноза ледовой конфигурации был написан небольшой фреймворк, позволяющий работать и с другими спутниковыми продуктами. В данном разделе описывается принятая в этом фреймворке модель данных. *Модель данных* — это совокупность структур данных и операций их обработки.

В основе модели лежат два понятия из области методов дистанционного зондирования, рассмотренные в разделе 1.2 : продукт и гранула. Понятие продукта преобразовалось в класс *GranulesCollection* (Коллекция гранул).

Гранулы. В фреймворке реализованы четыре класса гранул:

1. *GranuleAbstract* — абстрактный базовый класс, содержащий базовые методы и атрибуты, которые являются общими для всех классов гранул. Атрибутами для гранул являются: время создания, название продукта, имя файла, регион, спутник и временной интервал.
2. *Granule* — гранулы, связанные с файлами. Параметры в таких гранулах загружаются посредством OPeNDAP сервера, и после этого кешируются. Основным методом является метод получения данных с OPeNDAP сервера, позволяющий загрузить матрицу значений параметра.
3. *ForecastGranule* — гранулы, для прогнозирования. Класс предназначен для описания гранул полученных путем осреднения нескольких экземпляров класса *Granule*.
4. *PredictedGranule* — класс для описания гранул полученных в результате прогнозирования.

Так же для классов реализованы примеси¹ (mix in) для построения изображений и выделения масок земли, береговой линии и т.д. Например, на Рис. 2.4 приведено изображение гранулы за январь 1988 года с масками земли, береговой линии и неизвестной области.

Коллекции. Коллекцией является набор гранул. Класс *GranulesCollection* является оберткой над списком (list), позволяющей производить специфические для гранул операции:

- Запрос по атрибутам. Возвращает выборку гранул, удовлетворяющих параметрам запроса. Например, можно выбрать все гранулы за некоторый год или месяц.
- Подсчет площадей для всех гранул коллекции.

¹Примесь (англ. mix in) — элемент языка программирования (обычно класс или модуль), реализующий какое-либо чётко выделенное поведение. Используется для уточнения поведения других классов, не предназначен для порождения самостоятельно используемых объектов.

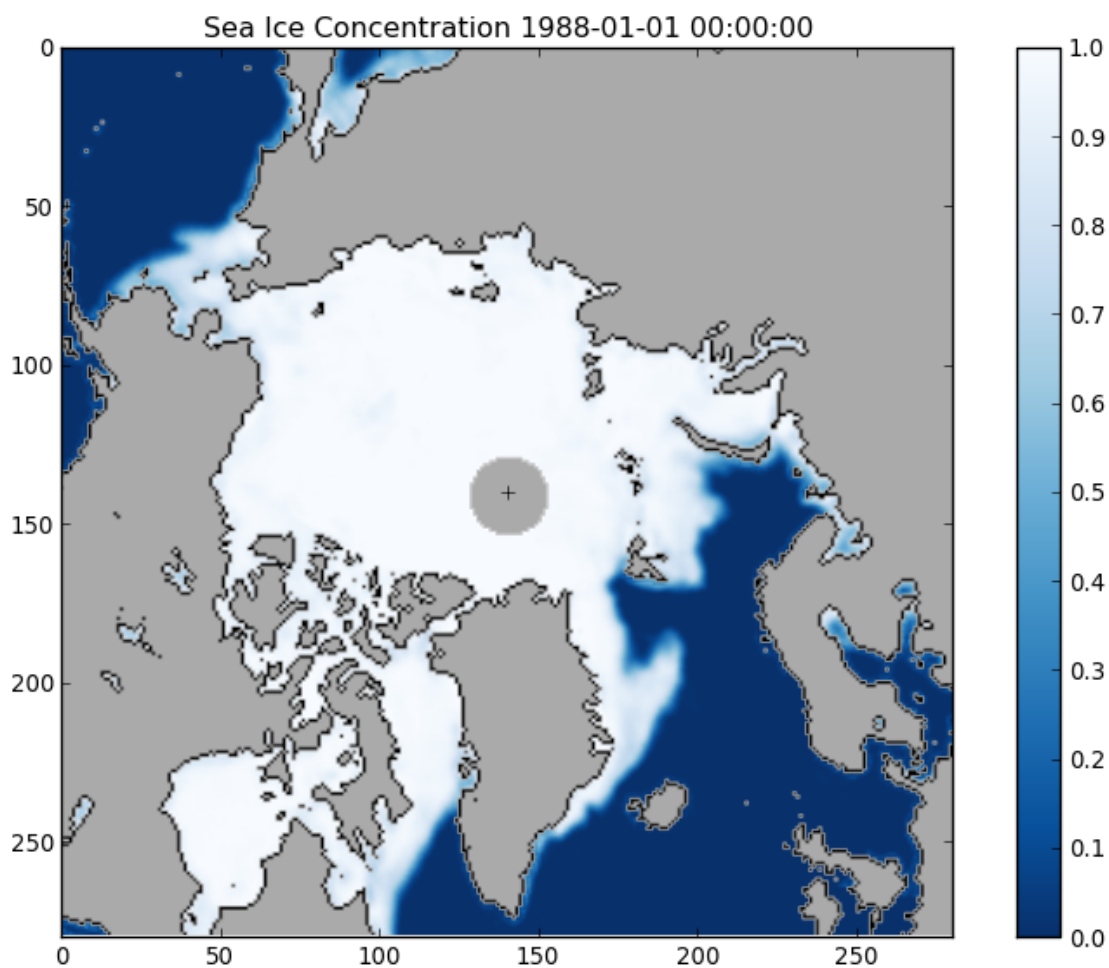


Рисунок 2.4 – Результат метода построения изображения.
Концентрация льда за сентябрь 2007 года

- Разбиение коллекции по годам, месяцам или неделям. Данный метод полезен когда необходимо рассматривать только конкретно заданный промежуток времени.
- Осреднение данных коллекции по заданному промежутку времени.

Так же реализованы аналитические методы, которые будут рассмотрены в разделе [2.4 Анализ данных](#).

- Подсчет площадей для всех гранул коллекции.
- Подсчет средней концентрации льда.
- Подсчет пространственной корреляции между значениями параметра во времени в конкретном пикселе.

- Подсчет тренда выраженного полиномом для значений концентрации льда во времени.

Так же как и для классов гранул реализованы различные примеси, позволяющие

- строить изображения для:
 - пространственных параметров,
 - временного ряда подсчитанных площадей,
 - временных рядов концентрации льда в точке всех гранул коллекции,
 - корреляционной карты;
- создавать видео;
- обеспечивать интерфейс доступа к гранул, аналогичный в списках.

У каждого экземпляра всех классов гранул есть атрибут **data**, в этот атрибут загружается матрица параметра, в данном случае концентрации льда. Значения элементов матрицы описаны в разделе 2.1.2. Если объединить матрицы всех гранул коллекции, то получим трехмерный массив, имеющий размерность $n \times 280 \times 280$, где n количество гранул в коллекции. Точечный срез, полученный путем выбора из каждой матрицы по одному значению стоящему на одинаковых местах, дает временной ряд значений параметра, прогноз которого и будет производиться.

2.3 Подготовка данных

Помимо спутниковых данных в архивах хранятся различные файлы, например, превью изображения. Стоит отметить что продукты имеют разную структуру. Однако, структуру архива можно не рассматривать т.к. регистрацию файлов можно производить при помощи рекурсивного обхода и поиска файлов, имена которых подходят под некоторое регулярное выражение:

```
GRANULE_NAME_RE = "nt_(?P<year>\d{4})(?P<month>\d{2})" \
    "(?P<day>\d{2})?_(?P<satellite>\w\d{2})" \
    "_(?P<version>(\w+))?(?P<region>\w+).bin"
```

Выше приведено регулярное выражение с именованными группами, позволяющее выделить:

- дату создания снимка;
- спутник, с борта которого производилась съемка;
- версию алгоритма обработки;
- регион: северное полушарие или южное.

При обходе создается коллекция гранул, в которой находятся объекты класса *Granule*. Гранулы в этой коллекции принадлежат двум продуктам NSIDC, рассмотренным в разделе [2.5 Выбор модели и метода прогнозирования](#). Для первого продукта уже есть осреднение по месяцам, так же в нем есть небольшая особенность: в 2007 году присутствуют данные с двух спутников f13 и f17, их необходимо осреднить. Для второго есть только дневные гранулы, следовательно для них необходимо произвести осреднение.

После осреднения гранул по месяцам, их необходимо объединить в одну коллекцию, и произвести загрузку параметра, то есть скачать матрицу концентрации льда с OPeNDAP сервера. Матрица параметра хранится в *ленивом* атрибуте *data*. Ленивым этот атрибут считается потому,

что его загрузка происходит в момент первого обращения. Такую особенность можно реализовать с помощью перегрузки метода `__getattr__`, который будет проверять загружены ли данные или нет и производить необходимые действия.

2.4 Анализ данных

Для того чтобы прогнозировать данные, необходимо понять особенности их поведения. В данном случае необходимо понять как ведет себя ледовый покров, когда достигаются экстремумы площади, как связаны концентрации льда в удаленных друг от друга точках и т.п.

Для удобства и наглядности создан видеоряд, позволяющий рассмотреть динамику конфигурации льда и изменения его концентрации. Видео размещено по ссылке: <http://bit.ly/Arctic-ice-concentration>.

2.4 Экстент ледяного покрова

Наиболее часто для оценки ледовой обстановки используется понятие экстента. Для каждой гранулы экстент (англ. Extent, что переводится как мера, размер, протяженность или пространство) рассчитывается по формуле: $extent = n * 625$, где n это количество пикселей, концентрация льда в которых больше 15% [31].

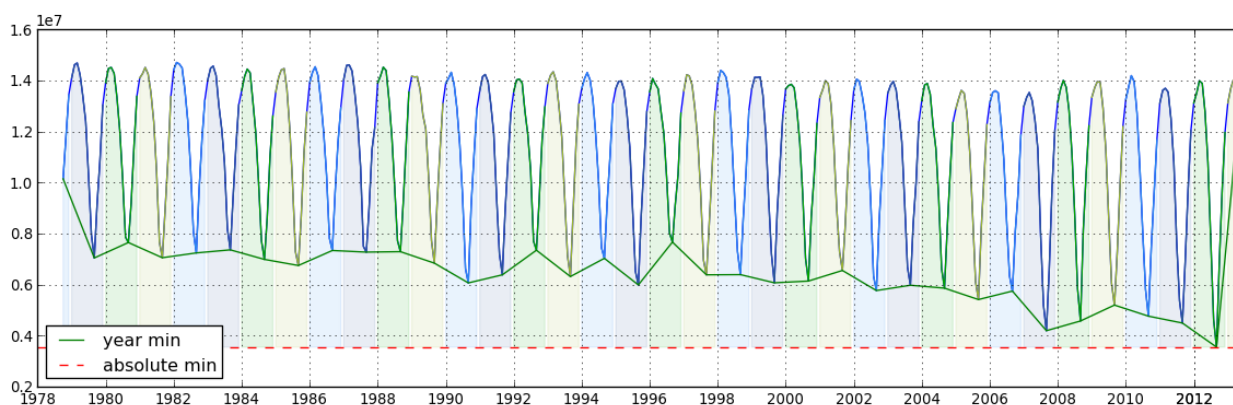


Рисунок 2.5 – Экстент ледового покрова с октября 1979 по июнь 2013

На Рис. 2.5 приведены значение экстента с октября 1979 по июнь 2013 года. Можно видеть, что идет уменьшение минимального значения экстента за год. То есть идет уменьшение площади ледового покрова. Минимумы обычно достигаются в сентябре. Во введении было указано, что по данным National Snow and Ice Data Center в сентябре 2012 года

был зарегистрирован самый низкий показатель с 1979 года, площадь оказалась вдвое меньше средней за 1979 и 2000 годы [23].

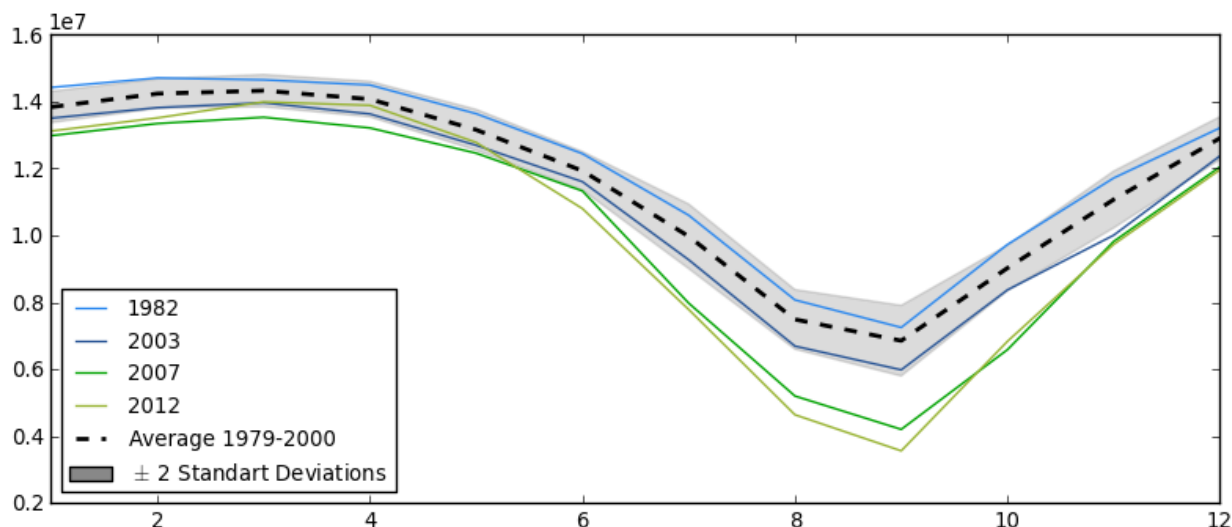


Рисунок 2.6 – Экстент ледового покрова по месяцам

На Рис. 2.6 сплошной линией изображены значения экстената по месяцам для некоторых годов, пунктирной линией обозначено усредненное значение, а серой областью обозначено удвоенное стандартное отклонение. По данному графику можно увидеть, что минимумы обычно достигаются в сентябре. Так же можно видеть, что колебания значения экстената растут.

Минимальный и максимальный экстенаты

На Рис. 2.7 приведены минимальный и максимальный экстенаты за весь доступный период. В написанном фреймворке, гранулу с минимальным подсчитанным экстенатом можно найти при помощи запроса:

```
max_granule = max(monthly_granules , key=lambda g: g.square)
min_granule = min(monthly_granules , key=lambda g: g.square)
```

Тенденции

Тренды позволяют увидеть общую тенденцию развития процессов. На основе трендов возможно построение прогнозов.

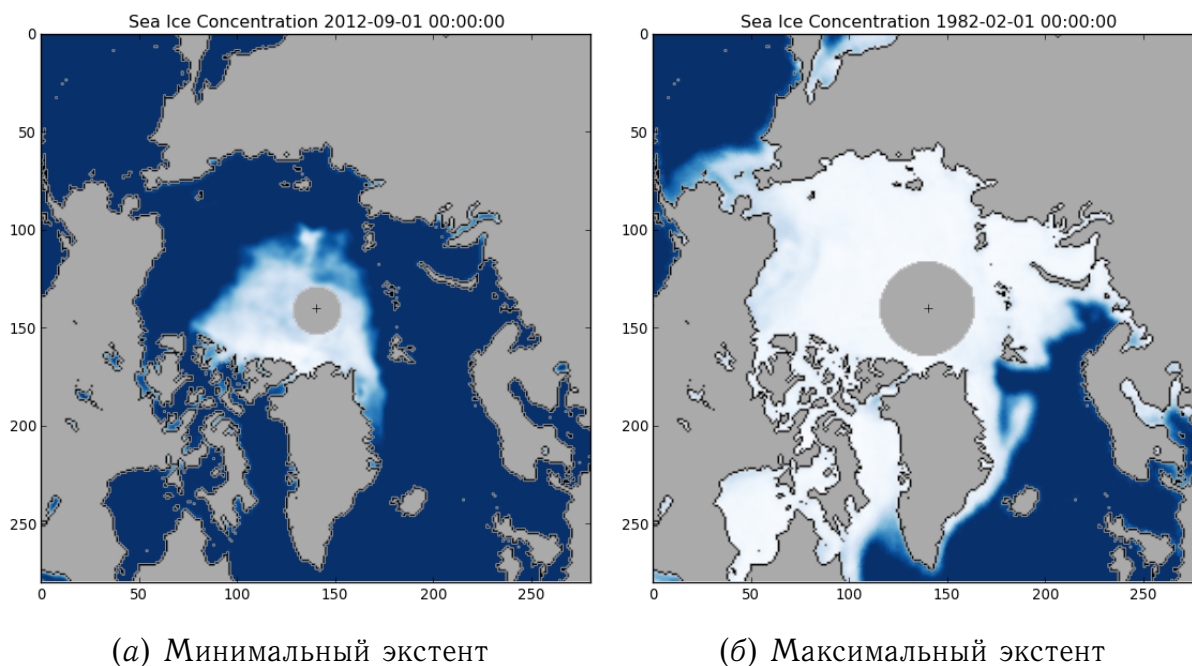


Рисунок 2.7 – Максимальный и минимальный экстенсты за период с 1979 по 2013 годы

На приведенном Рис. 2.8 видно, что происходит уменьшение экстенста. Так же видно, что линейный и полиномиальный тренды почти совпадают. Для лучшего понимания можно рассмотреть тенденции для месячных данных.

На основе графиков на Рис. 2.9 можно сделать предположение, что экстенст продолжит активно уменьшаться в июле, сентябре и январе. По тренду для марта можно сделать предположение, что площадь будет постоянно увеличиваться и при экстраполяции на длительный период будет стремиться к бесконечности. Эти моменты стоит учесть при построении пространственного прогноза.

2.4 Концентрация льда

Следующим этапом изучения ледовой обстановки является рассмотрение концентрации льда в различных точках. Изображение параметра концентрации льда имеет размеры 280 x 280 пикселей, что дает 78400 пикселей. Примерно половина из них закрыта маской земли, в итоге

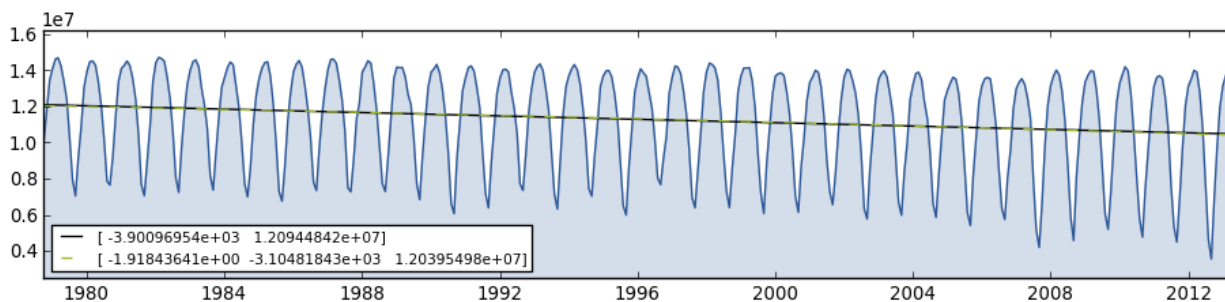


Рисунок 2.8 – Тренд экстенда за весь период. Черной прерывистой линией обозначен линейный тренд. Зеленой прерывистой линией обозначен тренд, описанный полиномом второго порядка.

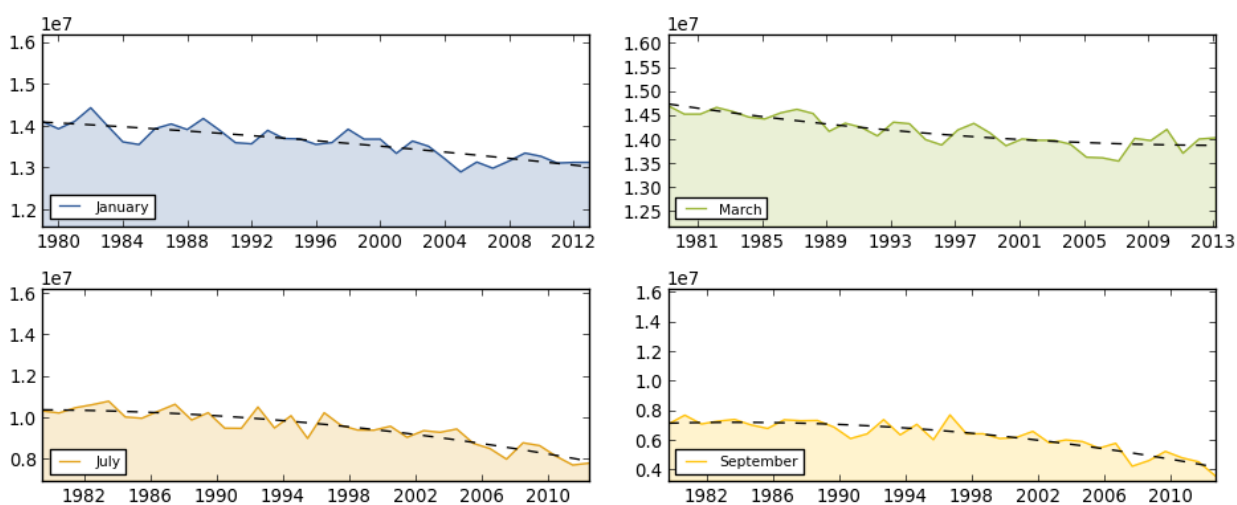


Рисунок 2.9 – Нелинейный тренд для данных, разделенных по месяцам.

остается 34929 элементов. То есть необходимо составить прогноз для 35 тысяч временных рядов. Рассмотрим некоторые концентрации льда для точек приведенных на Рис. 2.10.

На Рис. 2.12, *a* приведен временной ряд концентрации льда для точки (214, 206). В этой точке вода лишь иногда немного покрывается льдом. Это связано с тем, что здесь проходит *Северо-Атлантическое течение*, оно является мощным теплым океаническим течением, которое отделяется от Гольфстрима и несет теплые воды на север.

Справа на Рис. 2.12 отображена корреляционная карта для конкретного временного ряда со всеми остальными. Способ построения карты рассматривается дальше. На данный момент, стоит отметить что при

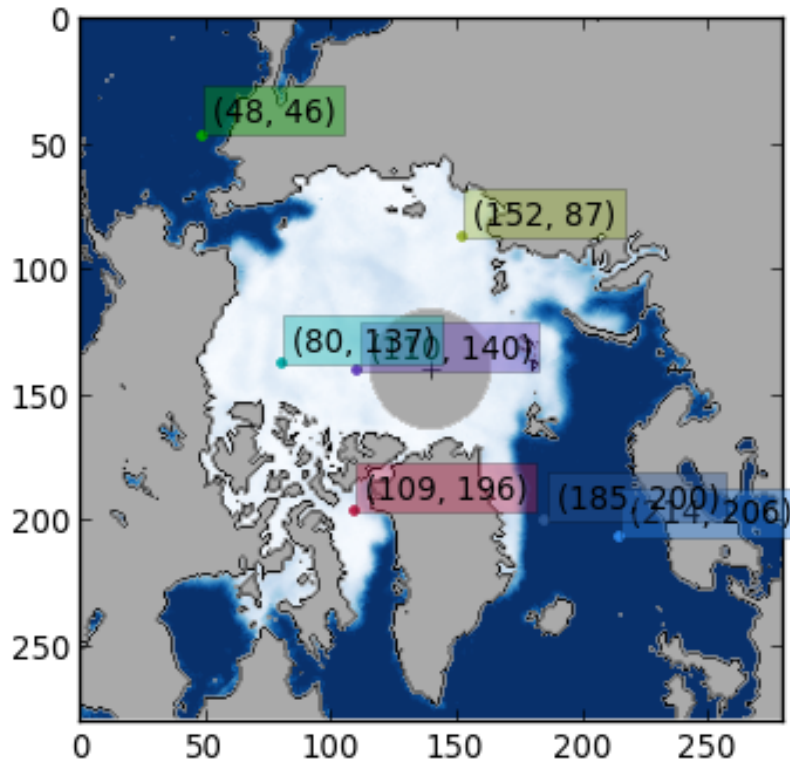


Рисунок 2.10 – Точки, для которых будет рассматривается концентрация льда.

значение корреляции больше 0.7, процессы принято считать сильно взаимосвязанными.

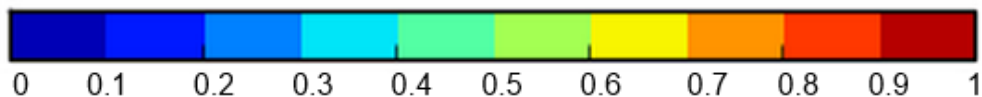
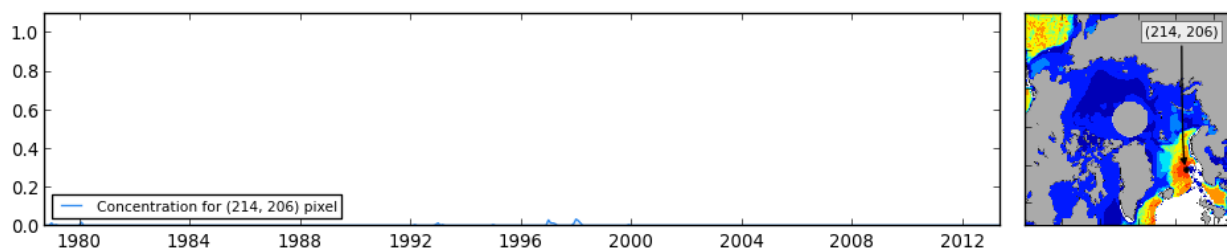


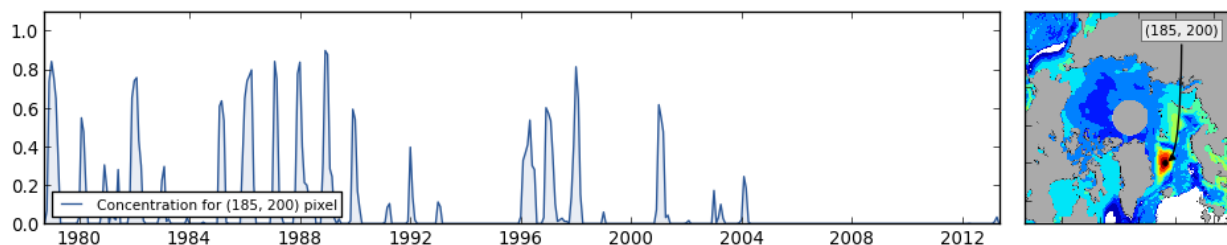
Рисунок 2.11 – Палитра цветов для корреляционной карты. Каждый цвет соответствует интервалу, обозначенному числами.

На Рис. 2.12, б и Рис. 2.13, а изображены концентрации льда для участков преимущественно свободных ото льда. Однако, можно видеть, что иногда граница льда доходит и до них и тогда концентрация льда в точке увеличивается.

Два следующих графика концентрации льда, приведенных на Рис. 2.13, б и Рис. 2.13, в схожи тем, что в сентябре значение концентрации льда значительно уменьшается. По карте корреляции можно



(а) Концентрация льда для точки (214, 206)

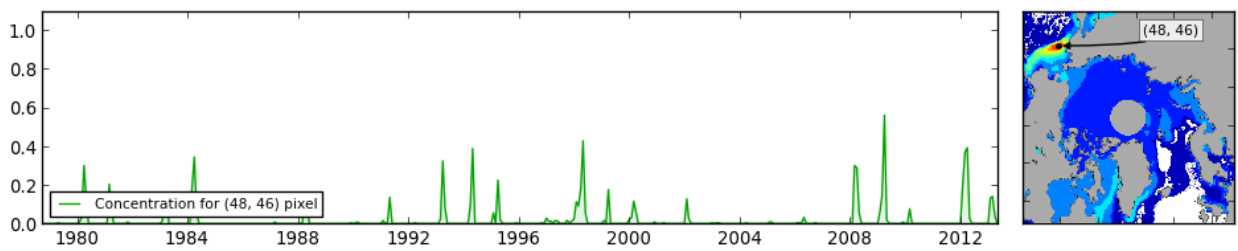


(б) Концентрация льда для точки (185, 200)

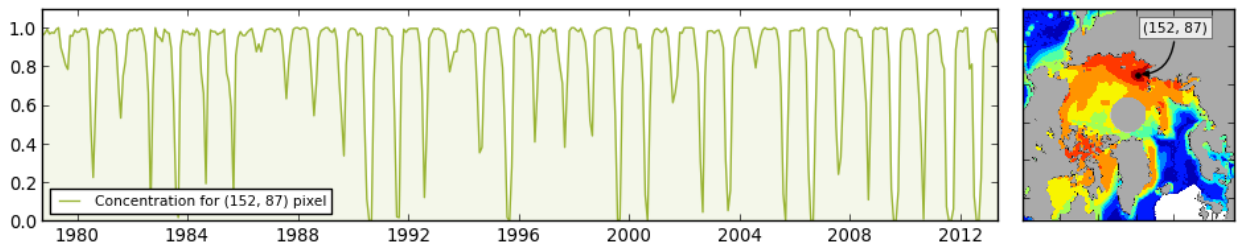
Рисунок 2.12 – Концентрации льда и корреляционная карта для выбранных точек. Часть 1

видеть, что таких рядов достаточно много. Схожие ряды обозначены кирпичным, оранжевым и красным цветами. Корреляция между данными точками лежит в интервале от 0.7 до 0.8, что является высоким показателем.

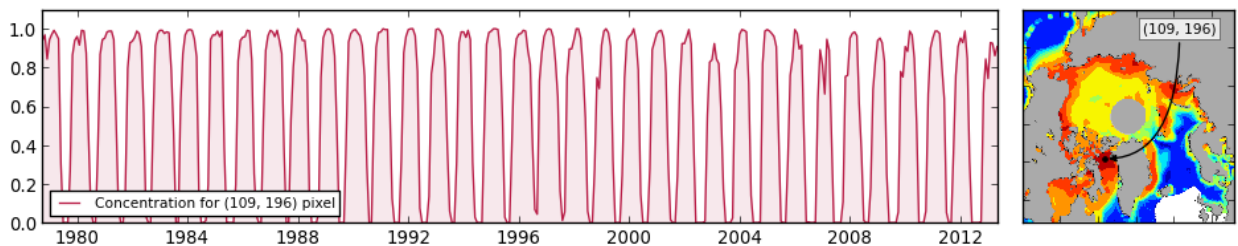
На Рис. 2.13, *з* и Рис. 2.13, *д* так же приведены два схожих графика концентрации льда. Эти области находятся достаточно близко друг к другу и концентрации льда ведут себя схоже. Но область на Рис. 2.13, *д* расположена дальше от северного полюса и концентрация льда в ней падает значительно ниже. Можно видеть, как начиная с 2006 года минимальное значение концентрации льда начинает сильно убывать. В сентябре 2012, когда был зафиксирован минимум экстенста, эта точка в подсчет уже не попала т.к. концентрация в ней составила меньше 15%.



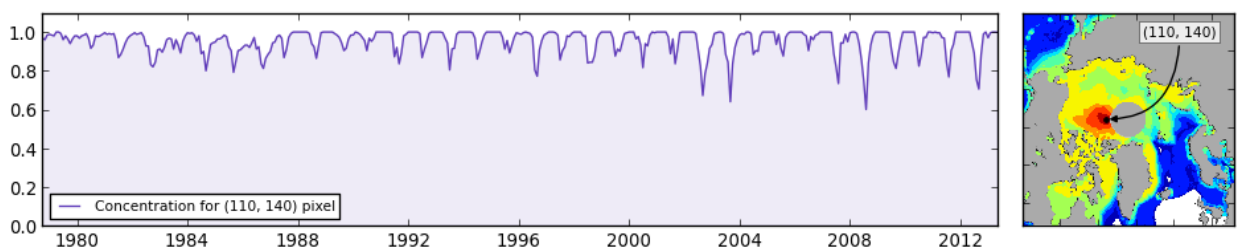
(а) Концентрация льда для точки (48, 46)



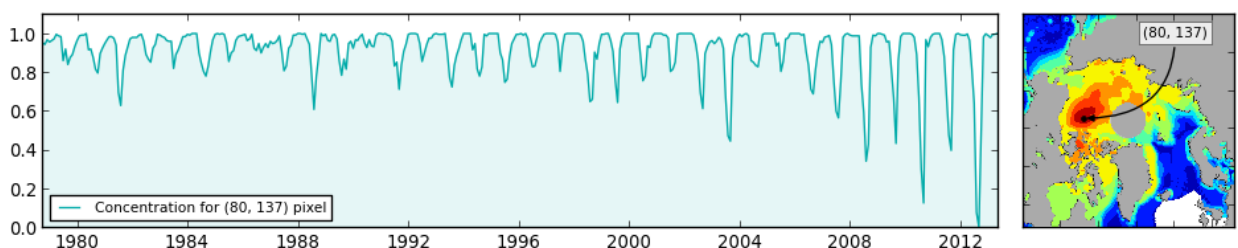
(б) Концентрация льда для точки (152, 87)



(в) Концентрация льда для точки (109, 196)



(г) Концентрация льда для точки (110, 140)



(д) Концентрация льда для точки (80, 137)

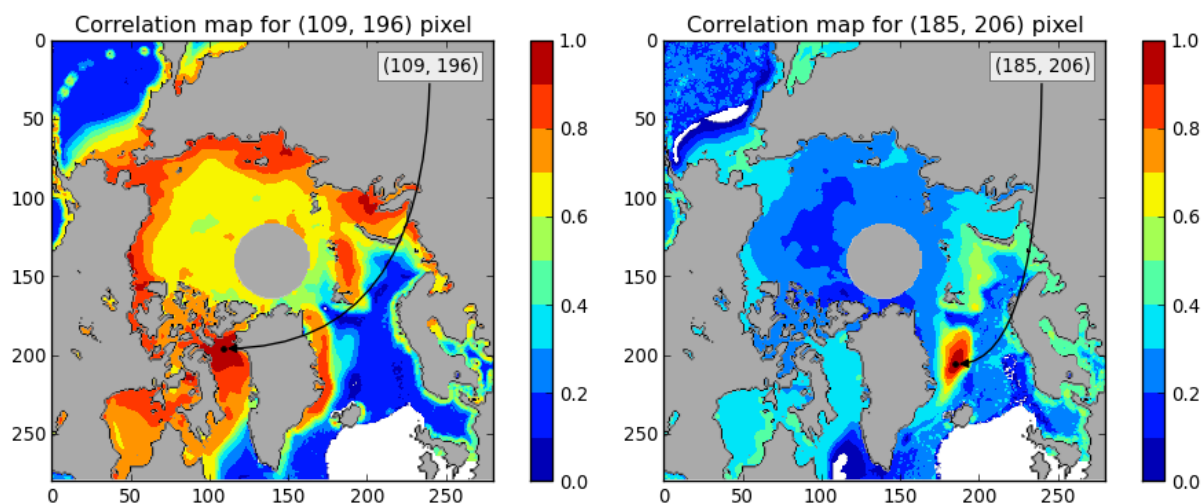
Рисунок 2.13 – Концентрации льда и корреляционная карта для выбранных точек. Часть 2

2.4 Пространственные характеристики

Карта коэффициентов корреляции

Карта коэффициентов корреляции для заданной точки отображает коэффициент корреляции Пирсона между рядом динамики концентрации льда в текущей точке с рядами динамики концентрации льда для всех остальных точек. Такая карта позволяет определить коррелирующие между собой области, используя, например, метод кластеризации многомерного шкалирования (multidimensional scaling, MDS)[6].

На Рис. 2.14 приведены карты корреляции для двух точек (109, 196) и (185, 206). Можно видеть что для ряда динамики первой точки есть много схожих других рядов динамик с высоким показателем корреляции.



(а) Корреляционная карта для точки (109, 196)

(б) Корреляционная карта для точки (185, 206)

Рисунок 2.14 – Максимальный и минимальный экстенды за период с 1979 по 2013 годы

Карта коэффициентов линейного тренда

Карта коэффициентов линейного тренда отображает для каждого временного ряда концентрации льда коэффициент наклона линейного трен-

да. Линейный тренд описывается уравнением прямой: $y = kt + b$, где k - коэффициент наклона, t - время, b - начальное значение.

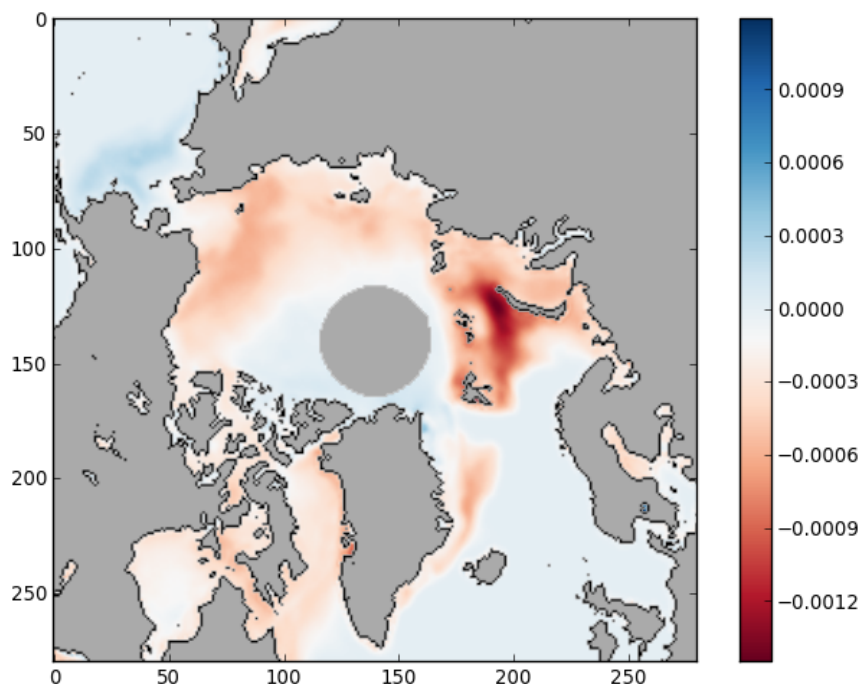


Рисунок 2.15 – Карта коэффициентов линейного тренда.

Красные области указывают на то, что концентрация льда в точках этих областей постоянно уменьшается. Если рассмотреть область северо-западнее архипелага Новая Земля - Баренцево море, и посмотреть видеоряд, описанный в разделе [2.4 Анализ данных](#), можно увидеть, что за последнее время эта область почти не замерзает.

Такие же карты можно построить для срезов данных по месяцам. Рассмотрим показатели за сентябрь и декабрь.

На Рис. [2.16](#) приведена карта коэффициентов наклона линейного тренда для сентября. В сентябре лед отступает в морях Бофорта и Чукотском. В декабре виден спад концентрации льда Баренцевом море, что совпадает с общим трендом.

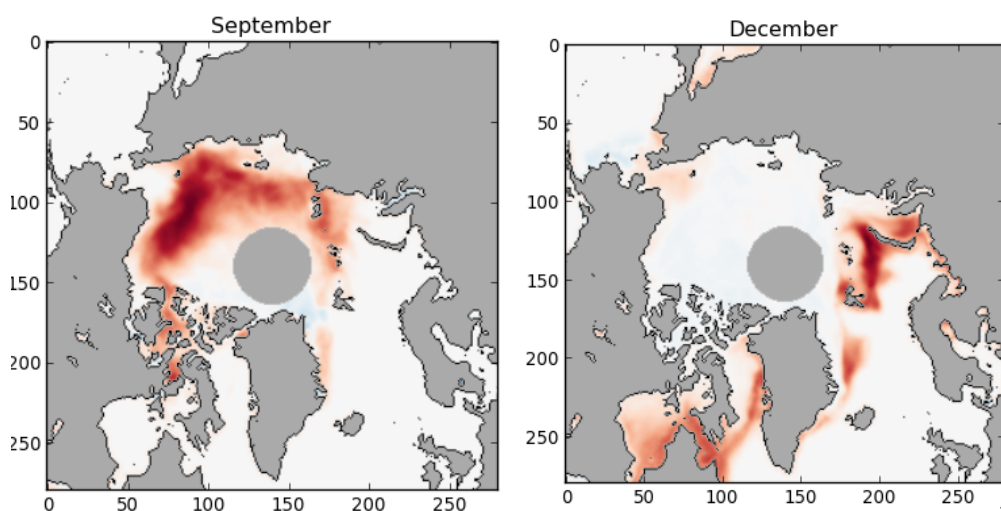


Рисунок 2.16 – Линейный тренд для данных за сентябрь и декабрь.

2.5 Выбор модели и метода прогнозирования

Задача прогнозирования конфигурации ледяного покрова сводится к прогнозированию концентрации льда во всех точках карты. В разделе [2.4 Анализ данных](#) было показано, что временные ряды концентрации льда в различных точках сильно отличаются друг от друга.

В связи с большим количеством временных рядов (34 929), для которых необходимо провести прогноз, использование авторегрессионной модели становится практически невозможным, т.к. данная модель включает большое число параметров, определение которых весьма затруднено. Проведя предварительную кластеризацию временных рядов на основе их корреляции друг с другом, можно обойти данное ограничение.

На Рис. [2.14](#) приведены корреляционные карты для двух точек. В качестве алгоритма кластеризации можно использовать метод кластеризации многомерного шкалирования (multidimensional scaling, MDS). Данный подход является будущим направлением работы.

В случае прогнозирования конфигурации льда, особый интерес вызывает месячная конфигурация. Ранее в разделе [2.4.1 Экстенд ледяного покрова](#) было показано, что временные ряды экстенда, построенные по отдельным месяцам, имеют специфические тенденции. Например, экстенд

льда в сентябре уменьшается намного быстрее, чем экстенд льда в декабре. В связи с этим принято решение строить независимый прогноз по месяцам и после этого объединять прогнозы в один.

Изначально предполагалось, что тенденция уменьшения площади ледового покрова имеет линейный характер. Однако при детальном рассмотрении были найдены механизмы ответной реакции. Например, по мере того как сокращаются снеговые и ледяные отражающие поверхности, более темные участки поглощают большее количество солнечной энергии, увеличивают потепление, что ускоряет процесс таяния льдов. В связи с этим было принято решение использовать для аппроксимации временных рядов концентрации льда полиномиальный тренд второго порядка.

Резюмируя вышесказанное, для прогнозирования конфигурации льда принято решение использовать следующий подход:

1. Исходные данные разделяются по месяцам на 12 коллекций;
2. В каждой коллекции оказывается от 33 до 35 снимков, описывающих концентрацию льда в точках;
3. Для каждого временного ряда концентрации льда в каждой коллекции при помощи метода наименьших квадратов строится аппроксимирующая модель на основе тренда второго порядка;
4. На основе полученных трендов концентрации льда в точках, восстанавливается общая динамика конфигурации льда;
5. Оценка адекватности полученной модели, производится на анализе остатков между экстендами, подсчитанным для изначальных и аппроксимированных данных, разделенных по месяцам.

Разделение данных по месяцам

В рамках полученного фреймворка разделение данных осуществляется с помощью метода коллекции *separate*:

```
monthly_granules.separate()
```

Данный метод возвращает список из двенадцати коллекций с гранулами, с одинаковыми месяцами.

Построение аппроксимирующей модели

Для построения аппроксимирующей модели для каждого временного ряда концентрации льда в точке в каждой коллекции при помощи МНК был найден полином второго порядка. Для наглядности результатов был написан метод, отображающий значения коэффициентов на карте.

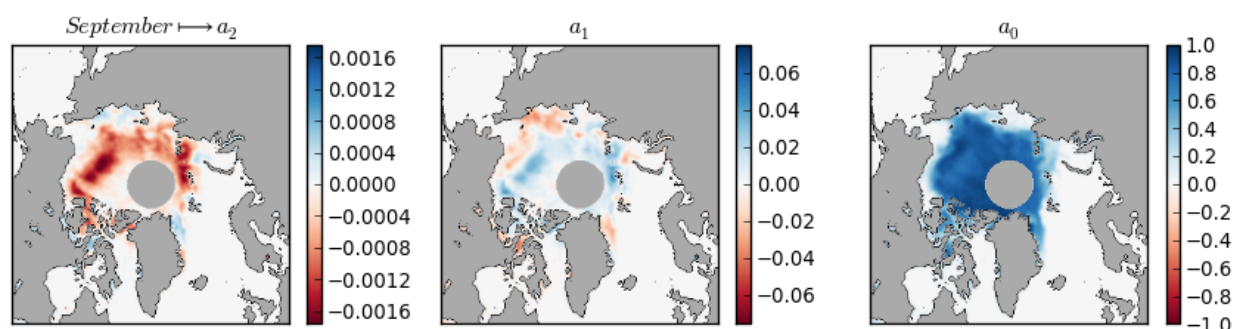


Рисунок 2.17 – Карта коэффициентов тренда второго порядка.

Результаты данного метода приведены на Рис. 2.17. На картах отображены значения коэффициентов уравнения полинома второго порядка: $y(t) = \alpha_2 t^2 + \alpha_1 t + \alpha_0$.

При объединении данных и подсчете для них экстента, получаем временной ряд приведённый на Рис. 2.18. Синей линией обозначен временной ряд экстента исходных данных, зеленой линией - временной ряд экстента для полученной модели, светло зеленой областью - временной ряд остатков.

Оценка адекватности модели

Большое количество временных рядов делает крайне затруднительной оценку адекватности каждой модели, аппроксимирующей временной ряд концентрации льда в точке. Для оценки адекватности было предложено три приближения:

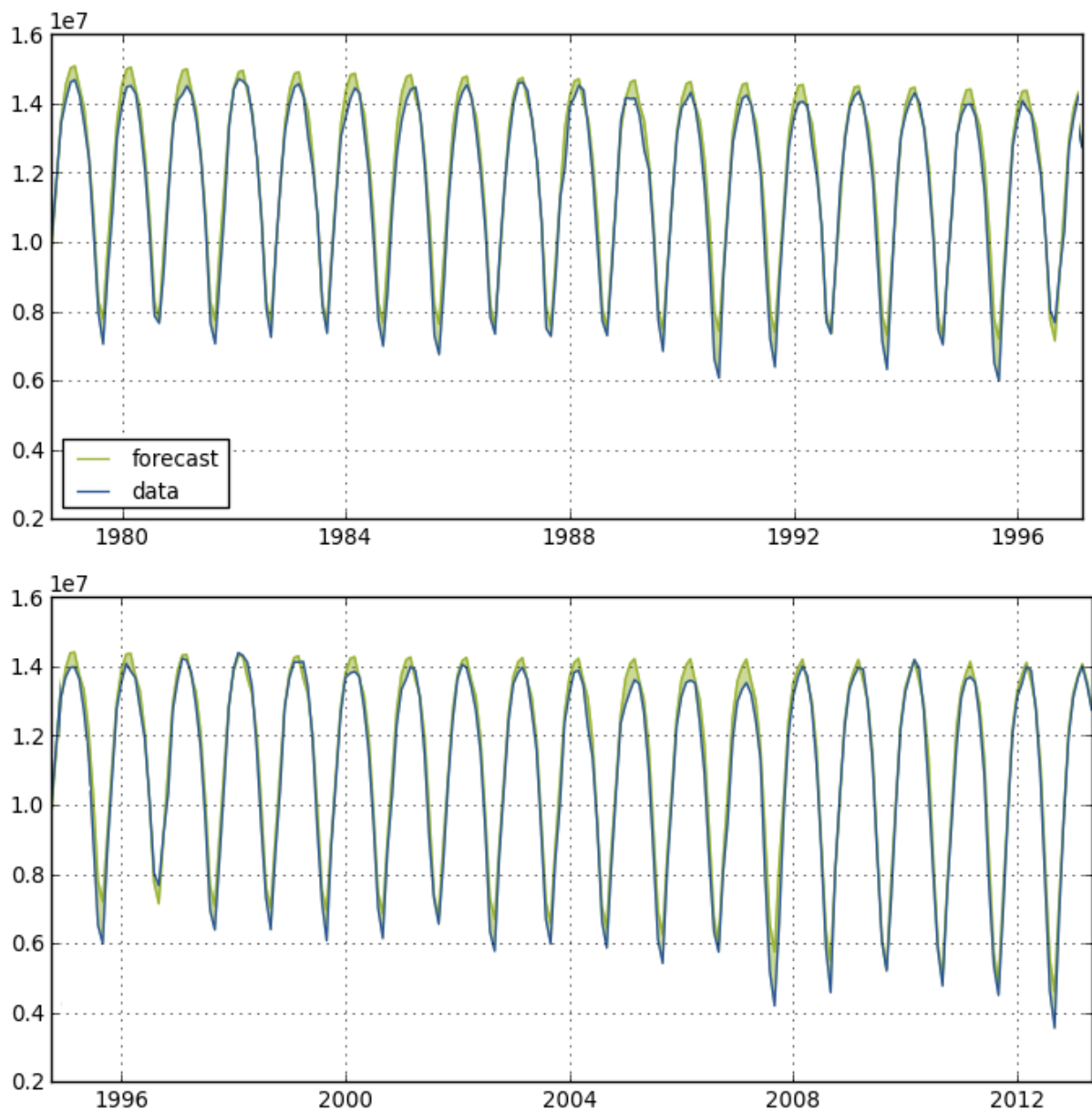


Рисунок 2.18 – Временные ряды значений экстремов, подсчитанных для исходных данных и данных модели.

1. Экстент, рассчитанный для исходных данных и полученной аппроксимирующей модели ведут себя одинаково в интервале, на котором данные доступны. Значения экстента аппроксимирующей модели сохраняют сезонность и тренд. Для оценки используется графический метод и оценка эксперта в области океанологии (в лучшем случае гляциологии).
2. Для оценки адекватности модели проводится оценка временных рядов значений экстента, разделенных по месяцам. Таким образом, будет необходимо проанализировать остатки для двенадцати временных рядов и проверить выполнение пяти предпосылок МНК.
3. Для оценки адекватности модели необходимо провести анализ рядов остатков для всех полученных моделей временных рядов концентрации льда. То есть необходимо проверить предпосылки МНК для 35 000 временных рядов.

Случайный характер ряда остатков. Выдвигаем гипотезу H_0 : ряд остатков случаен. Для проверки гипотезы применим критерий поворотных точек.

Критерий поворотных точек. Поворотной точкой называется точка p_t , удовлетворяющая одному из условий:

$$\begin{cases} p_t > \max(p_{t-1}, p_{t+1}) \\ p_t < \min(p_{t-1}, p_{t+1}) \end{cases} \quad (2.2)$$

Рассчитывается количество поворотных точек и, затем, проверяется условие $p \geq p_{cr}$, где

$$p_{cr} = \left[\frac{2}{3}(n - 2) - 1.96 \sqrt{\frac{16n - 29}{90}} \right] \quad (2.3)$$

Если условие нарушается, то гипотеза о случайности отвергается. Итак, для рассматриваемых рядов остатков имеем:

Из Таб. 2.1 видно, что условие выполняется для всех временных рядов, следовательно, гипотеза H_0 подтверждается и ряды остатков — случайны.

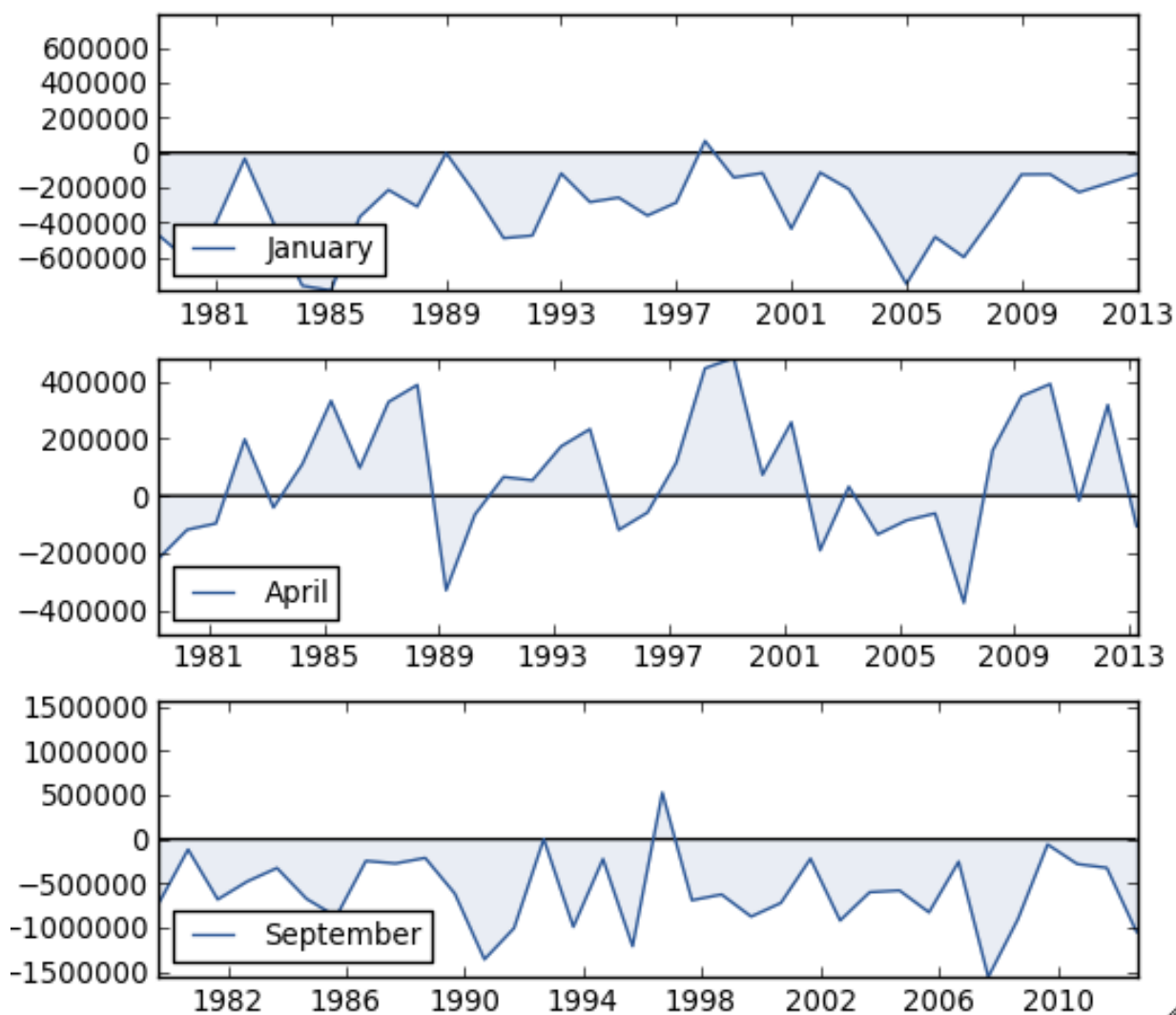


Рисунок 2.19 – Ряды остатков для января, апреля и сентября

Равенство нулю математического ожидания ряда остатков. Выдвигаем гипотезу $H_0 : M(e_t) = 0$. Для проверки гипотезы применим критерий Стьюдента.

Критерий Стьюдента. Рассчитывается статистика

$$t_{n-1} = \frac{|\bar{e}_t|}{S_e} \sqrt{n}, \text{ где } S_e = \sqrt{\frac{\sum (e_t - \bar{e}_t)^2}{n-1}}. \quad (2.4)$$

Если $T_n > t_{cr}(\alpha, n-1)$, то нулевая гипотеза отвергается. Для данных рядов $t_{cr}(0.05, 33) = 1.697$.

Результаты расчетов значения t для указанных рядов приведены в Таб. 2.2. Из таблицы видно, что данный критерий не выполняется. Включая в рассмотрение Рис. 2.19, можно сделать вывод, что модели дают

Месяц	p	p_{cr}	Месяц	p	p_{cr}
Январь	21	18	Июль	20	17
Февраль	19	18	Август	23	17
Март	19	18	Сентябрь	23	17
Апрель	22	18	Октябрь	26	18
Май	19	18	Ноябрь	22	18
Июнь	24	17	Декабрь	19	18

Таблица 2.1 – Расчет p_t и p_{cr} для остатков.

Месяц	t	Месяц	t
Январь	8.7948	Июль	8.3313
Февраль	10.063	Август	8.1358
Март	10.841	Сентябрь	8.0180
Апрель	1.9778	Октябрь	5.9249
Май	7.8563	Ноябрь	4.7382
Июнь	10.688	Декабрь	6.3513

Таблица 2.2 – Расчет t для остатков.

завышенное показание площади. При анализе результатов прогноза это необходимо учитывать.

Гомоскедастичность рядов остатков. Выдвигаем гипотезу H_0 : $D(e_t) = const$. Для проверки гипотезы применим критерий Гольдфельда-Куандта.

Критерий Гольдфельда-Куандта. Ряд остатков разбивается на две примерно одинаковые выборки ($n = n_1 + n_2$). Для обеих выборок рассчитывается значение дисперсии S_1^2, S_2^2 . Если статистика $F = \frac{S_{max}^2}{S_{min}^2}$, меньше $F_{cr}(\alpha, n_1 - k, n_2 - k)$, где $S_{max}^2 = \max(S_1^2, S_2^2), S_{min}^2 = \min(S_1^2, S_2^2), k$ - количество коэффициентов регрессии, то нулевая гипотеза принимается. Табличное значение $F_{cr}(0.05, 13, 13) = 2.53$.

Месяц	t	Месяц	t
Январь	1.17687	Июль	1.81335
Февраль	2.10160	Август	1.16440
Март	2.504013	Сентябрь	1.34222
Апрель	1.41677	Октябрь	1.01398
Май	1.42320	Ноябрь	1.69589
Июнь	1.32336	Декабрь	4.35583

Таблица 2.3 – Расчет статистики F для остатков.

Из Таб. 2.4 видно, что критерий $F < F_{cr}$, выполняется для всех рядов остатков кроме декабря. Для тех рядов, для которых условие выполняется, нулевая гипотеза принимается. Значения ряда остатков, условие для которого не выполнилось, являются гетероскедатичными, что означает неоднородность наблюдений, выражающуюся в неодинаковой дисперсии.

Отсутствие автокорреляционной зависимости в остатках. Выдвигаем гипотезу H_0 : уровни ряда остатков независимы. Для проверки гипотезы применим критерий Дарбина-Уотсона.

Критерий Дарбина-Уотсона. Рассчитывается статистика

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \approx 2(1 - r) \quad (2.5)$$

где r - коэффициент автокорреляции 1-го порядка. Если в остатках существует положительная автокорреляция, то $d \approx 0$. Если в остатках присутствует отрицательная автокорреляция, то $d \approx 4$. Если автокорреляция в остатках отсутствует, то $d \approx 2$. По специальным таблицам определяются критические значения критерия $d_H(\alpha, n, k)$ и $d_B(\alpha, n, k)$, где k – число степеней свободы. По этим значениям числовой промежуток $[0,4]$ разбивается на пять отрезков. Решение о принятии или отклонении гипотезы H_0 принимается в соответствии с Рис. 2.20.

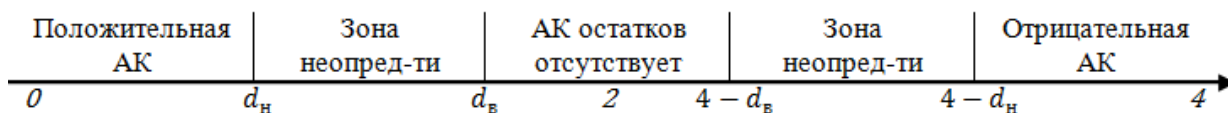


Рисунок 2.20 – Правило проверки гипотезы о наличии автокорреляции

Если значение статистики d попадает в зону неопределенности, то рассчитывается статистика

$$r(1) = \frac{\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=2}^n (e_t^2)} \quad (2.6)$$

и сравнивается с критическим значением коэффициентов корреляции r_{cr} . Если $r(1) \geq r_{cr}$, то нулевая гипотеза отвергается.

Месяц	t	Месяц	t
Январь	0.34604	Июль	0.77835
Февраль	0.31094	Август	0.80529
Март	0.30595	Сентябрь	0.83981
Апрель	1.38920	Октябрь	1.13994
Май	0.54874	Ноябрь	1.04494
Июнь	0.45268	Декабрь	0.85410

Таблица 2.4 – Расчет статистики d для остатков.

Полученные значения d попадают в интервал автокорреляции. Если бы строили прогноз для экстенда, то на полученных остатках можно было бы построить авторегрессионную модель и по ней строить прогноз.

Подчинение остатков нормальному закону распределения. Выдвигаем гипотезу H_0 : ряд остатков подчиняется нормальному закону распределения. Для проверки гипотезы воспользуемся правилом трех сигм. Правило трех сигм: если случайная величина имеет нормальное распределение с параметрами (m, σ^2) , то попадание ее в промежуток $(m - 3\sigma; m + 3\sigma)$ является практически достоверным.

При расчете все коэффициенты попали в допустимый интервал, следовательно, гипотеза H_0 выполняется, и что ряды остатков подчиняются нормальному закону распределения.

Таким образом, не все предпосылки регрессионного анализа выполняются. Как правило, нарушение предпосылок регрессионного анализа связано с тем, что на исходный ряд данных оказывают влияние факторы, не учтенные при построении модели.

2.6 Определение глубины прогноза

Наиболее распространенным методом определения глубины прогноза является предположение $n \leq \frac{B}{2...3}$, где n количество прогнозных значений ряда, B — длина имеющейся базы прогноза.

Для каждого месяца, имеется от 33 до 35 предшествующих значений, следовательно, для каждого ряда можно дать прогноз на 10 значений. Таким образом, можно построить прогноз на 10 лет.

Однако стоит отметить, что такой подход является не обоснованным для различных временных рядов. В данном подходе не учитывается качественные характеристики базы прогнозирования, например, такие как: степень зашумления, уровень колеблемости, степень взаимосвязь данных исходного ряда.

2.7 Прогнозирование

Поскольку не все предпосылки регрессионного анализа выполняются, можно сделать вывод, что модель не пригодна для точечного прогнозирования. Для наглядности результаты прогноза будут приведены в виде временного ряда экстенда с доверительными интервалами:

$$Y_{(n+l)} + t_{cr}S_l \leq Y_{(n+l)} \leq Y_{(n+l)} - t_{cr}S_l, \quad (2.7)$$

где \hat{Y}_{t+l} — точечный прогноз в момент времени $t = n + l, l = 1, 2, \dots, t_{cr} = t_{cr}(\alpha, n - k)$ — критическое значение Стьюдента, k — количество коэф-

коэффициентов регрессии, S_l^2 — оценка дисперсии точечного значения Y_{n+k} , определяемая по формуле:

$$S_l^2 = S^2 \left(1 + \frac{1}{n} + \frac{(n + l - \bar{t})^2}{\sum_{t=1}^n (t - \bar{t})^2} \right), \quad (2.8)$$

где S^2 — остаточная дисперсия.

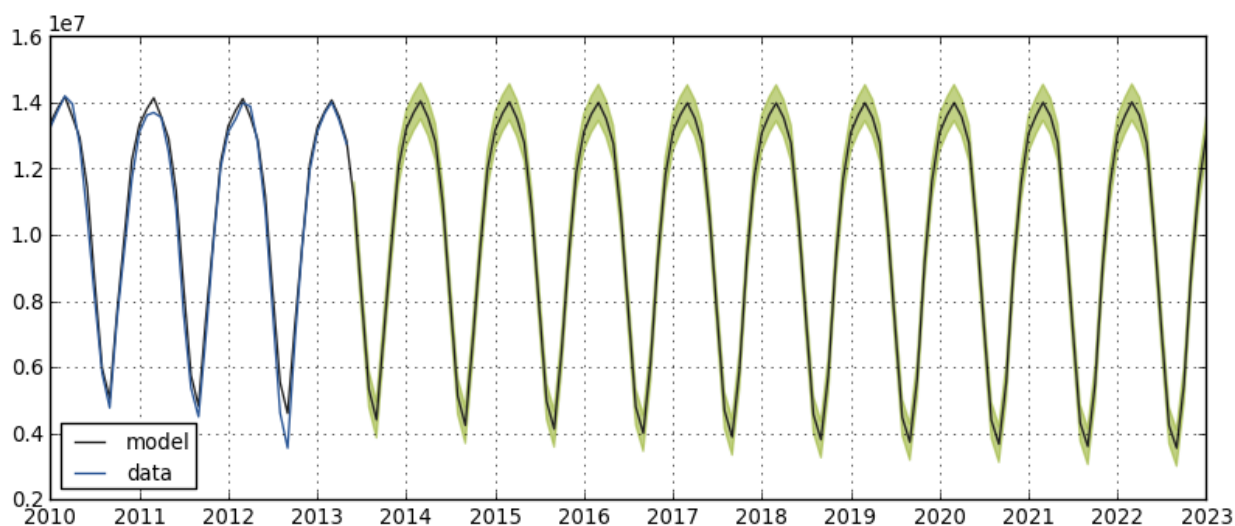


Рисунок 2.21 – Временной ряд экстенда, подсчитанного для полученного прогноза.

По графику на Рис. 2.21 видно, что значение показателя экстенда в сентябре продолжает уменьшаться. Такой прогноз не противоречит оценкам экспертов, в которых говорится, что площадь льда будет активно продолжать уменьшаться в летние месяцы и оставаться на одном уровне в зимние.

Стоит отметить недостаток данной модели: в ней не учтен объем. Ранее говорилось, что существуют спутниковые данные содержащие значения толщины льда, с помощью которых можно получить объем, однако доступ к ним получить не удалось.

Со временем площадь ледяного покрова будет уменьшаться, за счет того что таят многолетние льды. Как только многолетние льды растают, ледовая обстановка будет сильно зависеть от температуры и ветров в Арктическом регионе.

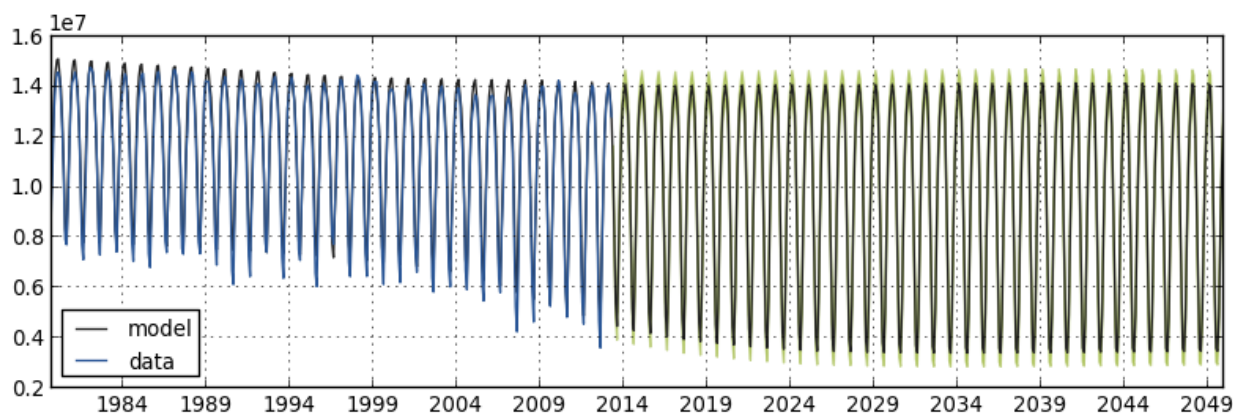


Рисунок 2.22 – Временной ряд экстенда, подсчитанного для полученного прогноза с экстраполяцией на длительный срок.

Предположим, что ледовая обстановка в Арктике сейчас находится в точке невозврата (момент, после которого уже нельзя повернуть назад), это даст возможность экстраполировать полученную модель на длительный период вперед. На Рис. 2.22 изображены временной ряд значений экстенда исходных данных и временной ряд экстенда, подсчитанного для спрогнозированных данных.

Начиная с 2029 года экстенд в сентябре перестает уменьшаться, это связано с тем что все временные ряды, для которых тренды имели отрицательные коэффициенты стали иметь постоянное значение 0, что соответствует открытой воде.

Стоит отметить, что по прогнозам SWIPA [21, стр. 9], будет наблюдаться рост температуры, что также окажет влияние на уменьшение экстенда, которое будет больше заметно в зимний период.

Для улучшения текущей модели, в неё можно добавить учет взаимосвязей между соседними пикселями, чтобы со временем концентрация льда в соседних пикселях не сильно отличалась.

2.8 Система мониторинга

В ходе написания магистерской диссертации был разработан прототип проекта по мониторингу акватории Арктического региона на основе данных дистанционного зондирования. Далее приведены: описание требований к функциональности системы, необходимые для отображения слои данных, описание архитектурного решения, а так же приведены снимки основных страниц веб-приложения с их описанием.

Задачи

Для предоставления пользователю возможности работать с данными необходимо реализовать базовую функциональность геоинформационных систем.

Визуализация данных в полярной стереографической проекции EPSG:3413.

Коллакация данных. Термин коллакация данных обозначает объединение слоев данных в одной области. Визуально коллакация может быть реализована с помощью наложения слоев с разными цветовыми градиентами и разными степенями прозрачности.

Временная шкала. Для отображения данных во времени необходим контроллер, который будет обеспечивать фильтрацию данных по времени.

Различные источники. Система должна обеспечивать возможность работы с различными источниками данных, а так же предоставлять панель управления для администрирования источниками и слоями отображаемыми на карте.

Модульность. Система должна поддерживать возможность создания модулей. Примеры модулей:

- Есть коллекция данных SSMI, позволяющая пользователю получить информацию о приводном ветре, водозапасе облаков и водя-

ном паре в атмосфере. Благодаря этой информации можно реализовать модуль, позволяющий идентифицировать на снимках полярные циклоны и получать о них информацию. Модуль записывает информацию в базу, либо выводит в файл в формате json или xml.

- Есть коллекция данных, из которой можно извлечь информацию о площади льда в Арктическом регионе. Можно реализовать модуль для расчета статистических показателей.
- Есть информация о путях передвижения полярных медведей и о покрытии льда. Можно определять медведей, оказавшихся в безвыходной ситуации.

Феномены. В примерах выше был описан модуль для поиска полярных циклонов. Результатом работы модуля является список полярных циклонов с их характеристиками. Такие полярные циклоны можно хранить в базе данных приложения и выводить их на общую карту.

Слой данных

В качестве базовых слоев для мониторинга акватории Арктического региона были выбраны:

1. Слои для подложки

- батиметрия: береговая линия, рельеф, изобаты, особые точки глубин;
- материка: контур, рельеф;
- административное деление: страны, столицы, населенные пункты;
- водные объекты: озера, реки (основные);
- порты: морские, речные.

2. Слои полученные в результате обработки данных дистанционного спутникового зондирования:

- температура морской поверхности;
- водяной пар в атмосфере;
- водозапас облаков;
- приводный ветер;
- морской лёд;
- абсолютные геострофические скорости;
- волны;
- абсолютная динамическая топография;
- полярные циклоны.

3. Сторонние слои и объекты:

- корабли;
- маршруты полярных медведей.

Архитектурное решение

На данный момент серверная часть прототипа системы мониторинга Арктики базируется на следующем программном обеспечении:

OPeNDAP Hyrax — сервер обеспечивающий доступ к спутниковым данным. Выбор данного сервера основывался на опыте французских коллег из компании Diateam (www.diateam.net). На основе данного сервера ими был разработан и поддерживается до сих пор проект Naiad (Dataminer спутниковых данных).

GeoServer — картографический сервер, получающий данные с OPeNDAP сервера и передающих их клиенту с использованием WMS спецификации. На сегодняшний день GeoServer является одним из наиболее мощных картографических серверов, распространяемых бесплатно по открытой лицензии. Кроме того GeoServer имеет хорошую документацию, позволяющую быстро решать поставленные задачи.

GeoDjango SIOWS — приложение, обеспечивающее централизованное хранение и доступ к данным. Выбор GeoDjango основывался на личном опыте автора работы с фреймворком Django. Данный фреймворк реально позволяет создавать в краткие сроки качественные полноценные веб-приложения.

В качестве базы данных используется СУБД PostgreSQL с расширением PostGIS. Использование PostgreSQL + PostGIS является наилучшей практикой в работе с GeoDjango приложениями.

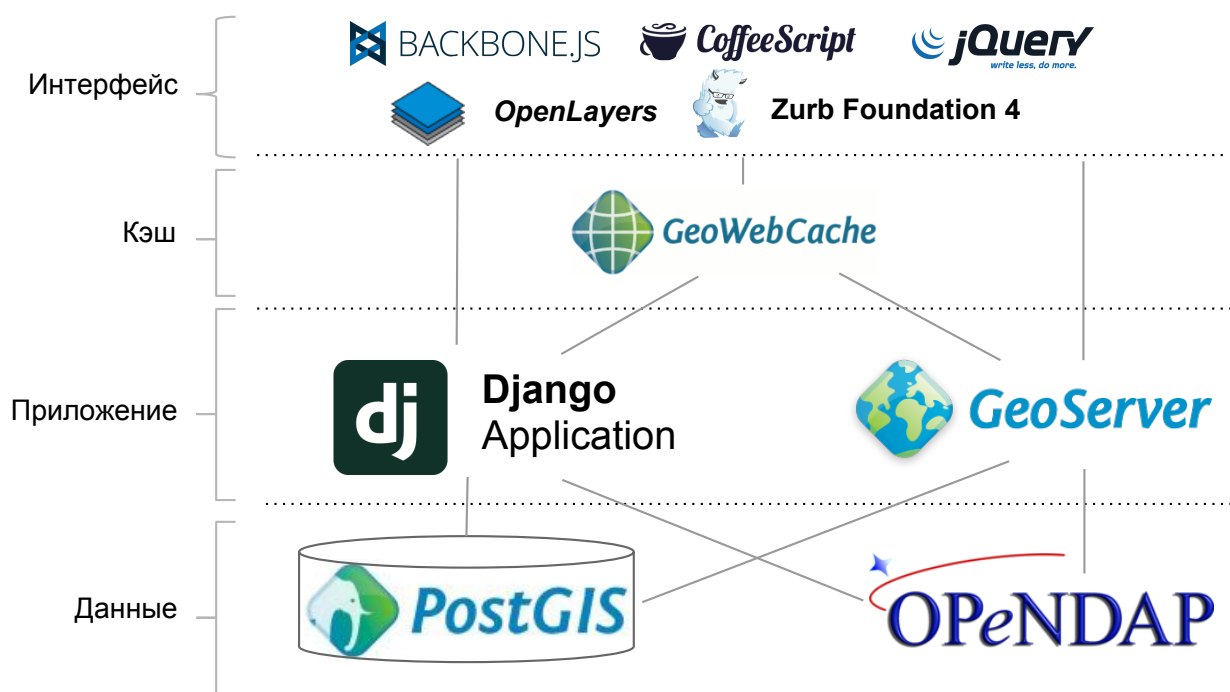


Рисунок 2.23 – Схематическое изображение архитектурного решения для системы мониторинга.

Далее перечислены, а также приведено краткое обоснование выбора, библиотек и фреймворков которые используются на стороне клиента. Описанные выбранного программного обеспечения приведено в разделе [1.4 Геоинформационные системы](#).

Backbone.js позволяет организовать структурированный javascript код. Будучи фреймворком Backbonejs обяывает front-end разработчика писать код с применением шаблона проектирования MVC. Фреймворк

принято использовать с библиотекой *underscore.js*, позволяющей упростить работу javascript структурами.

CoffeeScript является Ruby подобным языком программирования с приятным синтаксисом. CoffeeScript позволяет абстрагироваться от неполноценности javascript и писать код меньшего размера с улучшенной читаемостью.

jQuery, ставшая уже привычной всем библиотека, позволяющая удобно работать с элементами DOM.

OpenLayers наиболее распространенная бесплатная библиотека с открытым исходным кодом, позволяющая отображать карты в любой проекции на стороне пользователя. Библиотека OpenLayers выбрана из-за высоко уровня реализации и богатой функциональности.

Zurb Foundation 4 — молодой набирающий популярность веб-фреймворк для создания пользовательского интерфейса. Его выбор обусловлен интересом автора.

В будущем связка OPeNDAP и GeoServer, будет заменена на Рудар сервер, позволяющий предоставлять доступ ко всем необходимым данным и реализующий спецификацию WMS. Для перехода Рудар необходимо реализовать модуль для перепроецирования данных.

В случае осуществления перехода серверная сторона будет гомогенной (Рудар — сервер написанный на языке программирования Python, поддерживающий доступ к данные по opendap протоколу), что облегчает дальнейшую поддержку и развитие системы.

Пользовательский интерфейс

Текущее рабочее название системы мониторинга Арктики - SIOWS, что расшифровывается как Storm Ice Oil Watch System в переводе Система наблюдения за штормами, льдом и нефтяными разливами.

Пользовательский интерфейс разрабатывался на основе таких ГИС как MarineExplorer (ГИС для моря), 2GIS и GoogleMaps.

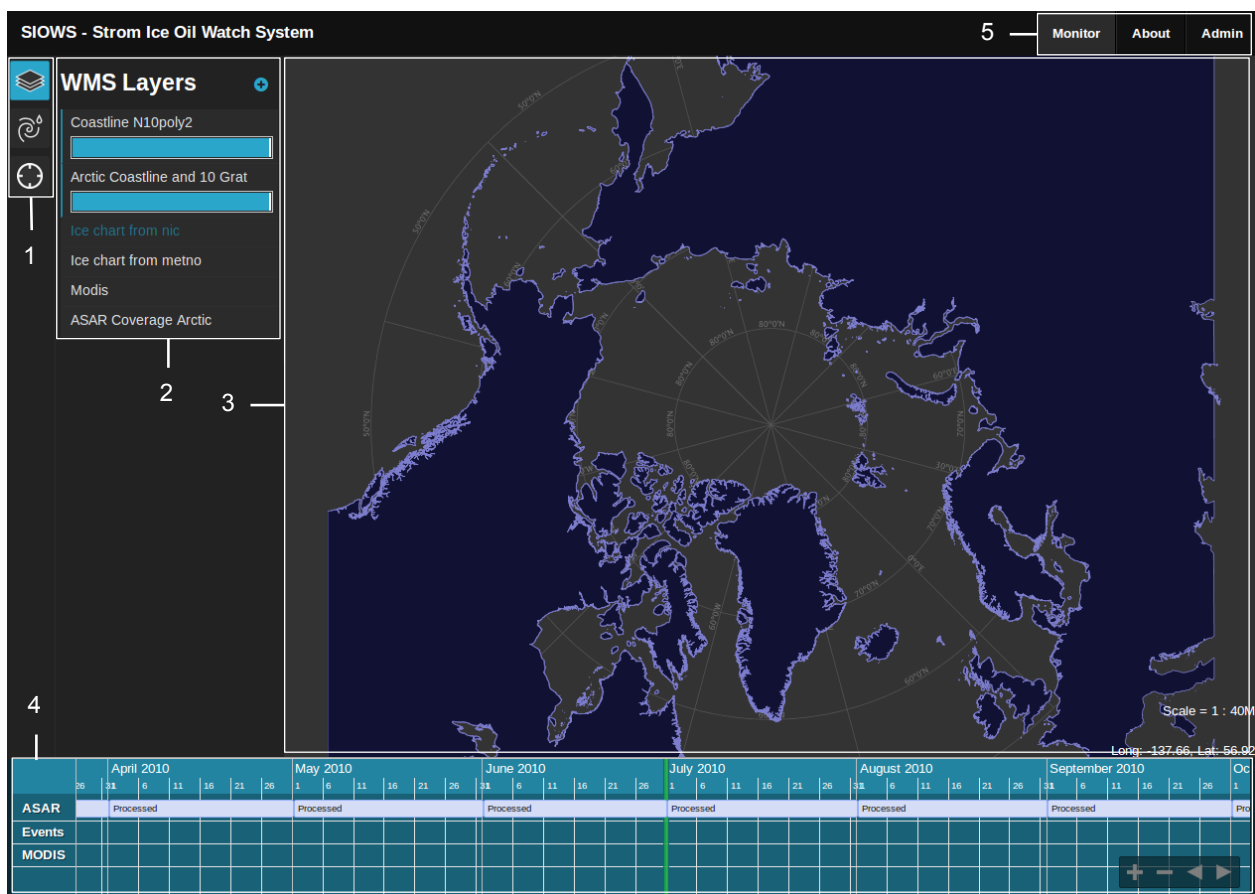


Рисунок 2.24 – Снимок главного экрана пользовательского интерфейса

Основные элементы пользовательского интерфейса:

1. **Вкладки боковой панели.** на данный момент реализованы вкладки для управления слоями, доступа к зарегистрированным феноменам и поиска географических объектов на карте.
2. **Активная вкладка.** На данный момент отображаются выбранные из архива слои, которые можно скрывать, настраивать прозрачность, а так же изменять порядок отображения.
3. **Карта.** На карту выводятся все активные слои. Сейчас активными слоями являются слой широтно-долготной сетки и слой земли с прибрежной линией.
4. **Временная шкала.** Временная шкала позволяет выводить на карту данные соответствующие времени. Так же на временной шкале

интервалом отображаются гранулы активных слоев, а точками - зарегистрированные феномены.

5. Панель доступа к административной панели.

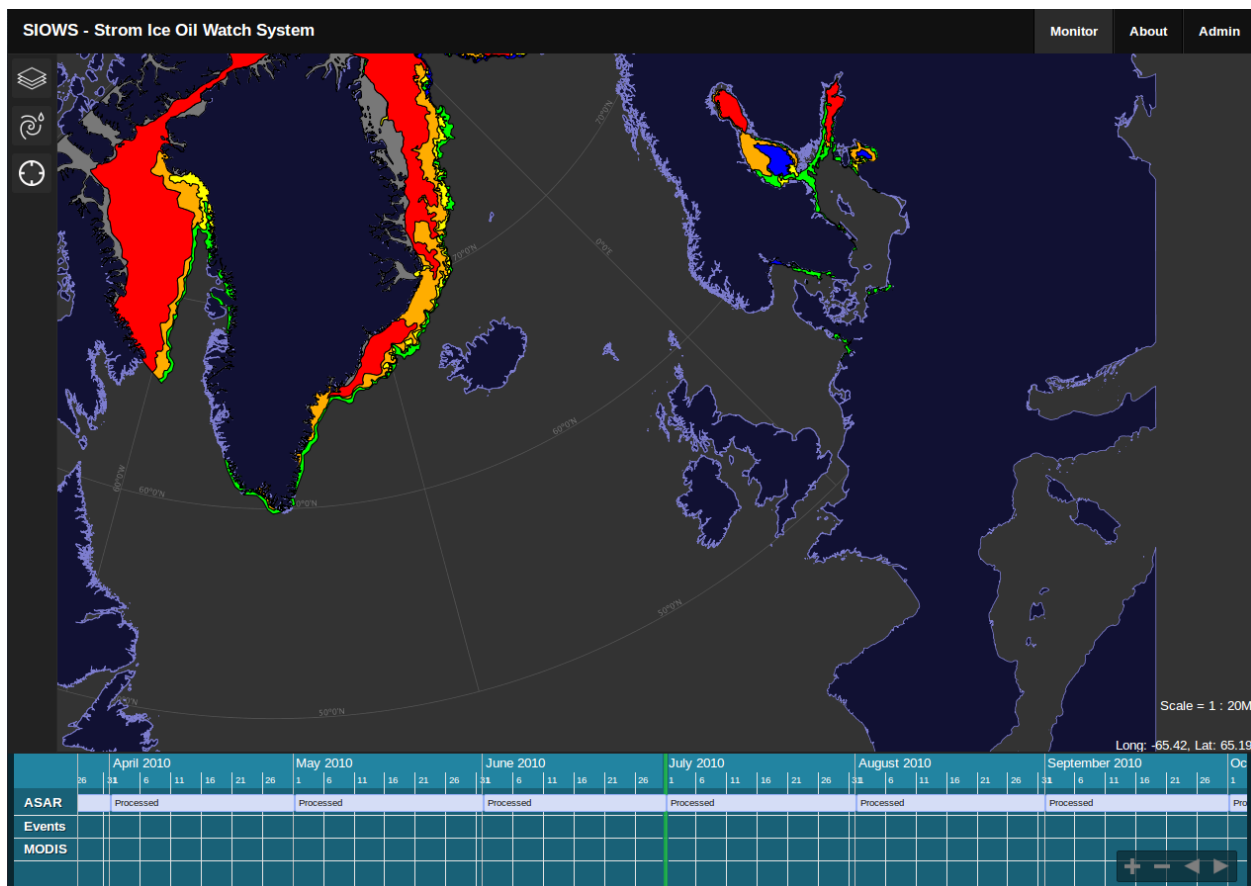


Рисунок 2.25 – Слой изолиний концентрации льда

На Рис. 2.25 на карте выведен слой с изолиниями концентрации льда. При переходе по ссылке с названием *Admin* пользователь попадает на страницу аутентификации, а после её прохождения и подтверждения прав доступа в административную панель. На Рис. 2.26 приведен снимок главного экрана административной панели.

Главный экран административной панели состоит из нескольких элементов:

1. панель быстрого доступа;
2. ссылки на страницы управления феноменами, в данном случае, полярные циклоны;

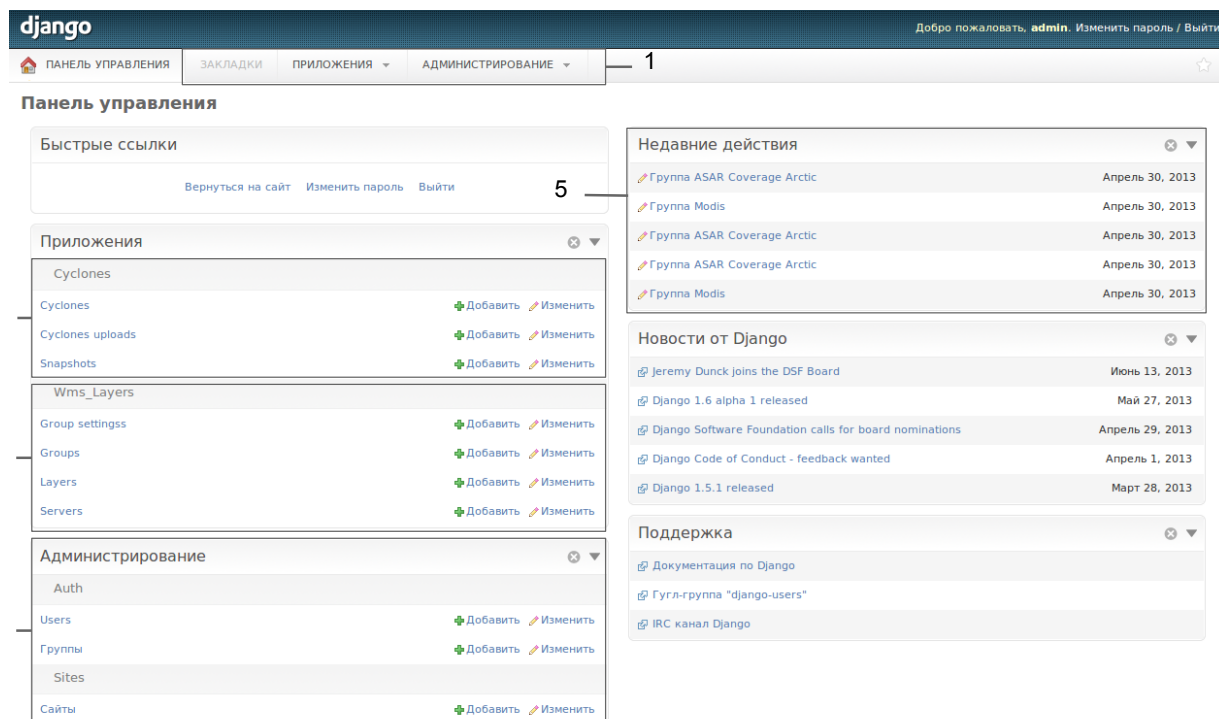


Рисунок 2.26 – Главный экран административной панели

3. ссылки на страницы управление слоями карты;
4. ссылки на страницы управление пользовательскими учетными записями;
5. список последних действия пользователя.

Приводить более детальное рассмотрение полученной в ходе написания магистерской диссертации системы нецелесообразно т.к. её создание не являлось целью данной работы.

Заключение

В данном параграфе перечислены выполненные задачи и описаны результаты, которые были получены в рамках магистерской диссертации.

На основе работ проекта SWIPA была рассмотрена, как текущая обстановка в Арктике, так и климатические прогнозы. В разделе [1.1 Ледовая обстановка в Арктике](#) приведены выводы из последнего отчета проекта SWIPA за 2011 год [21].

Рассмотрены область спутниковой океанографии и методы дистанционного зондирования необходимые для получения спутниковых данных, на основе которых производится анализ и строится прогноз ледовой обстановки.

В ходе построения прогноза создан фреймворк на языке Python с использованием библиотек Numpy, Scipy, pydap.client и matplotlib. Разработанный фреймворк позволяет:

- Производить сканирование архива данных на основе рекурсивного обхода каталога и выделения файлов по заданному регулярному выражению;
- Анализировать спутниковые данные. Базовые методы анализа описаны в разделе [2.2 Модель данных](#).
- Работать с коллекциями данных, выполнять запросы, подобные ORM.

Отличительной особенностью фреймворка от других существующих решений, является возможность работы с коллекциями данных.

Рассмотрены методы и модели прогнозирования, которые можно использовать для анализа и прогнозирования полей параметров представленных в спутниковых данных. Приведена классификация методов и моделей. В обзор методов вошли трендовые, регрессионные, авторегрессионные и модели экспоненциального сглаживания.

Построен долгосрочный интервальный прогноз конфигурации льда, который не противоречит оценкам экспертов. Для отображения прогноза, в работе приведен временной ряд экстенда, подсчитанного для спрогнозированных данных.

Предложен метод оценки адекватности аппроксимирующей модели. Для оценки адекватности аппроксимирующей модели предложено три приближения. В первом приближении в оценке адекватности учитывает схожесть модели и исходных данных. Оценка производится графическим методом или оценкой эксперта в области океанологии.

Рассмотрены стандарты, спецификации, протоколы, СУБД, библиотеки и фреймворки для создания веб-приложения с функциональностью геоинформационной системы.

Создан прототип системы визуализации данных во времени с возможностью коллокации различных слоев. Приведено обоснование выбранного инструментария.

Дальнейшая работа

В ходе выполнения магистерской диссертации были выявлены новые задачи, которые планируется решить в будущем.

Улучшение прогноза. Для улучшения прогноза будет полезна информация не о концентрации льда, а об его толщине. Такие исходные данные существуют, однако, своевременный доступ к ним получить не удалось. Так же прогноз можно улучшить за счет использования модели ARMA. Стоит помнить, что недостатком авторегрессионных моделей является большое число параметров модели, определение которых весьма затруднено и не может быть полностью автоматизировано. Использование модели ARMA может стать возможным если произвести кластеризацию временных рядов по их взаимной корреляции.

Фреймворк для анализа спутниковых данных. На данный момент основным отличием полученного фреймворка от других существующих библиотек является возможность удобной работы с коллекцией. Направлением дальнейшего развития будет интеграция с фреймворком Nansat (фреймворк для работы со спутниковыми данными), а также библиотекой Pandas (библиотека для анализа данных).

Прототип системы мониторинга Арктики. Дальнейшим шагом развития полученного прототипа является системы мониторинга Арктики. Для этого необходимо:

- Доработать WMS модуль Рудар сервера. Реализовать поддержку различных проекций.
- Доработать модуль полярных циклонов, позволяющий отображать на карте, зафиксированные полярные циклоны.

Литература

- [1] Анализ временных рядов прогноз и управление: Пер. с англ. Вып. 1 / Дж Бокс, Г Дженкинс, А Л Левшин, В Ф Писаренко. — Мир, 1974.
- [2] Трубина Л.К. и Быкова О.Г. Геоинформационные системы. — 2012.
- [3] Коновалова Н.В. и Капралов Е.Г. Введение в ГИС. — М., 1997.
- [4] Морские гидрологические прогнозы: Учебное пособие / Ксения Ивановна Кудрявая, Евгений Ильич Серяков, Лариса Ивановна Скриптунова и др. — Гидрометеиздат, 1974.
- [5] ООН Резолюция Генеральной Ассамблеи. 41/65 от 3 декабря 1986 г Принципы, касающиеся дистанционного зондирования Земли из космического пространства. — 1986. — 12.
- [6] Воронцов К. В. Лекции по алгоритмам кластеризации и многомерного шкалирования. — 2007.
- [7] США Геологическая служба. 90 Billion Barrels of Oil and 1,670 Trillion Cubic Feet of Natural Gas Assessed in the Arctic. — 2008. — URL: <http://www.usgs.gov/newsroom/article.asp?ID=1980>.
- [8] Худсон Дерек. Статистика для физиков. — М.: Мир, 1970.
- [9] Лоскутов АЮ. Анализ временных рядов. — 1988.
- [10] Практикум по эконометрике / ИИ Елисеева, СВ Курышева, НМ Гордеенко, ИИ Елисеева. — 2002.

- [11] Носко ВП. Эконометрика. Введение в регрессионный анализ временных рядов // М.: МФТИ. — 2002.
- [12] Тихонов ЭЕ. Прогнозирование в условиях рынка. — 2006.
- [13] Чучуева ИА. Модель прогнозирования временных рядов по выборке максимального подобия // Дисс. на соиск. уч. ст. к. т. н., Москва. — 2012.
- [14] Костикова Анна. Классификация картографических веб-сервисов OGC. — 2012. — URL: <http://gis-lab.info/qa/ogc-intro.html>.
- [15] Чучуева ИА. Классификация методов и моделей прогнозирования. — 2013. — URL: <http://habrahabr.ru/post/177633/>.
- [16] Ханнерц Фредрик. Арктические изменения и их глобальное влияние // Встреча Министров Охраны Окружающей Среды стран Арктического Совета. — 2013.
- [17] Иванов Данил. Проект SATIN. Построение информационного портала для предоставления доступа, поиска и визуализации данных дистанционного зондирования. // Четвертая международная Школа-семинар: «Спутниковые методы и системы исследования Земли» / Институт космических исследований РАН. — 2013.
- [18] Эконометрика / ИИ Елисеева, СВ Курышева, ТВ Костеева, ИВ Бабаева. — 2001.
- [19] МПР. Эвакуации станции «Северный полюс – 40» // Пресс-служба Минприроды России. — 2013. — 5. — URL: <http://www.mnr.gov.ru/news/detail.php?ID=130894>.
- [20] Accelerated decline in the Arctic sea ice cover / Josefino C Comiso, Claire L Parkinson, Robert Gersten, Larry Stock // Geophysical Research Letters. — 2008. — Vol. 35, no. 1.

- [21] Arctic Climate Issues 2011: Changes in Arctic Snow, Water, Ice and Permafrost. — 2011.
- [22] Arctic sea ice decline: Faster than forecast / Julienne Stroeve, Marika M Holland, Walt Meier et al. // Geophysical research letters. — 2007. — Vol. 34, no. 9.
- [23] Beitler Jane. Arctic sea ice extent settles at record seasonal minimum. — 2012. — URL: <http://nsidc.org/arcticseaicenews/2012/09/arctic-sea-ice-extent-settles-at-record-seasonal-minimum/>
- [24] Cavalieri D. C. Parkinson P. Gloersen, Zwally. H. J. Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data. — 1996, updated yearly. — Boulder, Colorado USA: NASA DAAC at the National Snow and Ice Data Center.
- [25] CryoSat-2 estimates of Arctic sea ice thickness and volume / Seymour W Laxon, Katharine A Giles, Andy L Ridout et al. // Geophysical Research Letters. — 2013. — Vol. 40. — P. 732–737.
- [26] CubeSat. Miniaturized satellite for space research. — 2013. — URL: <http://cubesat.org>.
- [27] Huler Scott. Defining the Wind: The Beaufort Scale and How a 19th-Century Admiral Turned Science into Poetry. — Crown, 2007.
- [28] Jung Thomas. Arctic climate change: Mechanisms, impacts and predictability // Arctic climate change: Mechanisms, impacts and predictability. — 2013. — 4.
- [29] Maslanik J., Stroeve J. Near-Real-Time DMSP SSM/I-SSMIS Daily Polar Gridded Sea Ice Concentrations. — 1999, updated daily. — Boulder, Colorado USA: National Snow and Ice Data Center. Digital media.

- [30] NASA National Aeronautics and Space Administration. Earth system science data resources: tapping into a wealth of data, information, and services, NP-2007-11-859-GSFC. — 2007.
- [31] NSIDC. Frequently Asked Questions on Arctic sea ice. — 2013. — URL: <http://nsidc.org/arcticseaicenews/faq>.
- [32] NSIDC. Monthly Sea Ice Extent Images. — 2013. — URL: http://nsidc.org/data/docs/noaa/g02135_seaice_index/#monthly_extent_image.
- [33] NSIDC. Near-Real-Time DMSP SSM/I-SSMIS Daily Polar Gridded Sea Ice Concentrations. — 2013. — URL: <http://nsidc.org/data/nsidc-0081.html>.
- [34] NSIDC. Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data. — 2013. — URL: <http://nsidc.org/data/nsidc-0051.html>.
- [35] Robinson Ian S. Measuring the oceans from space: the principles and methods of satellite oceanography. — Springer Verlag, 2004.
- [36] SATIN. Satellite Data Search and Manage Information Portal. — 2013. — URL: <http://satin.rshu.ru>.
- [37] Smedsrud Lars H. Arctic climate change: Local impacts of Global significance // Arctic climate change: Mechanisms, impacts and predictability. — 2013. — 4.