

**Федеральное государственное бюджетное образовательное  
учреждение высшего образования  
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА  
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ  
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»**

**Пойда А.А., Жижин М.Н., Трусов А.В.,  
Андреев А.В., Матвеев А.М.**

**Детектирование судовых огней на море**

Москва 2019

**Аннотация.** Лавинообразное увеличение объема спутниковых данных, получаемых современными сенсорами, стало основным препятствием к их использованию для ручного обнаружения рыболовных судов промысловыми агентствами и другими организациями. В связи с этим потребовалась разработка алгоритма и реализующей его автоматической системы детектирования ночных судовых огней по спутниковым данным и анализа их распределения. В данной работе представлен алгоритм детектирования ночных судовых огней по спутниковым данным сенсора VIIRS, описана программная система, реализующая разработанный алгоритм, описаны разработанные методы и средства анализа распределения ночных судовых огней, представлены результаты апробации разработанных методов.

Матвеев А.М., младший научный сотрудник международной лаборатории математических методов исследования социальных сетей ИПЭИ Российской академии народного хозяйства и государственной службы при Президенте РФ

Андреев А.В., инженер международной лабораторией математических методов исследования социальных сетей ИПЭИ Российской академии народного хозяйства и государственной службы при Президенте РФ

Жижин М.Н., ведущий научный сотрудник международной лаборатории математических методов исследования социальных сетей ИПЭИ Российской академии народного хозяйства и государственной службы при Президенте РФ

Пойда А.А., старший научный сотрудник международной лабораторией математических методов исследования социальных сетей ИПЭИ Российской академии народного хозяйства и государственной службы при Президенте РФ

Трусов А.В., заведующий международной лабораторией математических методов исследования социальных сетей ИПЭИ Российской академии народного хозяйства и государственной службы при Президенте РФ

Данная работа подготовлена на основе материалов научно-исследовательской работы, выполненной в соответствии с Государственным заданием РАНХиГС при Президенте Российской Федерации на 2018г.

# СОДЕРЖАНИЕ

Введение

Глава 1. Алгоритм детектирования судовых огней VIIRS Boat Detection (VBD)

1.1 Общая схема алгоритма

1.2 Базовые индексы (SMI, SI, SHI) и детектор молний

1.3 Фильтрация лунных бликов от волн и облаков

Глава 2. Реализация алгоритма детектирования судовых огней и его оптимизация

2.1 Реализация алгоритма детектирования судовых огней

2.2 Оптимизация алгоритма детектирования ночных огней

2.3 Оптимизация алгоритма межканальной корреляции в скользящем окне с учетом многоядерной архитектуры

Глава 3. Пространственно-временная база данных

Глава 4. Веб-приложение для визуализации и статистического анализа распределения судовых огней

4.1 Tile-сервер для отображения плотных облаков точек VBD детекций

4.2 Кластерный анализ карт VBD детекций

Глава 5. Анализ распределения детекций у Дальневосточных берегов Российской Федерации

5.1 Морские платформы

5.2 Зоны промыслового рыболовства у берегов п-ова Камчатка

5.3 Суда в портах и на рейдах

5.4 Детекции воздушных линий

Глава 6. Индекс эффективности ограничений рыбного промысла

6.1 Определение индекса эффективности ограничений рыбного промысла

6.2 Экспериментальные исследования и апробация индекса эффективности ограничений рыбного промысла

Глава 7. Инструмент для анализа данных «Google Earth Engine Assessment Tool»

7.1 Технические аспекты

7.2 Обзор приложения

7.3 Работа с приложением

**7.4 Пример использования**

Заключение

Литература

## ВВЕДЕНИЕ

С 70-х годов прошлого века известно, что при помощи высокочувствительных сенсоров, размещенных на спутниках, можно обнаруживать яркие судовые огни в ночное время суток (Croft, 1978; Croft&Colvocoresses, 1979). В первую очередь речь идет о бортовых огнях рыболовецких судов, которые используют электрическое освещение для привлечения рыбы в ночное время суток. Для рыболовной промышленности использование спутниковых данных имеет большое значение, что отражено во многих работах, например (Cho et al., 1999; Kiyofuji&Saitoh, 2004; Maxwell et al., 2004; Rodhouse et al., 2001; Waluda et al., 2002; Waluda et al., 2004; Zhang et al., 2013) и др. В 2015 году Уильям Страка с коллегами показали (Straka et al. 2015), что при помощи спутниковых данных можно аналогично обнаруживать и другие суда.

Первой системой для мониторинга судовых огней, был сенсор OLS, размещенный на спутнике DMSP, и ряд сервисов по хранению и каталогизации получаемых с него данных. В 2001 году был запущен новый метеорологический спутник Suomi NPP, который дал новый толчок в области анализа спутниковых данных, полученных с ночной стороны Земли (Miller et al., 2013; Miller et al., 2012; Elvidge et al., 2013; Schueler et al., 2013). Спутник оснащен мультиспектральной камерой видимого и инфракрасных (ИК) диапазонов с шириной охвата 3000 км. Сенсор VIIRS имеет существенные технические преимущества перед сенсором OLS: в 4 раза увеличенное разрешение (750 м на пиксель), увеличенный динамический диапазон (14 бит), постоянная радиометрическая калибровка, дополнительные спектральные диапазоны.

За счет улучшенных по сравнению с OLS технических характеристик сенсор VIIRS позволяет обнаруживать больше судов, но и объем данных, получаемых с сенсора VIIRS при этом, более чем в 200 раз превышает объем данных, получаемых с сенсора OLS. Такое значительное увеличение объема данных стало основным препятствием для использования спутниковых данных, получаемых с сенсора VIIRS, для «ручного» обнаружения рыболовных судов, а системы для автоматической обработки и обнаружения бортовых огней по спутниковым данным не было.

Чтобы решить проблему трудоемкости «ручного» обнаружения рыболовных судов на спутниковых данных, была разработана система для автоматического детектирования бортовых огней судов. В первую очередь система ориентирована на работу с данными сенсора VIIRS.

Разработанная система может быть использована как для оперативного мониторинга морского рыболовства, так и для анализа долгосрочных экологических последствий ограничений на вылов рыбы и морских экологических катастроф техногенного и природного происхождения.

# **ГЛАВА 1. АЛГОРИТМ ДЕТЕКТИРОВАНИЯ СУДОВЫХ ОГНЕЙ VIIRS BOAT DETECTOR (VBD)**

## **1.1 Общая схема алгоритма**

Алгоритм детектирования морских судов предназначен для автоматического выявления точечных источников излучения в видимом диапазоне на территории морей и океанов в ночное время суток. К таким источникам в первую очередь относятся морские суда (в том числе ведущие рыбный промысел с использованием света для привлечения улова) и морские платформы.

Изначально алгоритм ориентирован на использование данных, полученных при помощи сенсора VIIRS, установленного, например, на спутниках Suomi NPP и NOAA-20, но он может быть адаптирован и к другим сенсорам, получающим данные в видимом и ИК-спектрах, близких к спектрам VIIRS.

Схема алгоритма представлена на рисунке 1.1. Исходными данными являются изображения, полученные в видимом спектре и их геопривязка. Так как значения в DNB канала низкие, то для удобства они умножаются на 109 степени (перевод единиц измерения из Вт в нВт). В процессе обработки данные проходят ряд фильтров, описанных ниже.

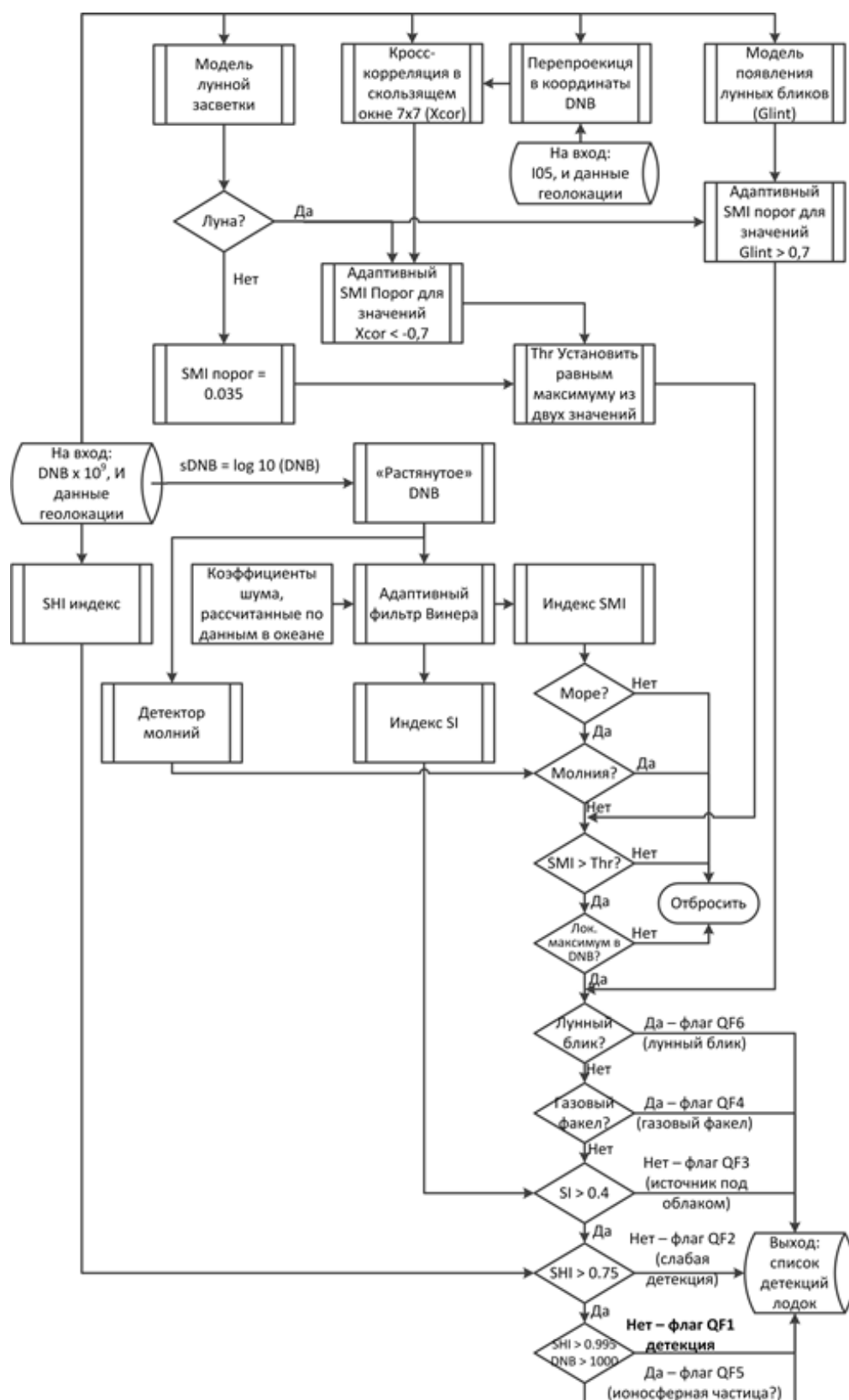


Рисунок 1.1 – Схема алгоритма детектирования морских судов

## 1.2 Базовые индексы (SMI, SI, SHI) и детектор молний

На первом этапе производится преобразование исходного изображения по формуле  $sDNB = \log_{10}(DNB)$ , где  $DNB$  - значение яркости исходного изображения в видимом диапазоне. Эта операция делается, чтобы уменьшить перепады значений яркости между сильными и относительно слабыми источниками. Полученное изображение называется “растянутым”.

На следующем этапе применяется фильтр Винера (Lim et al., 1990), подавляющий высокие частоты при низком уровне сигнала, которые рассматриваются как тепловой шум сенсора и алгоритма оцифровки. Этот шум распределяется вдоль линии сканирования неравномерно. В центре он ниже, а по краям достигает максимальных значений. Параметром фильтра Винера является вариация уровня сигнала на изображении вдоль линии сканирования, описываемая полиномом 6-й степени. Коэффициенты полинома выведены экспериментально на основании данных, полученных в ночное время в Тихом океане.

### *Высокочастотный (медианный) индекс SMI*

Результат применения фильтра Винера передается на следующий шаг, на котором к нему применяется медианный фильтр (Gonzales et al., 2002) в скользящем окне размера 3x3. Полученный результат вычитается из результата, полученного после применения фильтра Винера. Результат называется SMI индекс (Spike Median Index). Для дальнейшего анализа в качестве кандидатов в детекции выбираются пиксели исходного изображения, в которых наблюдается локальный максимум значений яркости в видимом диапазоне, и значение SMI индекса которых превышает заданный адаптивный порог. Порог SMI индекса меняется вдоль линии сканирования, чтобы превысить упомянутый ранее шум, уровень которого повышается на краях линии сканирования. На рисунке 1.2 показано изменение порога SMI индекса вдоль линии сканирования.

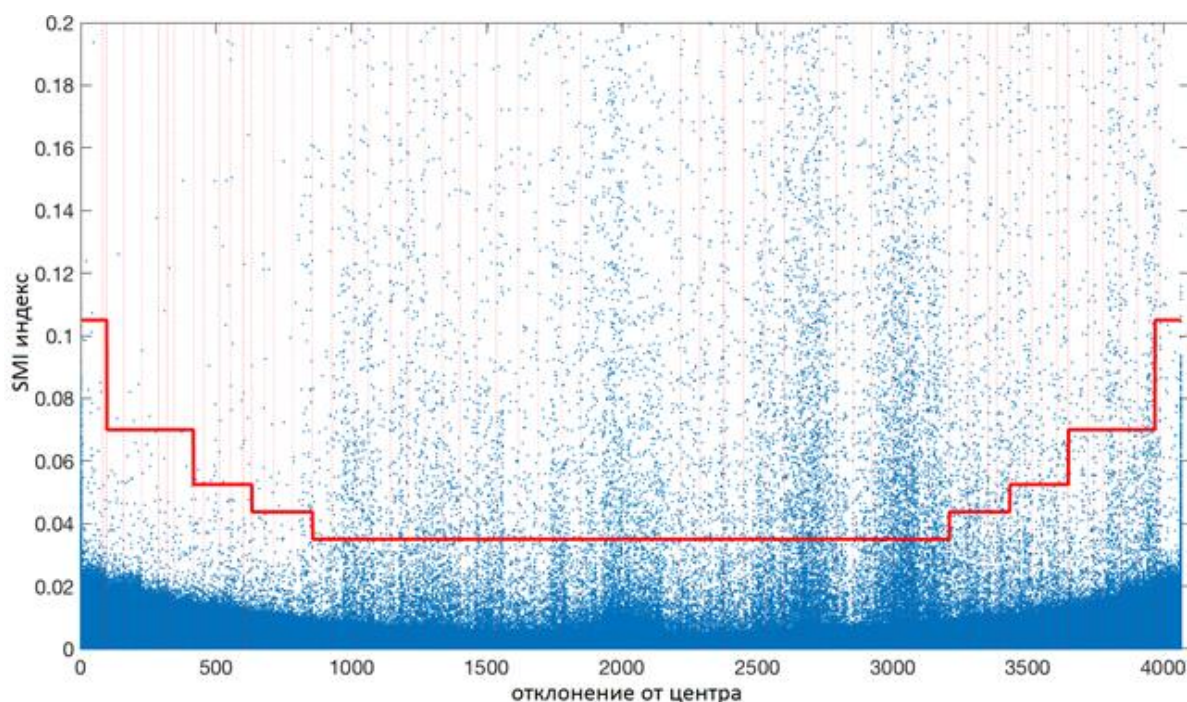


Рисунок 1.2 – Изменение SMI индекса вдоль линии сканирования

### Индекс локальной резкости SI

Фильтрация резкости сигнала позволяет обнаружить участки изображения, в которых сигнал размыт. Наиболее частой причиной размытия сигнала является наличие облачности. Но могут быть и другие причины, например, задымление. Размытые детекции не принимаются к дальнейшему рассмотрению или помечаются соответствующими флагами (в зависимости от уровня размытия).

Для определения меры локальной нерезкости изображения в заданной его области используется Sharpness Index (SI), специально разработанный для данного алгоритма. Вычисление индекса основывается на методе спектральной резкости (Vu et al., 2012), который использует тот факт, что на спутниковом изображении, полученном в ночное время, сигнал вокруг светового источника складывается из рассеянного сигнала от облаков и аэрозолей (рассеянное излучение) и прямого сигнала (не рассеянного), идущего напрямую от источника. Если в сигнале нет составляющей рассеянного излучения или оно мало (т.е. в этой области изображение является более резким), то в этой области перепады сигнала более резкие, и значит спектр Фурье будет быстрее затухать с ростом пространственной частоты.

На рисунке 1.3 представлено спутниковое изображение, полученное в канале DNB сенсора VIIRS в ночное время севернее Джакарты 27 сентября 2014 года в 18:34 UTC (слева) и соответствующая ему карта SI-индекса, вычисленного в скользящем окне 32x32 пикселя. На изображении видны огни газовых факелов, морских платформ и рыболовных судов. На карте SI-индекса видно, что более размытые участки имеют более



низкие значения SI-индекса (темные пятна). Все кандидаты в детекции, значение SI индекса которых не превышает 0,4, не рассматриваются в качестве детекций.

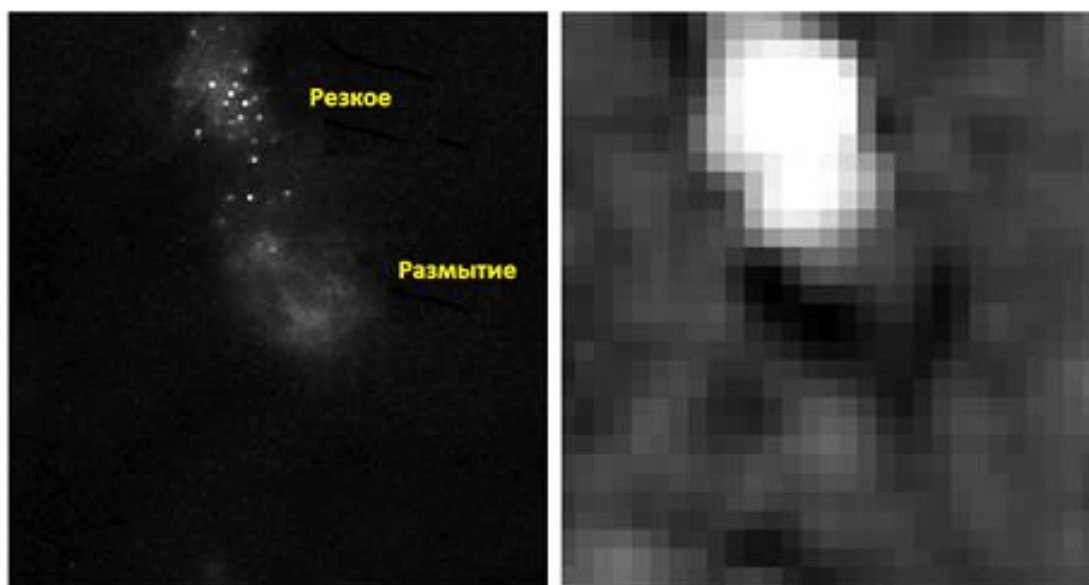


Рисунок 1.3 – Пример изменения SMI индекса

### Индекс относительной высоты спайков

Индекс SHI (Spike Height Index) используется для определения локальных ярких точек (т.е. точек, в которых значение яркости существенно превышает среднее значение в окружающих точках). Ниже представлена формула для расчета SHI-индекса.

$SHI = \max(SHI_H, SHI_V)$ , где:

$$SHI_H = \begin{cases} \frac{I_0 - (I_{H,PREV} + I_{H,NEXT}) / 2}{I_0}, & \text{if } I_0 > 0 \\ 0 & , \text{otherwise} \end{cases}$$

$$SHI_V = \begin{cases} \frac{I_0 - (I_{V,PREV} + I_{V,NEXT}) / 2}{I_0}, & \text{if } I_0 > 0 \\ 0 & , \text{otherwise} \end{cases}$$

$I_0$  - исследуемый пиксель,  $I_{V,PREV}$ ,  $I_{V,NEXT}$ ,  $I_{H,PREV}$ ,  $I_{H,NEXT}$  — пиксели, соседние исследуемому по строке и по столбцу.

Если значение SHI индекса меньше 0,75, то точка считается недостаточно изолированной, такая детекция все равно попадет в результирующий список, но будет иметь соответствующий флаг.

### **Детектор молний**

Детектор молний предназначен для фильтрации вспышек молнии, представляющихся на изображении в виде ярких полос одинаковой ширины и переменной длины. Постоянная ширина равная 16 пикселям объясняется тем, что камера VIIRS одновременно снимает 16 соседних линий скана. А длина линии зависит от того, сколько камера успела снять за время вспышки молнии.

Чтобы зафиксировать молнию, алгоритм проверяет стыки 16-строчных сканов, и если обнаруживает, что яркость соседних по столбцу пикселей, находящихся в разных сканах, превышает по относительному значению некоторый заданный порог, то помечает точку стыка. Если после этого обнаружится непрерывная цепочка таких помеченных точек, превышающая по длине заданный порог  $L$ , то в том скане, в котором значения яркости больше, отбрасывается ряд пикселей. Отбрасываются все пиксели скана, входящие в те же столбцы, в которые входили помеченные точки, а также в столбцы, отстоящие от помеченных точек на расстояние не превосходящее  $L$ . Для водной территории значение  $L$  следует брать меньше, чем для обработки спутниковых данных, полученных на территории суши, так как на суше большая плотность огней и вероятность случайного совпадения выше. Экспериментальным путем было установлено, что приемлемыми значениями  $L$  являются значения 24 и 16 для суши и моря соответственно.

### **1.3 Фильтрация лунных бликов от волн и облаков**

Для решения задачи фильтрации лунных бликов от поверхности воды и от облаков, решаются 3 подзадачи:

1. Определение факта наличия или отсутствия лунной засветки.
2. Определение вероятности появления бликов от волн в различных областях снимка.
3. Определение наличия облачности в анализируемой окрестности.

Для решения задачи определения факта наличия или отсутствия лунной засветки используется модель лунной засветки (Janiczek and Deyoung, 1987), позволяющая определить для заданного времени суток и координат величину лунной засветки в Люксах. Так как обрабатываемый снимок покрывает большую площадь, то он делится на 6 зон (Рисунок 1.4). Принадлежность пикселя к зоне определяется на основании величины лунной засветки в этом пикселе в сравнении с величиной лунной засветки в других пикселях обрабатываемого изображения. Для этого на всем множестве значений лунной засветки во всех пикселях изображения вычисляются перцентили, и на основании полученных перцентилей все пиксели причисляются к одной из 6 зон.

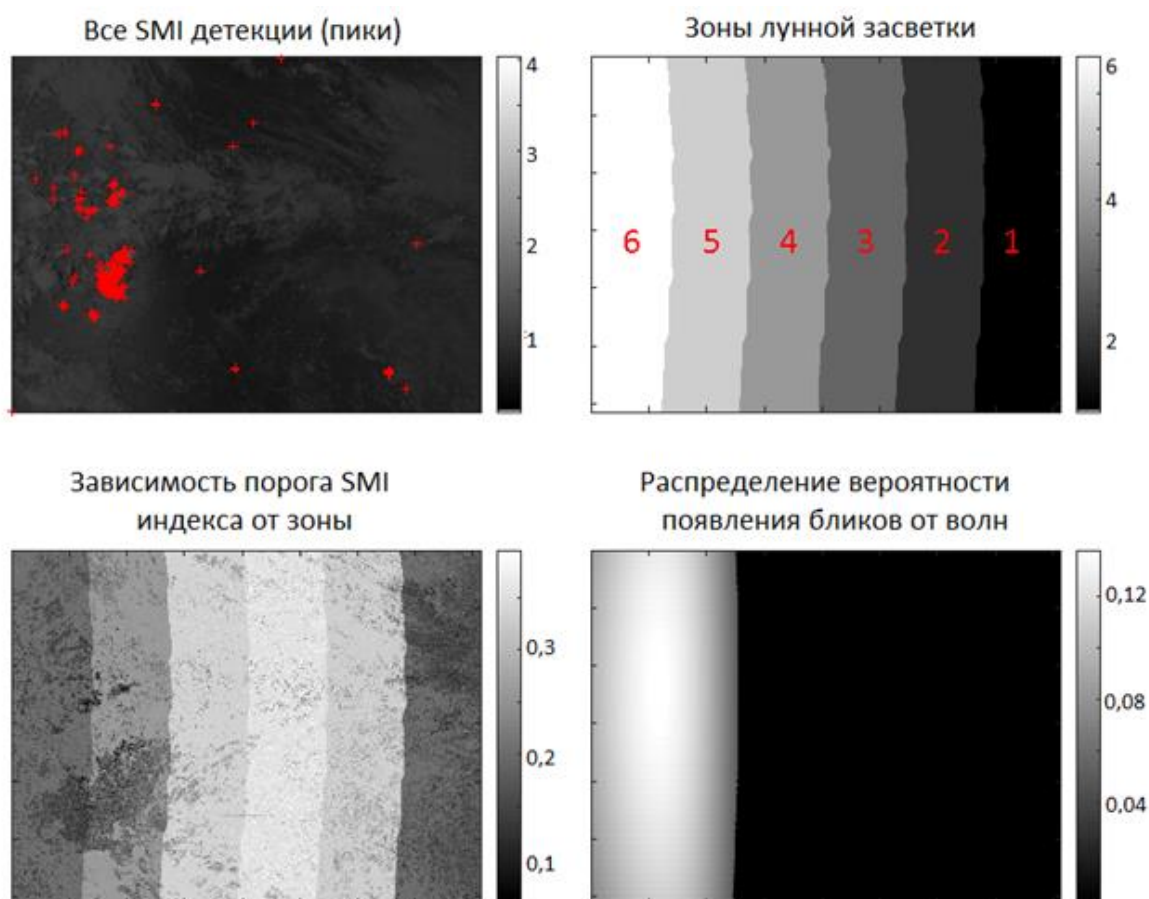


Рисунок 1.4 – Деление спутникового изображения на 6 зон по величине лунной засветки.

Разбиение изображения на зоны позволяет для каждой зоны вычислять свою величину засветки независимо от соседней зоны. Если верхняя граница зоны не превышает значения 0,0001 люкс, то считается, что для пикселей этой зоны нет лунной засветки. В этом случае фильтрация для пикселей данной зоны не делается. В противном случае применяется фильтрация лунных бликов как от поверхности воды, так и от облаков.

Для фильтрации лунных бликов от поверхности воды используется статистическая модель, предложенная Коксом и Манком (Cox, C. and Munk, W. 1954) для распределения яркости солнечных бликов на поверхности моря. Модель была адаптирована на случай лунной засветки. В качестве входных данных используется информация о положении спутника, луны, скорости и направлении ветра. Результат модели - вероятность появления бликов в заданной области. Ожидаемая яркость бликов определяется как математическое ожидание, умноженное на яркость лунной засветки с экспериментально выбранным масштабным множителем. В зависимости от ожидаемой яркости устанавливается порог SMI индекса, рассчитываемого для каждой потенциальной детекции. Все анализируемые точки, в которых значение SMI индекса меньше заданного порога, помечаются как маловероятные, и попадут в результирующий список с этой пометкой, если не будут отброшены на последующих

этапах анализа (см. фильтрацию лунных бликов от облаков). Для каждой из 6 зон порог SMI индекса устанавливается независимо от других зон.

Для фильтрации лунных бликов от облаков производится проверка факта наличия облачности в анализируемом районе. Проверка основывается на уровне корреляции между видимым и инфракрасным каналом в окрестностях точки-кандидата в детекцию. Для корреляции берутся значения пикселей в некоторой заданной окрестности исследуемой точки и выстраиваются в одномерный ряд. Получается два ряда: ряд значений в видимом диапазоне и ряд значений в ИК-диапазоне. Если в окрестности исследуемой точки присутствует облачность, то ожидается, что корреляция будет стремиться к отрицательным значениям, так как облака будут темнее воды (так как их освещает луна), но и температура их ниже, чем температура воды. В результате при попадании в район облака значения яркости в видимом канале будут увеличиваться, а значения в ИК-канале (тепловая температура) будут повышаться.

Если для анализируемого кандидата в детекции уровень корреляции меньше -0,6, то такая детекция далее не рассматривается; а если больше, то используется подход, схематично представленный на рисунке 1.5.

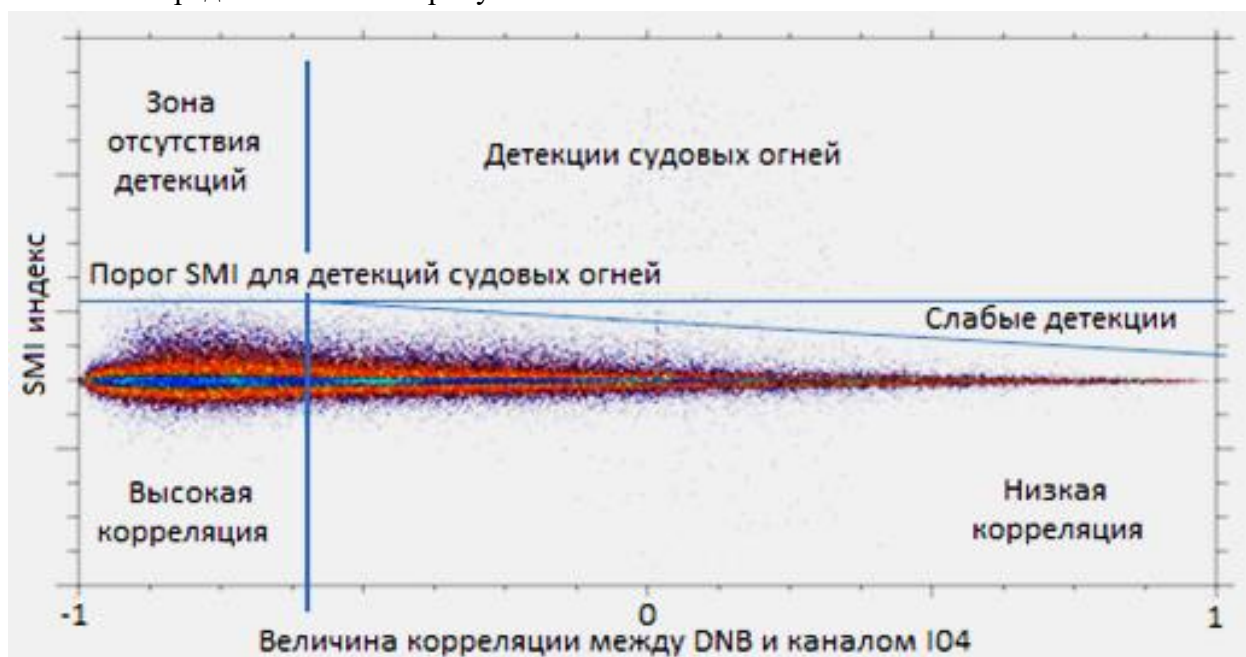


Рисунок 1.5 – Иллюстрация фильтрации лунных бликов от облаков

На рисунке по горизонтальной оси отложен уровень корреляции, по вертикальной - значения SMI индекса. Точки на рисунке обозначают кандидатов в детекции для некоторой одной зоны лунной засветки. Вертикальной линией отмечен порог корреляции -0,6, горизонтальной - порог SMI индекса для анализируемой зоны лунной засветки, наклонная линия соединяет точку пересечения двух указанных линий и точку с координатами (1; A), где значение A равно значению порога SMI индекса для анализируемой зоны лунной засветки, уменьшенному на четверть. Все детекции, попавшие ниже наклонной линии - не рассматриваются как кандидаты. Детекции, попавшие между наклонной и горизонтальной линией, попадают в результирующий

список, но как недостоверные (признак, присвоенный тем детекциям, значения SMI индекса которых ниже порога лунной засветки для соответствующих зон). Детекции, значения SMI индекса которых превышает уровень горизонтальной линии, продолжают рассматриваться в качестве полноценных детекций.

## **ГЛАВА 2. РЕАЛИЗАЦИЯ АЛГОРИТМА ДЕТЕКТИРОВАНИЯ СУДОВЫХ ОГНЕЙ И ЕГО ОПТИМИЗАЦИЯ**

### **2.1 Реализация алгоритма детектирования судовых огней**

Детектор судовых огней (VBD) реализован в виде программ на языках программирования Java, MATLAB и IDL, которые оркестрируются скриптом на языке программирования Python в соответствии с парадигмой MapReduce. Python-скрипт запускается автоматически каждый раз при поступлении новых спутниковых данных. Несколько экземпляров VBD могут работать на параллельном вычислительном кластере для ускорения обработки больших объемов данных.

Ниже перечислен набор входных файлов, необходимых для каждого запуска VBD детектора:

File_DNB	- видимый диапазон VIIRS (DNB) в формате HDF5 (* .h5)
File_GDNB	- файл с географической привязкой данных VIIRS DNB в формате HDF5 (* .h5)
File_LSM	- маска “вода-суша” в виде двоичной сетки с шагом 15 секунд сетка в формате ENVI
File_EEZ	- границы экономических зон - векторный файл в формате SHP (* .shp)
File_FLM	- название сетки маски факелов в формате ENVI
File_LI	- модель лунной освещенности для DNB изображений в формате ENVI
File_SVI	- инфракрасный диапазон VIIRS (SVI05), приведенный к проекции сенсора SVDNB и представленный в формате HDF5 (* .h5)
File_VNF	- CSV-таблица детекций инфракрасных источников VIIRS, полученных по алгоритму Nightfire (* .csv)
File_FMZ	- векторный файл зон рыбного промысла в формате SHP (* .shp)
File_MPA	- файл морских охраняемых районов в формате SHP (* .shp)
File_LTZ	- векторный файл локальных часовых поясов в формате SHP (* .shp)
File_RLP	- часто детектируемые (повторяющиеся) световые источники (NTL платформы) в векторном формате SHP (* .shp)



File\_RLV - файл, содержащий векторные полигоны границ для часто фиксируемых световых полей (например, мостов) в векторном формате SHP (\*.shp)

File\_TAI - таблица временной коррекции времени по атомным часам на борту спутника в международное время UTC (\*.txt)

File\_SAA - векторный файл, содержащий полигоны аномалий космических лучей (SAA) в формате SHP (\*.shp)

Результатом каждого запуска программы VBD для DNB изображения является CSV-таблица со списком детекций и KML-файл с картой их распределения. Пример карты детекций, построенной по данным, полученным в ночь 4 апреля 2016 года, показан на рисунке 2.1. Выходные файлы VBD для нескольких изображений DNB ежедневно объединяются в одну таблицу CSV и карту KML. Эти ежедневные сводные таблицы детекций служат в качестве входных данных для загрузки в базу данных детекций VIIRS, описанную в следующем подразделе.

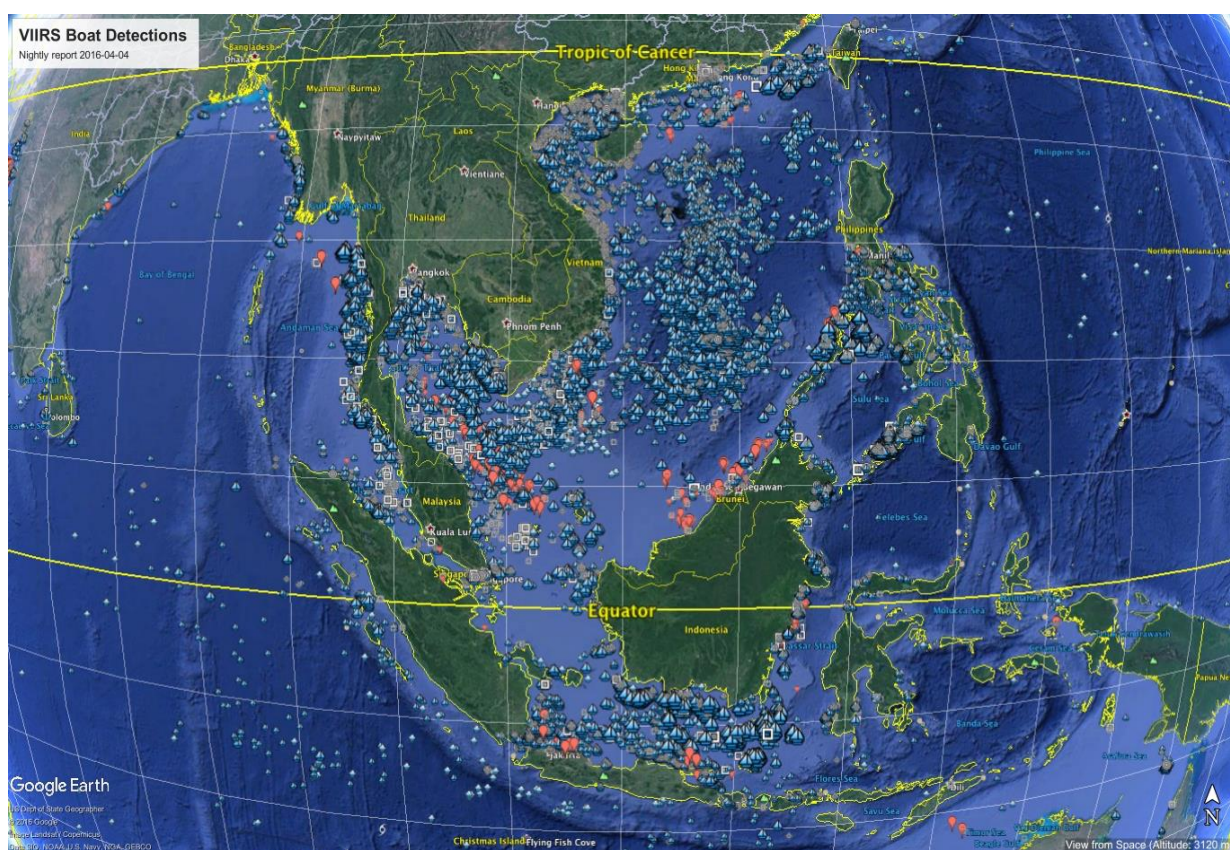


Рисунок 2.1 – Карта детекций VBD, построенная по данным, полученным в ночь 5 апреля 2016 года, в формате KML, визуализированная в Google Earth [116]

В таблице 2.1 представлен результат профилирования по времени программной реализации алгоритма детектирования судовых огней при обработке данных, полученных за 5 минут пролета спутника. Профилирование проводилось на рабочей станции, имеющей следующие характеристики:

- процессор: Intel Xeon E5640@2,67GHz (2 процессора);
- оперативная память: 8ГБ;
- разрядность операционной системы: 64 бит.

Таблица 2.1 – Результат профилирования программной реализации алгоритма детектирования судовых огней (наиболее времяемкие функции)

Логические блоки первого уровня	Логические блоки второго уровня	Логические блоки третьего уровня	Количес тво вызовов	Суммар ное время работы (сек.)	Процентное соотношение от времени родительского блока
Перепроекция из спутниковой проекции в картографическую			1	457	37%
Детектирование судовых огней			1	763	62%
	Вычисление Коэффици- ента размытия изображения		1	355	46%
		Двумерное преобразование Фурье	782073	128	37%
		Кросс- корреляция спектральных каналов	782073	132	39%
	Интерполя- ция данных shape-файлов		1	86	11%

В таблице перечислены только наиболее емкие по времени этапы, но даже этого достаточно для понимания, что общее время выполнения обработки на одной вычислительной машине соизмеримо (и даже превышает) время получения данных. Чтобы сократить общее время выполнения программы, можно использовать

вычислительный кластер, распараллеливая обработку данных, а можно оптимизировать код программы.

## 2.2 Оптимизация алгоритма детектирования ночных огней

Для обработки потока мультиспектральных спутниковых данных, получаемых с сенсора VIIRS, без накопления задержки, можно использовать вычислительный кластер с большим объемом оперативной памяти и дискового пространства. Но даже это не сильно помогает. Так, например, на вычислительном кластере, состоящем из 12 четырехпроцессорных узлов, скорость обработки данных VIIRS лишь в три раза превышает скорость поступления новых. Это приемлемо для обработки текущих данных, но провести архивную обработку уже не представляется возможным.

Чтобы ускорить обработку спутниковых данных, была проведена оптимизация наиболее ресурсоемких среди алгоритмов, наиболее часто используемых для обработки мультиспектральных спутниковых данных, получаемых с ночной стороны Земли. В первую очередь сюда вошли: двумерное преобразование Фурье и использующие его алгоритмы (например, алгоритм вычисления коэффициента размытия изображения), а также кросс-корреляция наборов данных в скользящем окне (таблица 2.1).

В первую очередь реализация целевых алгоритмов была перенесена с языка программирования MATLAB на язык программирования C++, а во вторую – полученная реализация была дополнительно оптимизирована для эффективной работы на мультиядерных процессорах семейства Intel, в частности: уменьшено количество переходов, увеличена степень векторизации циклов, проведено выравнивание массивов, внедрены готовые программные примитивы для Intel платформ, использованы библиотеки OpenMP, использована оптимизация компиляторов Intel.

В таблицах 2.2-2.4 приведены результаты сравнения производительности реализации алгоритма вычисления коэффициента размытия изображения на языке C++ с оптимизацией и без оптимизации. В случае оптимизированной реализации рассматривались два варианта:

1. Кол-во OpenMP потоков равно количеству физических ядер на узле.
2. Кол-во OpenMP потоков равно количеству виртуальных потоков на узле с учетом hyperthreading.

Таблица 2.2 – Сравнение производительности оптимизированной и неоптимизированной реализации алгоритма вычисления коэффициента размытия изображения на языке программирования C++. Процессор (Intel(R) Xeon Phi(TM) CPU 7250, 68 ядер)

Неоптимизированная	Оптимизированная	
4мин 28 с.	68 потоков	272 потока
	15,588 с.	22,044 с.



Таблица 2.3 – Сравнение производительности оптимизированной и неоптимизированной реализации алгоритма вычисления коэффициента размытия изображения на языке программирования C++. Процессор (Intel(R) Xeon(R) CPU E5-2697 v3, 2 процессора по 14 ядер)

Неоптимизированная	Оптимизированная	
51,389 с.	28 потоков	56 потока
	5,003 с.	4,418 с.

Таблица 2.4 – Сравнение производительности оптимизированной и неоптимизированной реализации алгоритма вычисления коэффициента размытия изображения на языке программирования C++. Процессор (Intel(R) Xeon(R) CPU E5-2698 v4, 20 ядер на процессор, 2 процессора на узле)

Неоптимизированная	Оптимизированная	
48,832 с.	40 потоков	80 потоков
	4,563 с.	4,310 с.

Как видно из приведенных таблиц, оптимизированная реализация алгоритма позволила существенно повысить производительность и сократить общее время работы.

### 2.3 Оптимизация алгоритма межканальной корреляции в скользящем окне с учетом многоядерной архитектуры

#### Базовая версия алгоритма

Из таблицы 2.1 видно, что наиболее ресурсоемким алгоритмом в составе кода является алгоритм кросс-корреляции спектральных каналов, используемый для вычисления коэффициента размытия изображения. Чтобы ускорить работу кода, мы этот алгоритм оптимизировали отдельно.

Вычисление корреляции в скользящем окне является очень распространенным методом статистической оценки взаимосвязанности двух наборов данных, используемым в обработке изображений (Bastiaans, 1985), анализе сигналов (Burkhauser et al., 2000), медицине (Campello et al., 2013), геофизике (Cho et al., 1999) и т.д. Функции вычисления параметров в скользящем окне (сумма, среднее арифметическое, корреляция и др.) входят во многие программные пакеты (например, MATLAB, R, S-PLUS) и активно используются при анализе временных рядов в финансовом анализе (Croft, 1978).

В настоящей работе мы рассматриваем линейную корреляцию Пирсона, вычисляемую в скользящем окне между наборами пространственных данных на регулярной сетке, преимущественно между двумя изображениями. Прикладная задача относится к автоматическому поиску судовых огней на мультиспектральных спутниковых данных, полученных в ночное время суток (Croft and Colvocoresses 1979). Задача заключается в фильтрации ложных детекций, вызванных отражением лунного света от облаков, путем вычисления корреляции в скользящем окне между данными в видимом и инфракрасном каналах. При наличии облачности коэффициент корреляции должен быть близок к -1, так как значения в видимом канале увеличиваются из-за отражения лунного света от облаков, в то время как значения в инфракрасном канале уменьшаются вследствие пониженной температуры облаков по сравнению с поверхностью Земли.

Базовый алгоритм может быть представлен в виде синхронного прохождения двух наборов данных скользящим окном заданного размера с заданным шагом, и вычисления линейной корреляции Пирсона между значениями исследуемых наборов в этом окне. При этом значения в n-мерных скользящих окнах вытягиваются в одномерные векторы.

Для вычисления корреляции в каждом окне используется классическая формула линейной корреляции Пирсона:

$$c = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}, \text{ где:}$$

$c$  – коэффициент корреляции Пирсона;

$x_i$  и  $y_i$  – соответствующие значения двух дискретных рядов, для которых вычисляется корреляция;

$\bar{x}$  и  $\bar{y}$  – арифметическое среднее рядов  $x$  и  $y$  соответственно;

суммирование ведется от 1 до  $n$  – по всем членам рядов  $x$  и  $y$ .

Вычисление корреляции в скользящем окне с использованием классической формулы приводит к существенным накладным расходам, что проиллюстрировано на примере обработки двумерных изображений, представленном на рисунке 2.2.

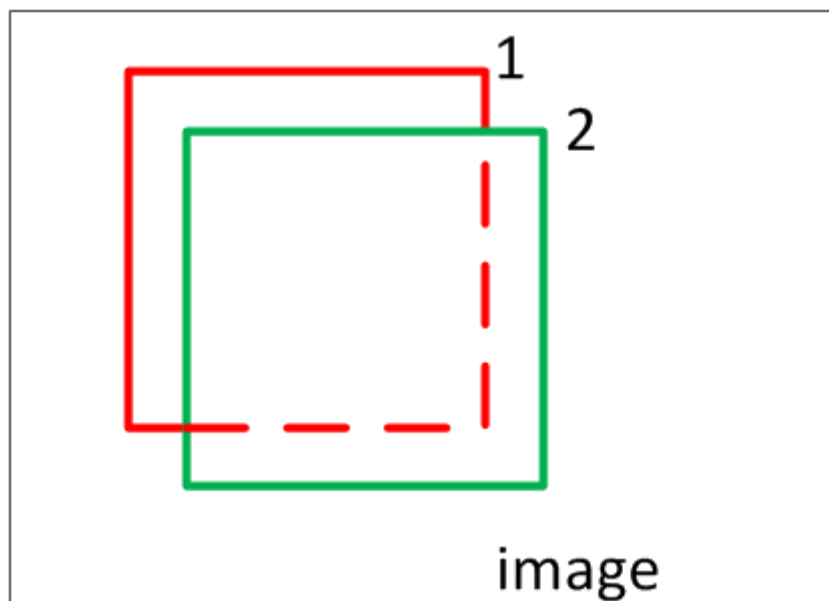


Рисунок 2.2 – Вычисление корреляции в скользящем окне по классической формуле на примере двумерных изображений

Окна 1 и 2 имеют большое перекрытие. При расчете корреляции в каждом окне, элементы, которые принадлежат обоим окнам, будут обработаны дважды. Если шаг скользящего окна достаточно мал, например, один пиксель, а сам размер окна велик, то каждый элемент изображения будет обработан столько раз, сколько окон его накроет.

В данной работе мы оптимизировали базовую версию алгоритма межканальной корреляции в скользящем окне с учетом многоядерной архитектуры и провели тестирование оптимизированной версии на графических процессорах.

### Оптимизированная версия алгоритма

Чтобы уменьшить накладные расходы, вызванные многократной обработкой элементов, мы использовали модифицированную формулу расчета корреляции, которую часто используют, чтобы сократить количество вычислительных операций, связанных с вычитанием из каждого значения среднего арифметического:

$$c = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}} \quad (2.1)$$

Выгодным отличием данной формулы для исследуемой в данной работе задачи является выражение значения корреляции через суммы величин  $x$ ,  $y$ ,  $xy$ ,  $x^2$ ,  $y^2$ . Такие суммы для одномерных рядов могут быть получены линейным алгоритмом вычисления суммы в скользящем окне *movesum* (аналогично скользящему среднему *moveaverage*). Чтобы перейти от одномерного ряда к двумерной матрице изображения, мы последовательно применяем "скользящее суммирование" сначала по одному измерению, а потом по второму (Рисунок 2.3).

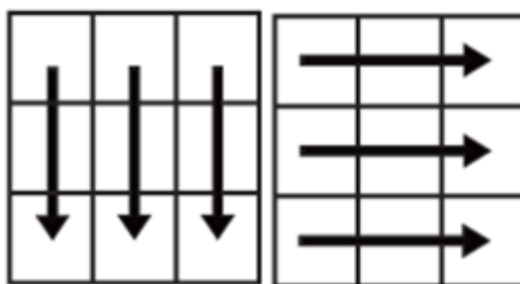


Рисунок 2.3 – Схема вычисления суммы в скользящем окне для двумерного случая

В результате суперпозиции двух одномерных суммирований с окном длины  $k$  получается матрица такого же размера, как исходная, но на каждой позиции которой будет стоять сумма элементов этой матрицы, взятых в окне размером  $k \times k$  с центром в данной точке (при условии, что одномерное суммирование также помещает результат в центр окна).

Таким образом, общий алгоритм выглядит следующим образом:

1. Поэлементным умножением матриц  $x$  и  $y$  вычисляются значения матриц  $xy$ ,  $x^2$ ,  $y^2$ .
2. Для каждой матрицы  $x$ ,  $y$ ,  $xy$ ,  $x^2$ ,  $y^2$  производится суммирование со скользящим окном длины  $k$  по всем строкам (результаты обозначим как  $Ax$ ,  $Ay$ ,  $Axy$ ,  $Ax^2$ ,  $Ay^2$ ).
3. Для каждой матрицы  $Ax$ ,  $Ay$ ,  $Axy$ ,  $Ax^2$ ,  $Ay^2$  производится суммирование со скользящим окном длины  $k$  по всем столбцам, результаты обозначим как  $Sx$ ,  $Sy$ ,  $Sxy$ ,  $Sx^2$ ,  $Sy^2$  (матрицы сумм в окнах с центром в данном элементе).
4. Для каждого элемента результирующей матрицы вычисляем корреляцию по формуле 2.1.

Чтобы обобщить приведенный алгоритм на  $n$ -мерный случай, надо последовательно повторить суммирование со скользящим окном не два раза (шаги 2-3), а  $n$  раз по одному для каждого измерения.

Стоит отметить, что оптимизацию можно делать не только с помощью скользящего окна, но и с помощью кумулятивных сумм (Delmas and Gimel'farb). Мы сделали такую реализацию, но для решаемой прикладной задачи, описанной во введении, экспериментальные исследования показали проигрыш по времени выполнения данной реализации по сравнению с реализацией алгоритма, предложенного в данном разделе. Теоретические выкладки (которые мы не приводим здесь, чтобы сэкономить место) также показывают проигрыш по количеству операций при использовании кумулятивных сумм по сравнению с суммированием в скользящем окне.

### Распараллеливание оптимизированной версии алгоритма

Описанный в предыдущем подразделе алгоритм может быть эффективно распараллелен на случай многоядерной архитектуры, например GPU:

1. На шаге 1 каждый поток обрабатывает элементы из небольшой окрестности исходных данных.
2. На шаге 2 каждый поток обрабатывает одну или более строк в зависимости от размера данных и степени параллелизма.
3. На шаге 3 каждый поток обрабатывает один столбец или более в зависимости от размера данных и степени параллелизма.
4. На шаге 4 каждый поток обрабатывает элементы из небольшой окрестности исходных данных.

Для n-мерного случая распараллеливание производится аналогично.

### **Реализация алгоритмов и тестирование**

Описанные в предыдущих подразделах алгоритмы были реализованы в программном обеспечении. Базовая и оптимизированная версии были реализованы на языке программирования MATLAB. Распараллеленная оптимизированная версия реализована на языке программирования C++ под библиотеки CUDA. В программной реализации вычисляется корреляция в скользящем окне между двумя двумерными наборами данных. Шаг окна - одна ячейка. Оба набора данных должны иметь одинаковый размер. Результирующий набор имеет такой же размер, как аргументы. Для каждой пары окон, результат корреляции записывается в среднюю ячейку. Данные могут содержать пропущенные значения, которые идентифицируются числом меньшим или равным -999. Если в коррелируемых окнах встречается хотя бы одно пропущенное значение, то результирующее значение для двух данных окон устанавливается равным -2. Значения по краям результирующей матрицы, для которых соответствующие им корреляционные окна выходят за границы матриц-аргументов, устанавливаются равными -2.

Тестирование и сравнение программных реализаций производилось на двух спутниковых изображениях, перепроецированных на регулярную географическую сетку. В частности, мы использовали изображения в видимом и инфракрасном каналах, полученных сенсором VIIRS, установленном на спутнике Suomi NPP. Размер изображений равен 4064x3072 пикселей. Размер окна корреляции равен 7x7 пикселей. Тестирование MATLAB версии проводилось на рабочей станции со следующими характеристиками:

- процессор: 2 x Intel Xeon E5640 (8 ядер, 2,66 ГГц)
- объем оперативной памяти: 8 ГБ
- операционная система: Windows 7

Тестирование GPU версии проводилось на вычислительной машине со следующими характеристиками:

- процессор: 2 x Intel Xeon E5-2650 v2 (8 ядер, 2,60 ГГц)
- объем оперативной памяти: 128 ГБ
- графическая карта: 2 карты NVIDIA Tesla K80
- операционная система: Ubuntu 14.04.3

Тестирование показало следующие результаты:

- Базовая версия: около 42 секунд, из них около 1 секунды - чтение данных и порядка 41 секунды - работа самого алгоритма корреляции.
- Оптимизированная версия: около 3,3 секунды, из них около 1 секунды – чтение данных и порядка 2,3 секунды - работа самого алгоритма корреляции.
- Распараллеленная оптимизированная версия: около 13 секунд, из них: почти 11 секунд - инициализация GPU, 1,5 секунды - чтение данных и порядка 0,6 секунды - работа самого алгоритма, включая выделение памяти на GPU и трансфер данных на устройство и с него.

Таким образом, оптимизация ускорила работу алгоритма для тестового примера более чем в 15 раз, а распараллеливание и перенос на GPU ускорили его еще примерно в 4 раза.

Если произвести теоретический расчет количества арифметических операций при обработке двумерных данных, то получаем:

- Для базовой версии:  $[2*n*n \{ \text{вычисление двух средних арифметических} \} + 2*n*n \{ \text{вычитание из каждого элемента среднего арифметического} \} + n*n \{ \text{перемножение в числителе} \} + n*n \{ \text{сложение в числителе} \} + 2*n*n \{ \text{возведение элементов во вторую степень в знаменателе} \} + 2*n*n \{ \text{два суммирования в знаменателе} \} + 2 \{ \text{вычисление квадратного корня в знаменателе и деление} \}] * (A-n+1)*(B-n+1) \{ \text{количество окон} \} \approx 10*n^2*A*B$ , где  $n$ -размер одного измерения окна,  $A$  и  $B$  - размеры обрабатываемого двумерного набора данных.

- Для оптимизированной версии:  $3*A*B \{ \text{поэлементное вычисление значений } x, x^2, y^2 \} + 5*SM + 3*A*B \{ \text{два поэлементных произведения и вычитание} \} + 2*3*A*B \{ \text{два вычисления подкоренных выражения, каждое из которых включает два поэлементных умножения и поэлементную разность} \} + 3*AB \{ \text{два поэлементных вычисления квадратного корня} \} + A*B \{ \text{поэлементное деление} \} = 16*A*B + 5*SM$ , где  $A$  и  $B$  - размеры обрабатываемого двумерного набора данных, а  $SM$  - количество операций, необходимых для вычисления двумерного суммирования со скользящим окном.  $SM = A \{ \text{количество строк} \} * [ B \{ \text{количество суммирований элементов в одной строке} \} + (B-n) \{ \text{количество вычитаний элементов в одной строке} \} ] + B \{ \text{количество столбцов} \} * [ A \{ \text{количество суммирований элементов в одном столбце} \} + (A-n) \{ \text{количество вычитаний элементов в одной строке} \} ] = A*[B+(B-n)] + B*[A+(A-n)] \approx 4*A*B$ . Итого, для оптимизированной версии алгоритма получаем  $16*A*B + 5*SM = 36*A*B$  арифметических операций.

Конечно, приведенные выкладки не могут точно отражать временное соотношение, так как не все операции выполняются одинаковое время, а также потому что мы не учитывали компьютерные операции, например, присвоение значений, но тем не менее данные формулы можно использовать для приблизительного сравнения сложности алгоритмов. Если подставить в приведенные формулы значения, используемые в эксперименте, то получим:

- для базовой версии алгоритма:  $10*n^2*A*B = \{n=7\} = 490*A*B$ ;

— для оптимизированной версии алгоритма:  $36 \cdot A \cdot B$ .

И их отношение равно  $490/36 \approx 13,5$ , что довольно близко к отношению времени выполнения этих алгоритмов, полученное при тестировании.

Анализируя полученные формулы можно сделать вывод: количество арифметических операций в базовой версии алгоритма квадратично зависит от размера окна, в то время как в оптимизированном алгоритме количество арифметических операций вообще не зависит от размера окна.

Провести аналогичные теоретические подсчеты для распараллеленной версии можно, но они будут сильно расходиться с практической реализацией по следующим причинам:

1. Мы не можем в теоретических расчетах учесть накладные расходы, связанные с инициализацией и поддержкой одного вычислительного потока. Более того, различные вычислительные архитектуры будут требовать различных накладных расходов.

2. Нам сложно учитывать степень параллелизма в конкретной архитектуре и особенности конкретной реализации. Так, например, не все параллельные потоки могут выполняться параллельно. Например, в случае GPU конкретное разбиение пространства данных на блоки может сильно влиять на производительность.

3. Сложно учитывать затраты на обмен данными. Например, опять же в случае с GPU компилятор может оптимизировать программу, размещая данные на регистрах или в локальной памяти. А это приводит к увеличению или уменьшению времени работы программы.

Еще один момент, который следует затронуть по результатам эксперимента - точность. Наиболее критично этот аспект проявляется в тех случаях, когда результирующее значение корреляции должно дать некорректный результат, например, в случае, когда все значения в одном из окон имеют одинаковое значение (в этом случае, согласно формуле, возникает деление на ноль). Но в случае даже небольших погрешностей, деление на ноль может обернуться делением на очень маленькие значения, что приведет не к ошибке, а к некоторому вполне определенному значению. Разумеется, данная проблема актуальна только в тех случаях, когда рассматриваются числа с плавающей точкой. Как показали результаты тестирования, базовый алгоритм дает наиболее точные результаты в случае экстремумов (деление на ноль), в то время как оптимизированный алгоритм и его параллельная версия часто выдают некорректный результат. Это связано в первую очередь с самой оптимизированной формулой: как показали эксперименты два суммирования с перемножением с последующим нахождением разницы, производимые в числителе оптимизированной формулы, дают большую погрешность, чем нахождение среднего арифметического с последующим его вычитанием из каждого элемента и перемножением, производимые в классической формуле. Во вторую очередь, это связано с накоплением ошибки при вычислении суммы в скользящем окне.

Тем не менее, если не учитывать экстремальные значения, то погрешность алгоритмов достаточно мала: если в случае проведенного эксперимента взять результаты базовой версии в качестве эталонного результата, то для оптимизированной версии максимальное поэлементное отклонение составило менее 0.001 по абсолютному значению.

## **Выводы**

Разработанная оптимизированная версия алгоритма вычисления межканальной корреляции в скользящем окне позволила сократить время выполнения задачи фильтрации ложных детекций судовых огней, вызванных отражением лунного света от облаков, более чем в 15 раз для вычислительной архитектуры с центральным процессором и еще в 2 раза для CUDA GPU (время инициализации GPU не берем в расчет, так как инициализация может производиться заранее, при старте программы). Разработанный алгоритм и его реализация могут быть использованы не только для обработки спутниковых изображений, но и в других областях.

## **ГЛАВА 3. ПРОСТРАНСТВЕННО-ВРЕМЕННАЯ БАЗА ДАННЫХ**

Ежедневные выходные файлы VBD в формате CSV могут быть успешно использованы для операционных служб, таких как система предупреждения о незаконном промысле. Однако этот выходной формат является неподходящим для организации архива детекций судов, основанного на нескольких годах непрерывных спутниковых наблюдений. Чтобы ускорить поиск по географическим координатам и статистический анализ в долгосрочном архиве судов, мы разработали специальную базу данных на основе СУБД с открытым исходным кодом PostgreSQL и ее пространственного расширения PostGIS.

База данных имеет всего одну таблицу, которая называется VBD. Таблица VBD содержит все те же самые атрибуты (включая их названия и типы), которые содержит выходной CSV-файл. База данных проиндексирована по координатам детекций при помощи пространственного индекса VBD\_GIX. Ниже приведен фрагмент кода SQL для создания и индексации отношения VBD.

```
CREATE TABLE IF NOT EXISTS vbd (  
  PGID SERIAL PRIMARY KEY,  
  id INTEGER,  
  id_Key TEXT UNIQUE,  
  Date_Proc TIMESTAMP,  
  Lat_DNB DOUBLE PRECISION,  
  Lon_DNB DOUBLE PRECISION,  
  Rad_DNB DOUBLE PRECISION,  
  Date_Mscan TIMESTAMP,  
  ...
```



```
LUNA_GDNBO DOUBLE PRECISION);
SELECT AddGeometryColumn('public', 'vbd', 'geom', 4326, 'POINT', 2);
CREATE INDEX vbd_gix ON vbd USING GIST(geom);
```

Например, чтобы сделать базу данных VBD детекция за 2016 год по Южной Азии, мы использовали все ночные DNB изображения, охватывающие прямоугольную область интереса (ROI) с координатами северо-западного угла 50 с.ш., 90 в.д. и координатами юго-восточного угла 15 с.ш., 165 в.д. ROI охватывает такие страны как Индонезия, Бруней, Филиппины, Малайзия, Вьетнам, Таиланд, Камбоджа, Мьянма и Тимор-Лешти. Обработка всех спутниковых данных за 2016-й год по выбранному региону при помощи алгоритма VBD заняла несколько часов на серверной вычислительной машине под управлением ОС Linux. Загрузка в базу данных результатов обнаружения лодок в 2016 году в Южной Азии занимают несколько часов на серверном компьютере Linux. Общее количество записей в базе данных (детекций) составило 5 288 058. Размер базы данных составил 6,9 ГБ, включая размер таблицы VBD, равный 5,9 ГБ, и размер пространственного индекса, равный 1 ГБ. Распределение количества детекций по дням приведено на графике на рисунке 3.1 синей линией. Также на графике красной линией показана фаза лунного цикла. Видно, что ежедневное количество детекций сильно зависит от фазы луны. В лунные ночи количество детекций уменьшается по причине зашумления изображений бликами от воды и отражением от облаков. В результате, чтобы подавить шумы, приходится понижать порог чувствительности алгоритма, что ведет к уменьшению обнаруженных детекций. Максимальное число детекций составило 33 613 и было зафиксировано в новолуние 11 мая 2016 года. Минимальное количество (240) детекций было зафиксировано 11 апреля 2016 года из-за засветки луной верхних облаков, создающей помехи для использования VBD в исследуемой области.

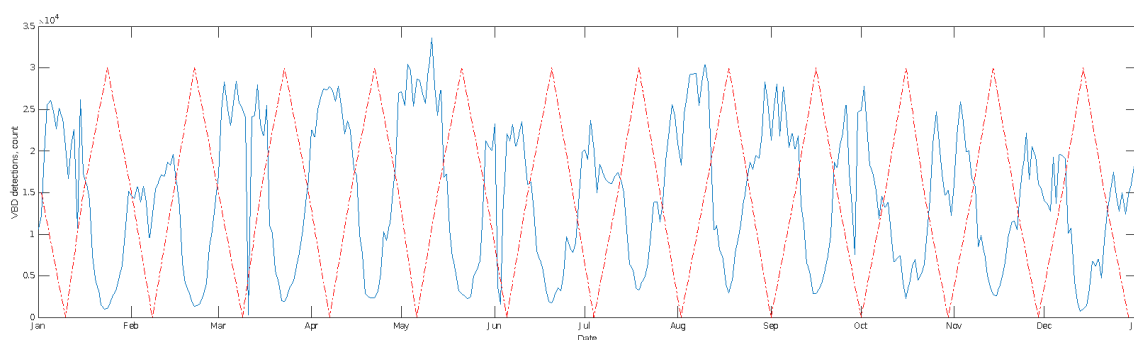


Рисунок 3.1 – Число детекций в день (сплошная синяя линия) и фаза луны (красная пунктирная линия) за 2016 год в базе данных VBDDB\_v22

### *Примеры SQL-запросов к базе данных детекций судов*

Ниже приведен пример запроса к базе данных для выбора записей о детекциях судов, полученных в новолуние ( $LI < 0,0001$ ), и имеющих значение флага QF (см. главу

1), равное 1,2,3,4,8 или 11 (что соответствует судам или часто фиксируемым источникам) внутри интересующей прямоугольной области (ROI) с углами NW и SE равными (-3 с.ш., 106 в.д.) и (-6 с.ш., 109 в.д.) соответственно (на 1 градус к северу от Джакарты):

```
SELECT id_Key, Date_Proc, Lat_DNB, Lon_DNB, Rad_DNB, Date_Mscan,
Date_LTZ, Line_DNB, Sample_DNB, SMI, Thr_SMI, SI, Thr_SI, SHI, Thr_SHI, LI, Thr_LI,
Glint, Thr_Glint, Xcorr, QF_Detect, QF_Bitflag, Land_Mask, EEZ, FMZ, MPA, File_DNB
FROM vbd WHERE ST_Intersects(ST_SetSRID(ST_MakeBox2D(ST_MakePoint(107, -
6),ST_MakePoint(106, -5)),4326), geom) AND (Qf_Detect = 1 OR Qf_Detect = 2 OR
Qf_Detect = 3 OR QF_Detect = 4 OR QF_Detect = 8 OR QF_Detect = 11) AND (LI < 0.0001)
ORDER BY Date_Mscan;
```

Выполнение запроса занимает 5 секунд на четырехъядерном процессоре Intel i7 2,5 ГГц с 16 ГБ оперативной памяти. Результат запроса включает 14 542 детекций судов и морских платформ, отображенных на карте на рисунке 3.2.

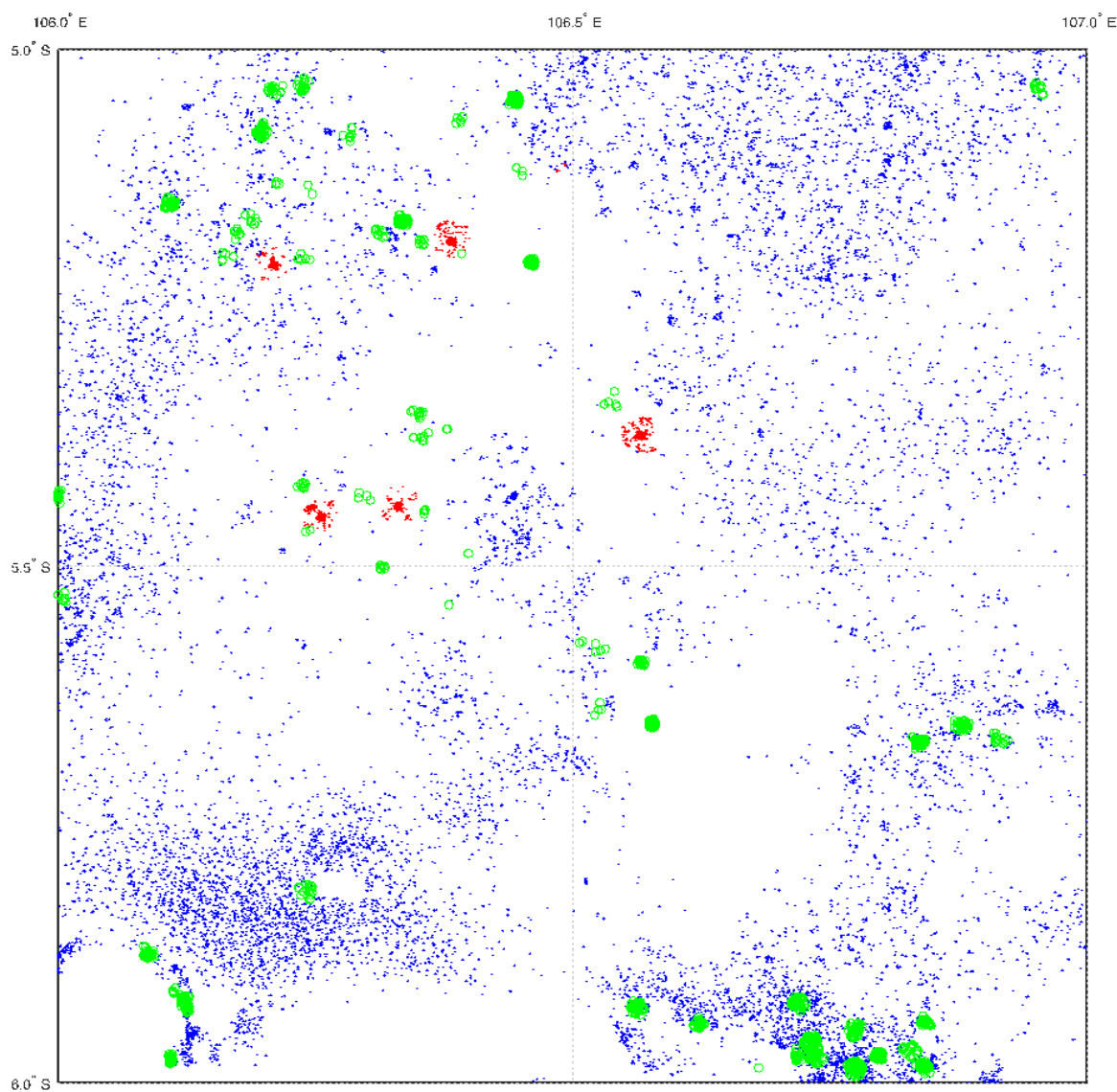


Рисунок 3.2 – Карта выборки детекций судов из базы данных в квадрате 1x1 градус к северу от Джакарты в 2016 году. Синий цвет соответствует детекциям судов, полученным в новолуние; зеленый цвет соответствует морским платформам; красный цвет соответствует газовым факелам

Второй пример запроса к базе данных - поиск судов в пределах окружности с радиусом 1 градус с центром в точке с координатами 15.5 с.ш., 114 в.д. в Южно-Китайском море. Запрос занимает 8 секунд и дает 19 229 детекций судов и морских платформ за 2016 год, представленных на рисунке 3.3.

```
SELECT id_Key, Date_Proc, Lat_DNB, Lon_DNB, Rad_DNB, Date_Mscan,
Date_LTZ, Line_DNB, Sample_DNB, SMI, Thr_SMI, SI, Thr_SI, SHI, Thr_SHI, LI, Thr_LI,
Glint, Thr_Glint, Xcorr, QF_Detect, QF_Bitflag, Land_Mask, EEZ, FMZ, MPA, File_DNB,
ST_Distance_Sphere(ST_Centroid(geom), ST_SetSRID(ST_MakePoint(114,15.5), 4326)) AS
Dist FROM vbd WHERE ST_DWithin(geom, ST_SetSRID(ST_MakePoint(114, 15.5), 4326),
1) ORDER BY Date_Mscan;
```

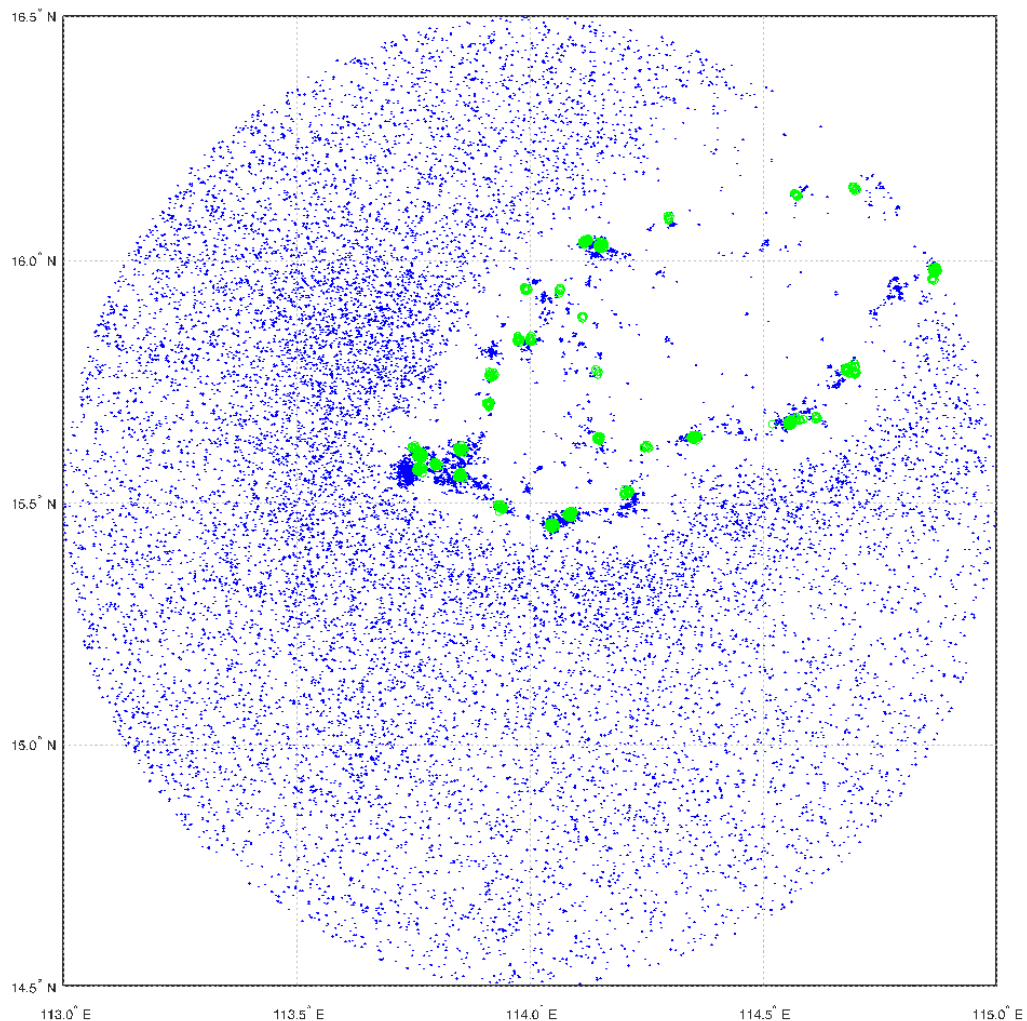


Рисунок 3.3 – Карта выборки детекций судов и морских платформ в окружности с радиусом 1 градус и центром в точке 15,5 с.ш., 114 в.д. Синий цвет соответствует детекциям судов, зеленый - детекциям морских платформ

## ГЛАВА 4. ВЕБ-ПРИЛОЖЕНИЕ ДЛЯ ВИЗУАЛИЗАЦИИ И СТАТИСТИЧЕСКОГО АНАЛИЗА РАСПРЕДЕЛЕНИЯ СУДОВЫХ ОГНЕЙ

### 4.1 Tile-сервер для отображения плотных облаков точек VBD детекций

Для понимания распределения вероятности обнаружения судов в заданном географическом регионе необходима интерактивная визуализация большого количества VBD детекций, полученных за длительный период времени. Чтобы реализовать данную возможность, мы разработали веб-приложение для визуализации и статистического анализа распределения судовых огней. В качестве входных данных для работы приложения используется результат SQL-запросов к базе данных VBD детекций, описанной в главе 3. Разработанное приложение позволяет создавать интерактивные карты детекций при помощи специально разработанного tile-сервера, а также выполнять интерактивный статистический анализ при помощи программного инструмента Jupyter Notebook. Пользователи могут создавать и редактировать скрипты, реализующие построение графиков и написанные на языке программирования Python, прямо на веб-странице и запускать их на веб-сервере Jupyter.

В качестве подложки мы используем векторную карту границы “море-суша” и карту государственных границ из проекта OpenStreetMap. Облако точек VBD детекций отображается как дополнительный полупрозрачный слой, накладываемый поверх базового слоя, полученного с сервера OpenStreetMap. Чтобы поддерживать возможность интерактивного масштабирования и перемещения по карте для слоя VBD детекций, мы переводим облако точек детекций в набор растровых тайлов, имеющих такой же размер и географическую привязку, как и тайлы, из OpenStreetMaps.

В качестве базового программного приложения для создания VBD тайл-сервера мы использовали проект с открытым исходным кодом, разработанный Майклом Фоглеманом для визуализации данных о такси в Нью-Йорке, и поддерживающий работу с данными, состоящими из 77 миллионов точек. Тайл-сервер реализован на языке программирования Go, и имеет хранилище типа “ключ-значение”, реализованный на базе NoSQL СУБД Cassandra. Чтобы ускорить просмотр карт и избежать повторения запросов к базе данных тайл-сервера, тайлы кэшируются на стороне сервера. Исходный проект Майкла Фоглемана был модифицирован и адаптирован под структуру данных VBD, а также для обеспечения возможности определения местоположения курсора и внедрения средства измерения расстояния на интерактивной веб-карте<sup>1</sup>.

Карта распределения VBD детекций в Юго-Восточной Азии, построенная по данным за 2016 год, приведена на рисунке 4.1. Наибольшая плотность детекций

---

<sup>1</sup> <https://github.com/ljagis/leaflet-measure>

наблюдалась вблизи Таиланда, Вьетнама и Индонезии. При увеличении карты в облаке точек VBD детекций можно заметить кластерные структуры, концентрирующиеся в районах часто детектируемых световых источников и в районах рыбной ловли в ночное время суток.

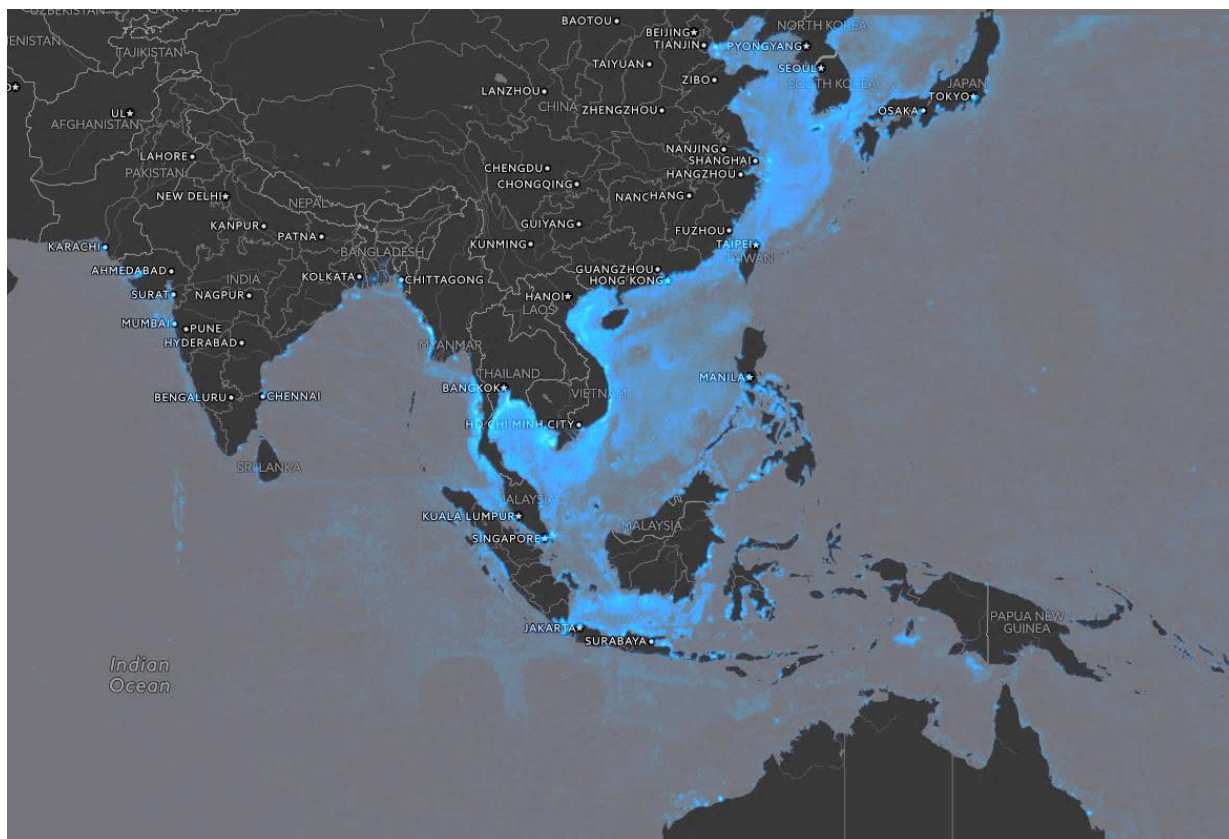


Рисунок 4.1 – Карта распределения VBD-детекций в Юго-Восточной Азии на тайл-сервере. В качестве фона используются карта границы “суша-море” и государственных границ OpenStreetMap

Например, к северу от Джакарты можно видеть набор часто детектируемых источников (кластеров детекций VDB) прямоугольной формы (рис. 4.2). Размер прямоугольника составляет 750-800 м, что указывает на то, что изолированный источник света часто попадает в один квадратный пиксель изображения DNB такого же размера. Используя снимки с высоким разрешением (0,5 м) от Terraserver (Terraserver.com, 2018), мы можем подтвердить это наблюдение для кластера, расположенного в точке с координатами -5.974632 с.ш. и 106.798782 в.д. Большое судно размещается в середине кластера VBD детекций по крайней мере в течение последних 5 лет (рисунок 4.2). Проверка с помощью снимков Terraserver нескольких соседних прямоугольных скоплений указывает, что большинство из них являются морскими нефтяными платформами.



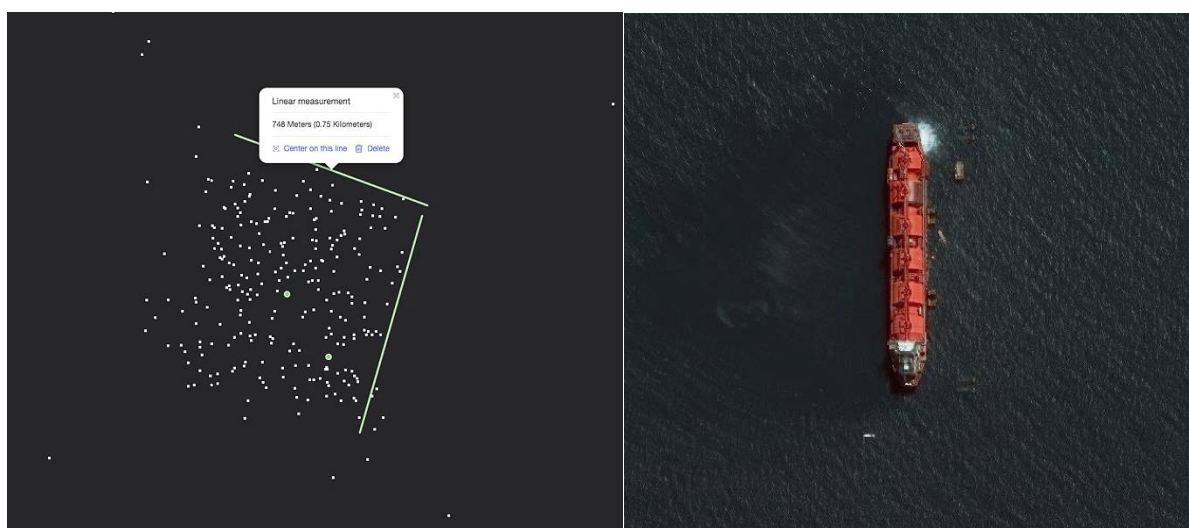
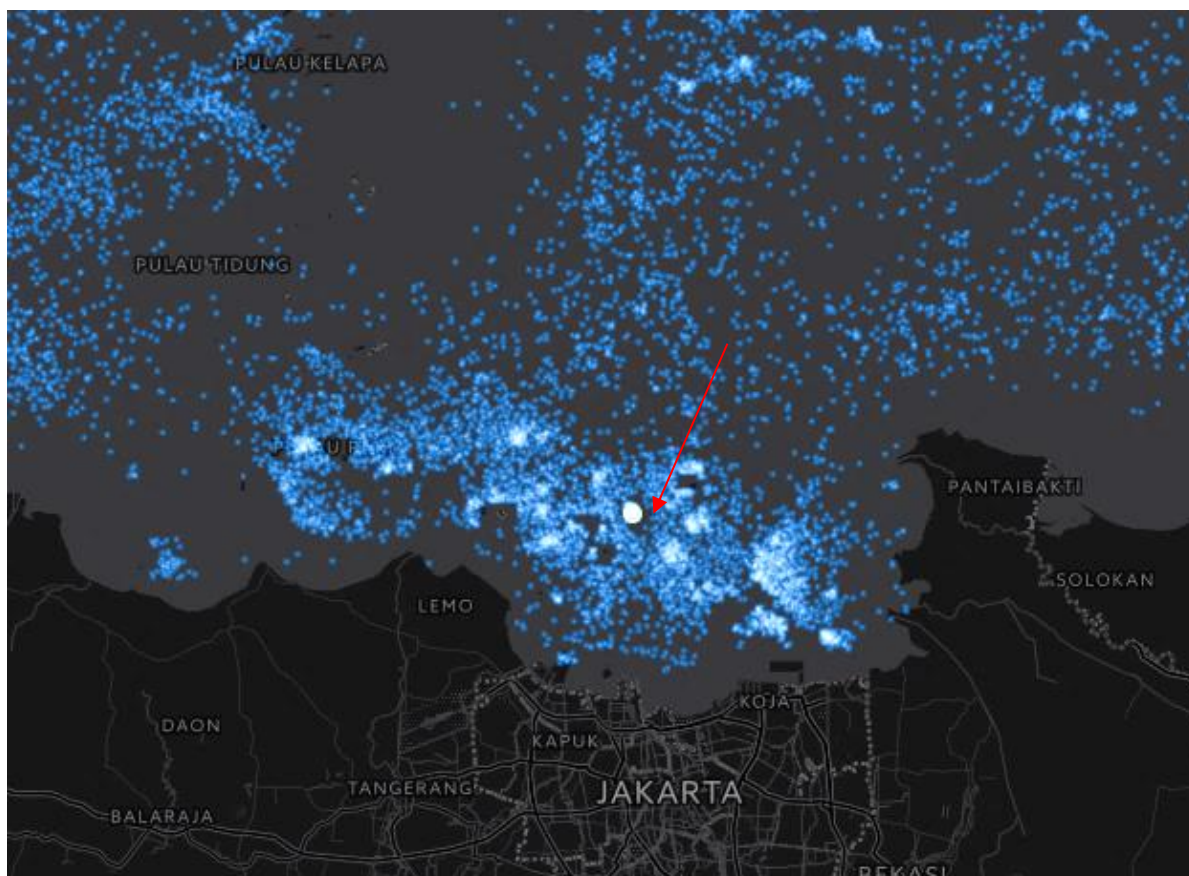


Рисунок 4.2 – Кластеры VDB детекций прямоугольной формы с центром в точке - 5,974632 с.ш. и 106,798782 в.д., а также соответствующее изображение с разрешением 0,5 м, взятое в 2017-05-31

Совершенно уникальная кластерная структура была обнаружена при помощи разработанного тайл-сервера в Сиамском заливе. Она выглядит как регулярная сетка, покрывающая сотни квадратных километров с регулярным шагом в обоих направлениях, равным 1 морской миле (рис. 4.3). Вероятным объяснением этой структуры является то, что в узлах этой сетки расположены рыболовные приманки, но это утверждение требует дополнительной проверки.

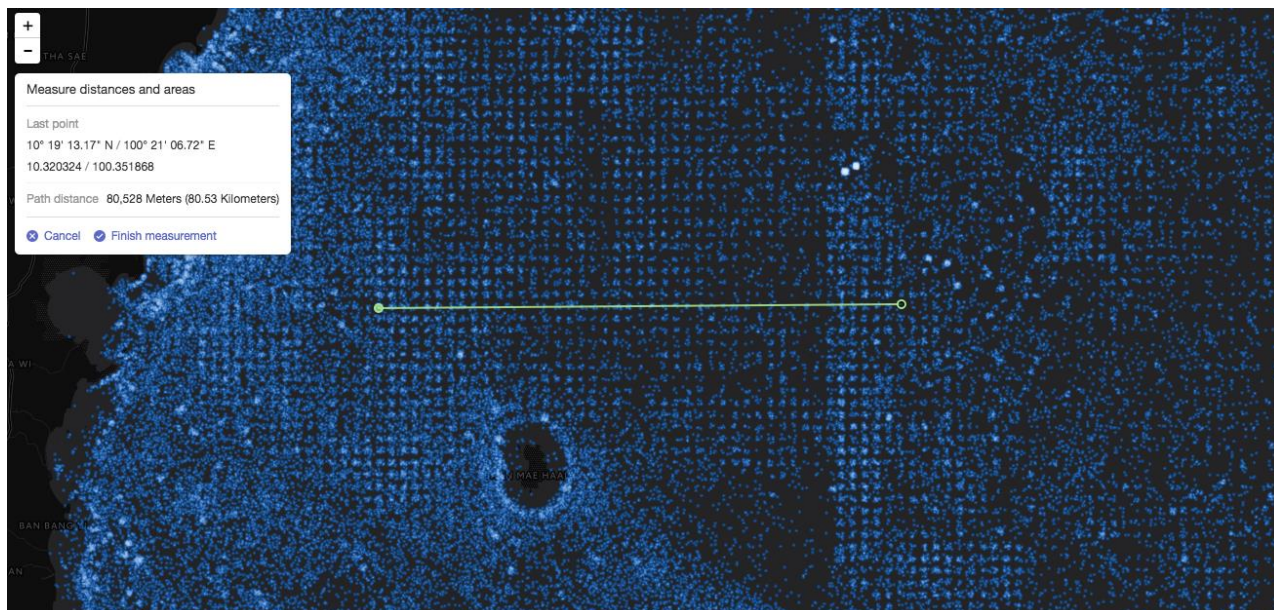


Рисунок 4.3 – Регулярная сетка, покрывающая сотни квадратных километров в Сиамском заливе с регулярным шагом в обоих направлениях, равным 1 морской миле

## 4.2 Кластерный анализ карт VBD детекций

Пространственные закономерности на картах распределения VDB детекций, построенные по данным, охватывающим длительные периоды времени (от нескольких месяцев до нескольких лет) могут быть проанализированы с помощью алгоритмов кластеризации. В большинстве случаев количество пространственных кластеров на карте заранее неизвестно, а число точек в облаке велико (от тысяч до миллионов). Это сужает набор возможных для применения алгоритмов кластеризации до одного алгоритма DBSCAN<sup>1</sup>, (Burkhauser et al., 2000). Наши предварительные испытания показали, что DBSCAN можно использовать как для микрогранулярного анализа часто детектируемых огней (размером порядка 100 м) как это показано на рисунках 4.2 и 4.3, так и для изучения крупномасштабных распределений рыболовных зон (размером порядка 100 км). Несмотря на то, что геопространственный модуль в базе данных Postgres имеет внутреннюю реализацию алгоритма DBSCAN, в приведенных ниже примерах мы использовали Hierarchical DBSCAN (Campello et al., 2013), (функция hdbscan) из пакета Scientific Python.

Ниже приведен пример скрипта анализа для программной системы Jupyter notebook:

Jupyter скрипт

Configure Python and plotting library

In [1]:

<sup>1</sup> <https://www.aaii.org/Papers/KDD/1996/KDD96-037.pdf>

```

import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import psycpg2

%matplotlib inline

sns.set_context('poster')
sns.set_style('white')
sns.set_color_codes()
plot_kwds = {'alpha': 0.5, 's': 8, 'linewidths': 0}

```

Parse result set from VBD database. Only low-moon detections are selected.

In [2]:

```

def search_roi(database, neLat=90, swLat=-90, swLon=180, neLon=-180):
    con = psycpg2.connect("dbname='"+database+"' user='jjn' host='localhost'
password='')
    cur = con.cursor()

    sql = "SELECT "
    sql += "Lon_DNB, "
    sql += "Lat_DNB "
    sql += "FROM vbd "
    sql += "WHERE ST_Intersects(ST_SetSRID(ST_MakeBox2D(ST_MakePoint(%s,
%s)," % (swLon, swLat)
    sql += "ST_MakePoint(%s, %s)),4326), geom) " % (neLon, neLat)
    sql += "AND Qf_Detect IN (1, 2, 3, 8) AND (LI < 0.0001) "
    sql += "AND Date_Mscan >= '2016-01-01' AND Date_Mscan < '2017-01-01' "
    sql += "ORDER BY Date_Mscan;"

    print sql

    cur.execute(sql)
    npar = np.array(cur.fetchall())

    con.close()
    return npar

```

In [3]:

```

database = 'vbddb_v23'
test_data = search_roi(database, neLat="-3.5", neLon="105", swLat="-
6.5", swLon="110")

```



```
SELECT Lon_DNB, Lat_DNB FROM vbd WHERE
ST_Intersects(ST_SetSRID(ST_MakeBox2D(ST_MakePoint(110, -6.5),ST_MakePoint(105, -
3.5)),4326), geom) AND Qf_Detect IN (1, 2, 3, 8) AND (LI < 0.0001) AND Date_Mscan >=
'2016-01-01' AND Date_Mscan < '2017-01-01' ORDER BY Date_Mscan;
```

In [4]:

Run HDBSCAN for given min\_cluster\_size = 100

In [5]:

```
import hdbscan
plot_clusters(test_data, hdbscan.HDBSCAN, (), {'min_cluster_size':100})
```

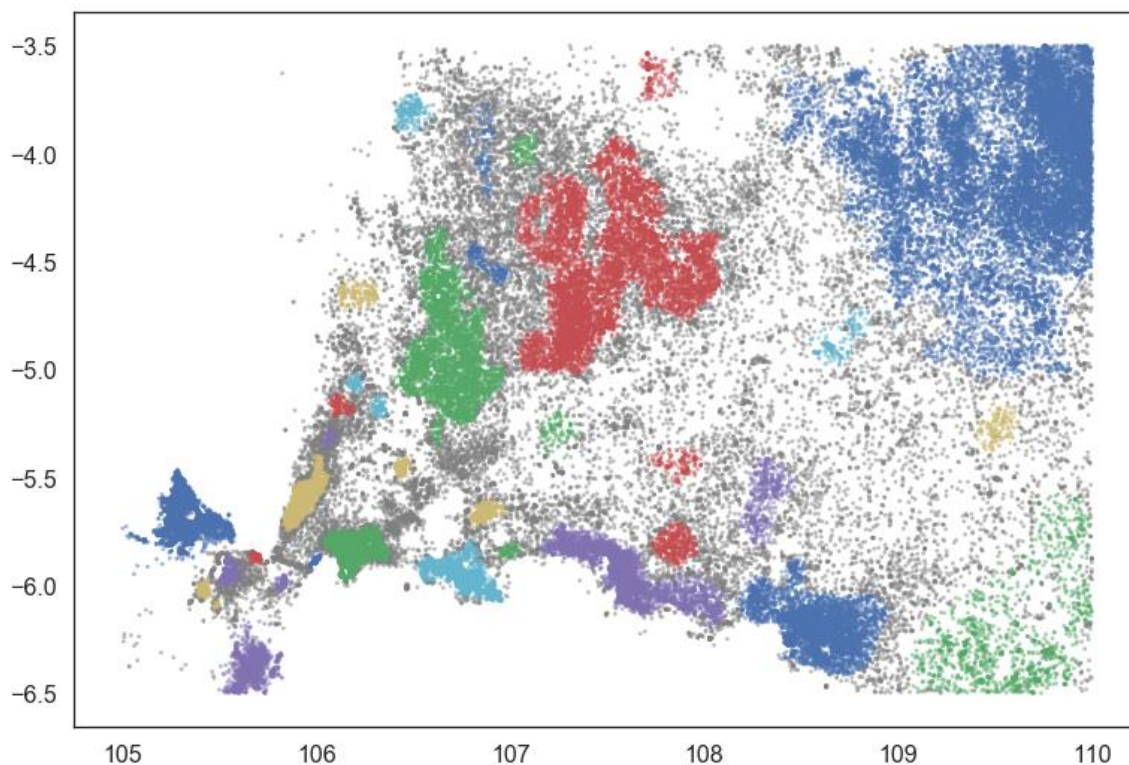


Рисунок 4.4 –карта кластеров детекций судовых огней, построенная с использованием алгоритма DB Scan

## ГЛАВА 5. АНАЛИЗ РАСПРЕДЕЛЕНИЯ ДЕТЕКЦИЙ У ДАЛЬНЕВОСТОЧНЫХ БЕРЕГОВ РОССИЙСКОЙ ФЕДЕРАЦИИ

В рамках отчетного периода был проведен анализ распределения VBD детекций на морской территории, граничащей с Дальним востоком. Сюда вошли: Камчатский край, Магаданская и Сахалинская область, а также Хабаровский край. Проведенный анализ показал, что основной объем детекций на море в исследуемом районе распределяется между:

- огнями рыболовных судов, ведущих рыбный промысел;
- бортовыми огнями судов в районах портов и рейдов;
- морскими платформами и обслуживающими их судами;

— воздушными линиями.

Проведенный анализ показал, что выявленные VBD детекции хорошо согласуются с основными промысловыми зонами Охотского моря и Курильских островов как с точки зрения их пространственного расположения, так и с точки зрения сезонности (на примере зоны промысловой добычи сельди).

Для анализа мы взяли временной период с января 2017 года по конец февраля 2018 года (14 месяцев).

### **5.1 Морские платформы**

На рисунке 5.1 показаны кластеры детекций около восточного побережья о. Сахалин. На рисунке явно прослеживаются 4 кластера. Это морские платформы. В качестве подтверждения мы взяли список факелов сжигания ПНГ за 2017 год и отображали их на карте приложения Google Earth (Рисунок 5.2). Видно, что выделенные кластеры действительно соответствуют факелам сжигания ПНГ, расположенным на морских платформах. В качестве дополнительного подтверждения, на рисунке 5.3 представлена фотография одной из таких платформ, полученная из сервиса Google Earth (платформа “Орлан”).

Сами платформы сосредоточены в центре кластеров, и представляются в виде плотного облака точек. На рисунке 5.4 приведен пример увеличенного кластера вокруг платформы “Орлан”. Вокруг плотного облака точек, соответствующего платформе, проявляется облако детекций, который мы относим к детекциям судов, обслуживающих платформу, а также к размытым детекциям из-за облаков.

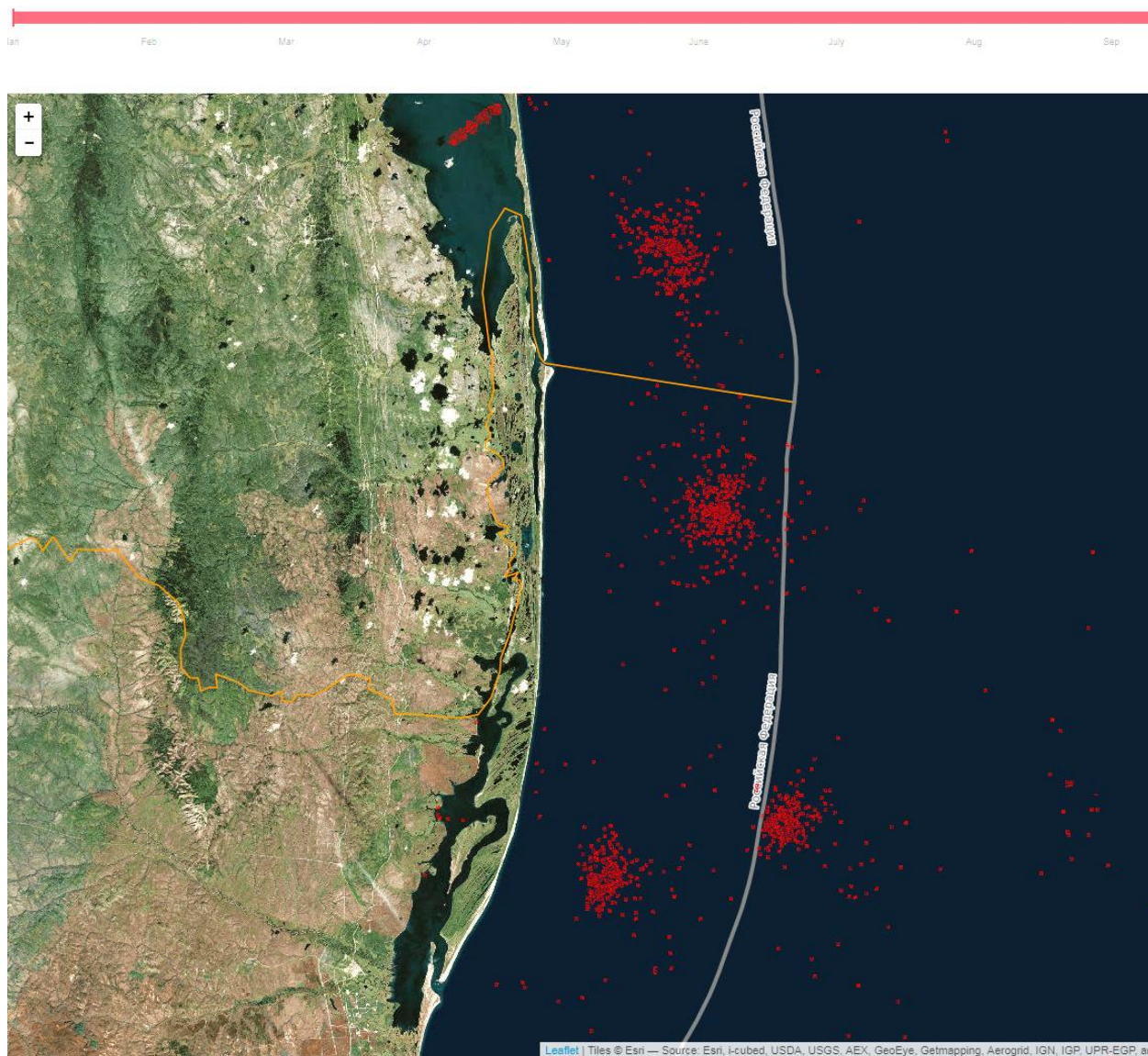


Рисунок 5.1 – Кластеры VBD детекций в районе восточного побережья острова Сахалин (VBD – акроним англ. VIIRS Boat Detector, алгоритм детектирования ночных огней на море)



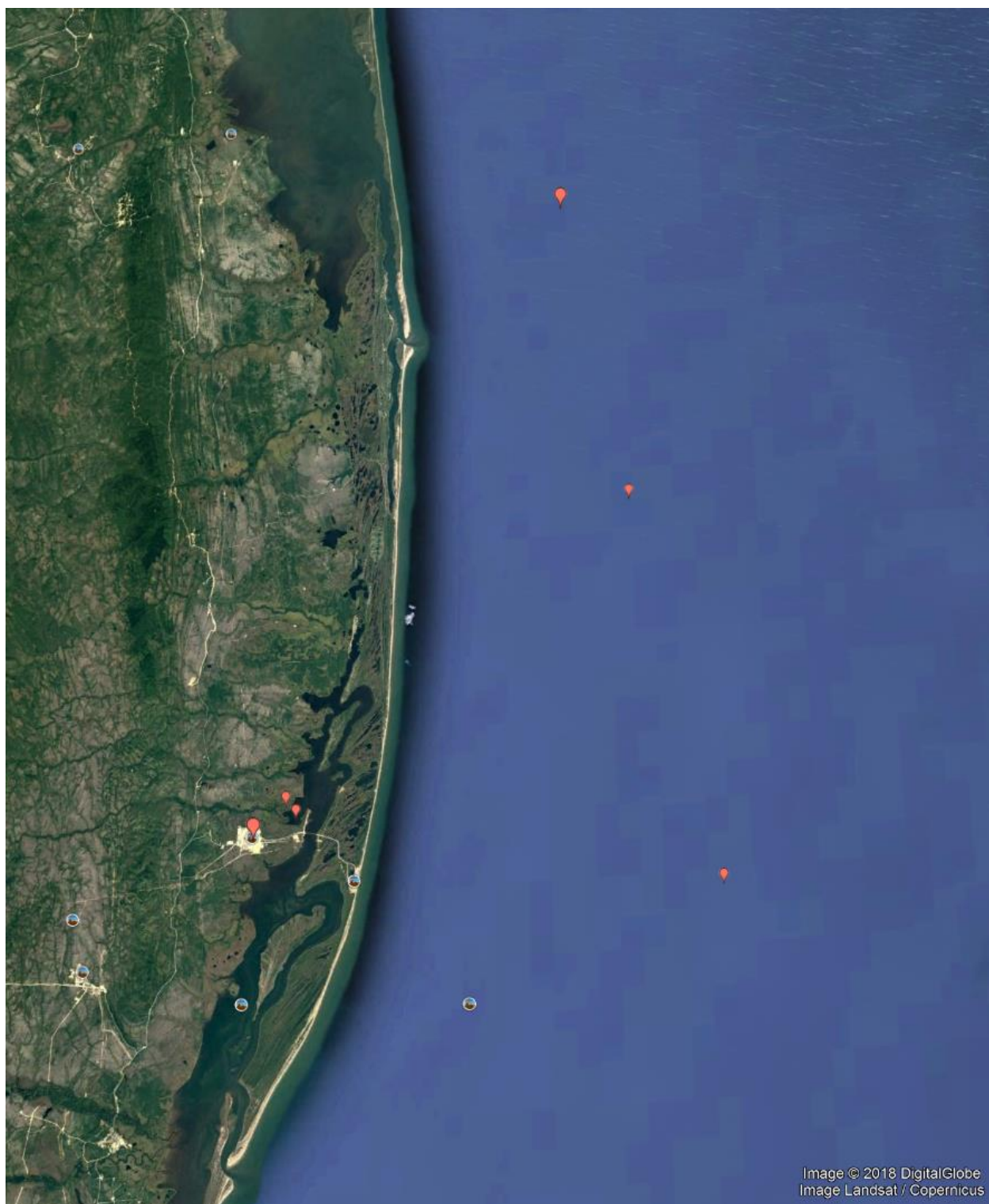


Рисунок 5.2 – Карта факелов сжигания природного газа за 2017 год в районе восточного побережья острова Сахалин

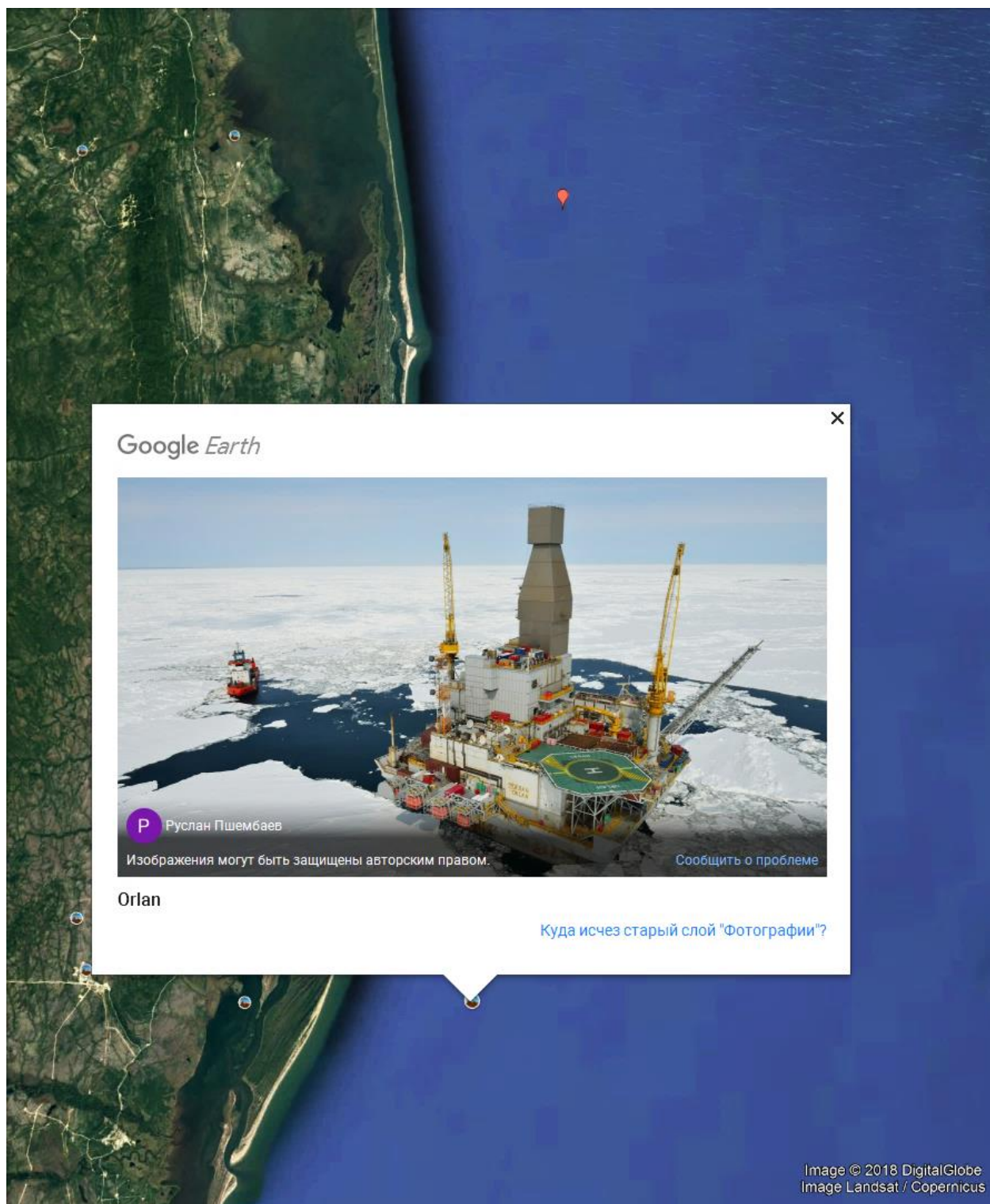


Рисунок 5.3 – Фотография платформы “Орлан” на сервисе Google Earth

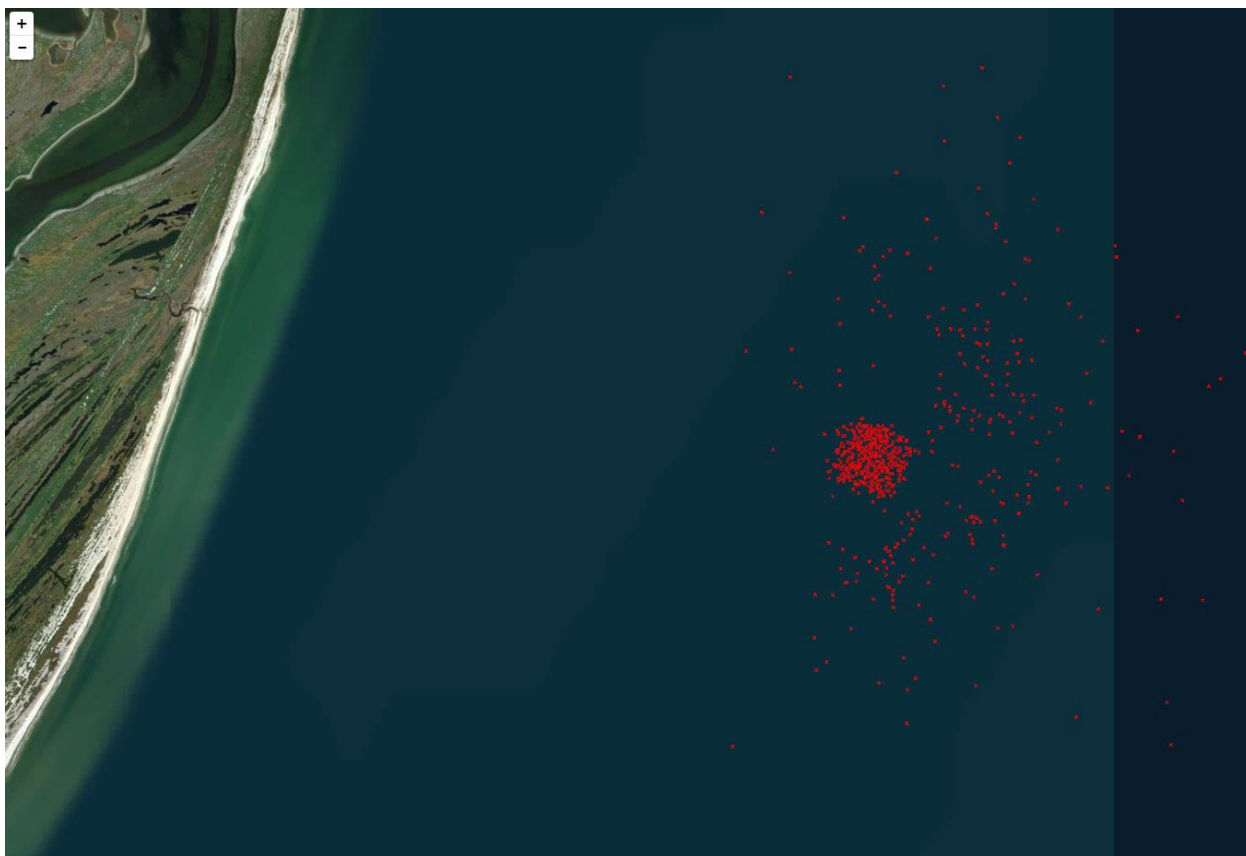


Рисунок 5.4 – Кластер VBD-детекций в районе морской платформы “Орлан”

Отдельный вопрос вызывают детекции (см. рисунок 5.1) на территории залива Пильтун (вверху рисунка). Детекции появились только в период с декабря по февраль 2018 года и соответствуют по расположению подводного перехода нефтепровода Эксон НЛ через залив Пильтун. С декабря 2017 по апрель 2018 «Эксон НЛ» вела ремонтные работы на этом нефтепроводе, для чего на льду залива была наморожена специальная ледовая дорога шириной около 30 метров и толщиной до нескольких метров. Существует экспертное мнение<sup>1</sup>, что это было одной из причин массовой гибели сельди в заливе в 2018 году. Мы считаем, что обнаруженные VBD детекции соответствуют транспорту, участвовавшему в ремонтных работах.

## 5.2 Зоны промыслового рыболовства у берегов п-ова Камчатка

На рисунке 5.5 приведена карта VBD детекций исследуемого района. Детекции представлены красными точками. Линией, проходящей около берега, представлена граница территориальных вод. На карте можно видеть концентрации детекций в районах юго-западного побережья Камчатки, в районе Курильских островов, в северной части охотского моря, а также у берегов Сахалина и, особенно, у его южной оконечности

<sup>1</sup> "Трагедия в Пильтунском заливе" // "Советский Сахалин" 09.07.2018 [http://debri-dv.com/article/19586/tragediya\\_v\\_piltunskom\\_zalive](http://debri-dv.com/article/19586/tragediya_v_piltunskom_zalive)



между Сахалином, Курильскими островами и островом Хоккайдо. Центральная же часть охотского моря (в сторону южной его оконечности) бедна детекциями.

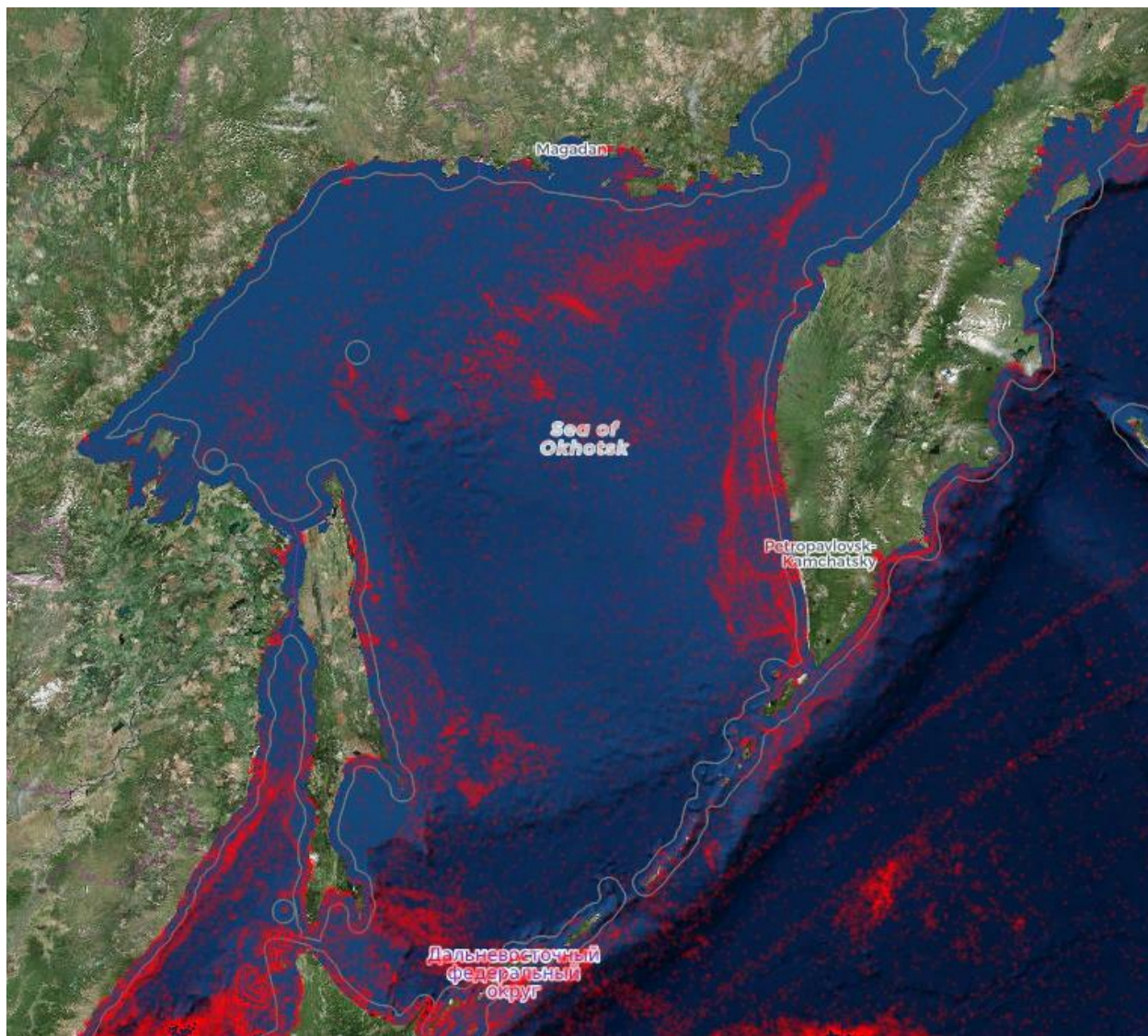


Рисунок 5.5 – Карта VBD детекций исследуемого района

Местоположение и распределение указанных кластеров хорошо соответствует распределению основных зон рыбного промысла, представленного на рисунке 5.6, на котором также видно, что зона в центре охотского моря не используется для промыслового лова рыбы.



Рисунок 5.6 – Карта основных зон рыбного промысла в Охотском море (с сайта <https://geographyofrussia.com/>)

На рисунке 5.7 увеличена карта детекций в области полуострова Камчатка. Мы видим, что распределение относительно полуострова симметрично: на западной стороне (в Охотском море на Западно-Камчатском шельфе) детекций больше и они простираются дальше от берега. Это хорошо согласуется с данными по рыбному промыслу, например, с картой распределения добывающего снюрреводного флота в прикамчатских водах (Рисунок 5.8). Несмотря на то, что карта получена по данным 2002-2003 годов, можно предполагать, что зона лова снюрреводного флота изменилась несильно. Конечно, снюрреводный флот не может быть описывать рыболовные суда всех типов, тем более что основная часть изъятия рыбных ресурсов в прикамчатских водах приходится на



разноглубинные тралы<sup>1</sup>. Тем не менее второе место по количеству промысловых усилий и вылову в регионе принадлежит снюрреводному лову, поэтому распределение судов именно этого типа дает хорошее представление о распределении рыбопромысловых судов. Таким образом, логично сделать заключение, что те детекции, которые мы наблюдаем на рисунке 5.7, соответствуют именно рыболовным судам.

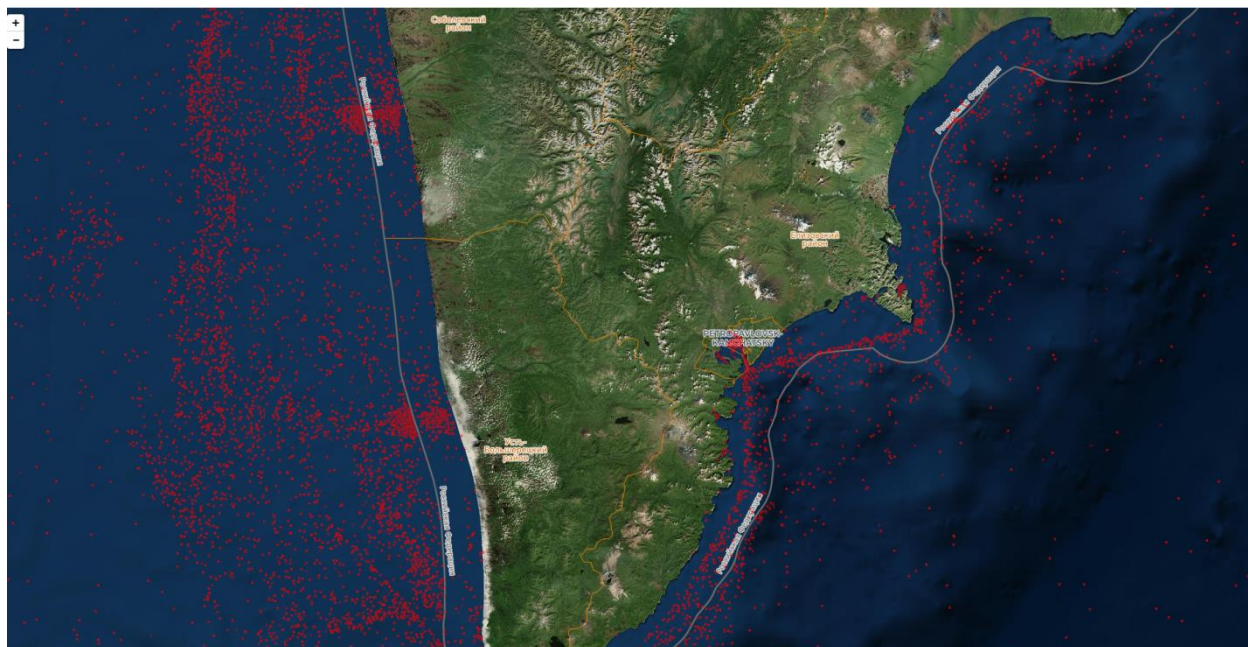


Рисунок 5.7 – Карта VBD детекций в области камчатского полуострова

---

<sup>1</sup> <http://oxotskoe.arktifiksh.com/index.php/ekonomika/282-perspektivy-snyurrevodnogo-promysla-mintaya>

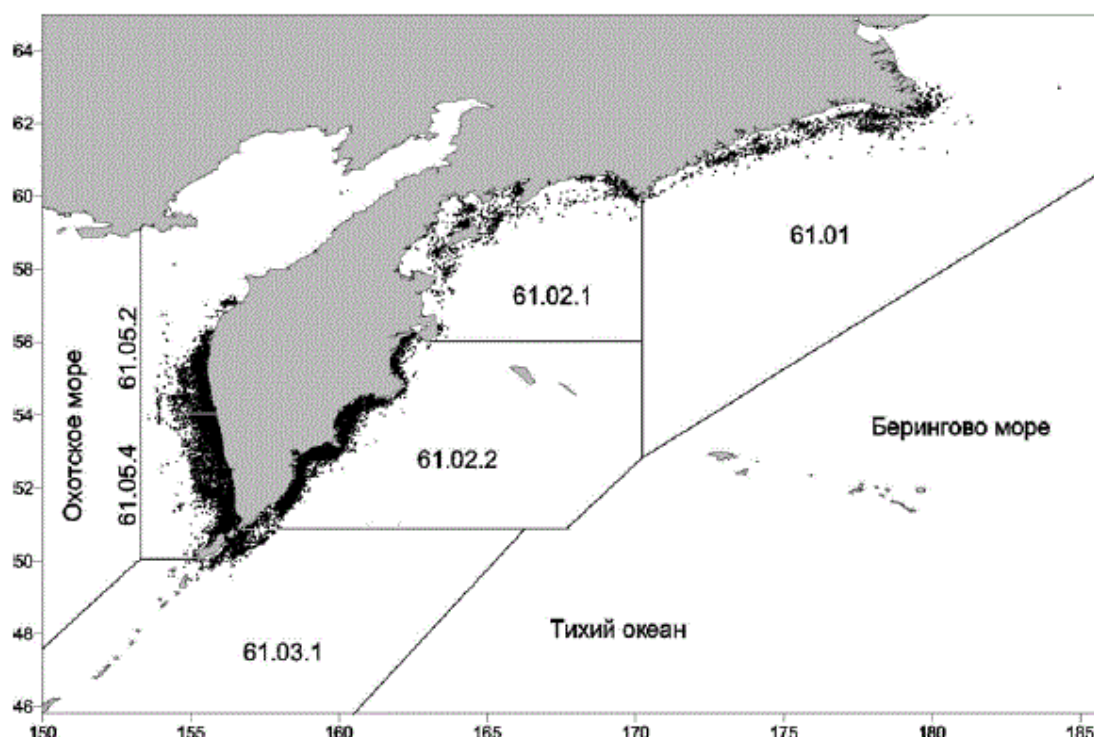


Рисунок 5.8 – Распределение добывающего снюрреводного флота в прикамчатских водах в 2001-2003 гг. (с сайта “Охотское море”<sup>1</sup>)

Рассмотрим еще одну зону рыбного промысла - зону лова корфо-карагинской сельди у северо-восточного побережья Камчатки. На рисунке 5.9 представлена карта распределения флота, занятого специализированным траловым промыслом сельди в западной части Берингова моря в путину 2012 г. Мы анализируем данные за 2017-2018 года, а не за 2012 год, но можно предположить, что распределение флота не сильно изменилось за 5 прошедших лет.

Анализ детекций VBD в районе Берингова моря даст распределение, представленное на рисунке 5.10. На рисунке проявляется кластер как раз в интересующем районе.

<sup>1</sup> Сайт “Охотское море”. <http://oxotskoe.arktikfish.com/index.php>

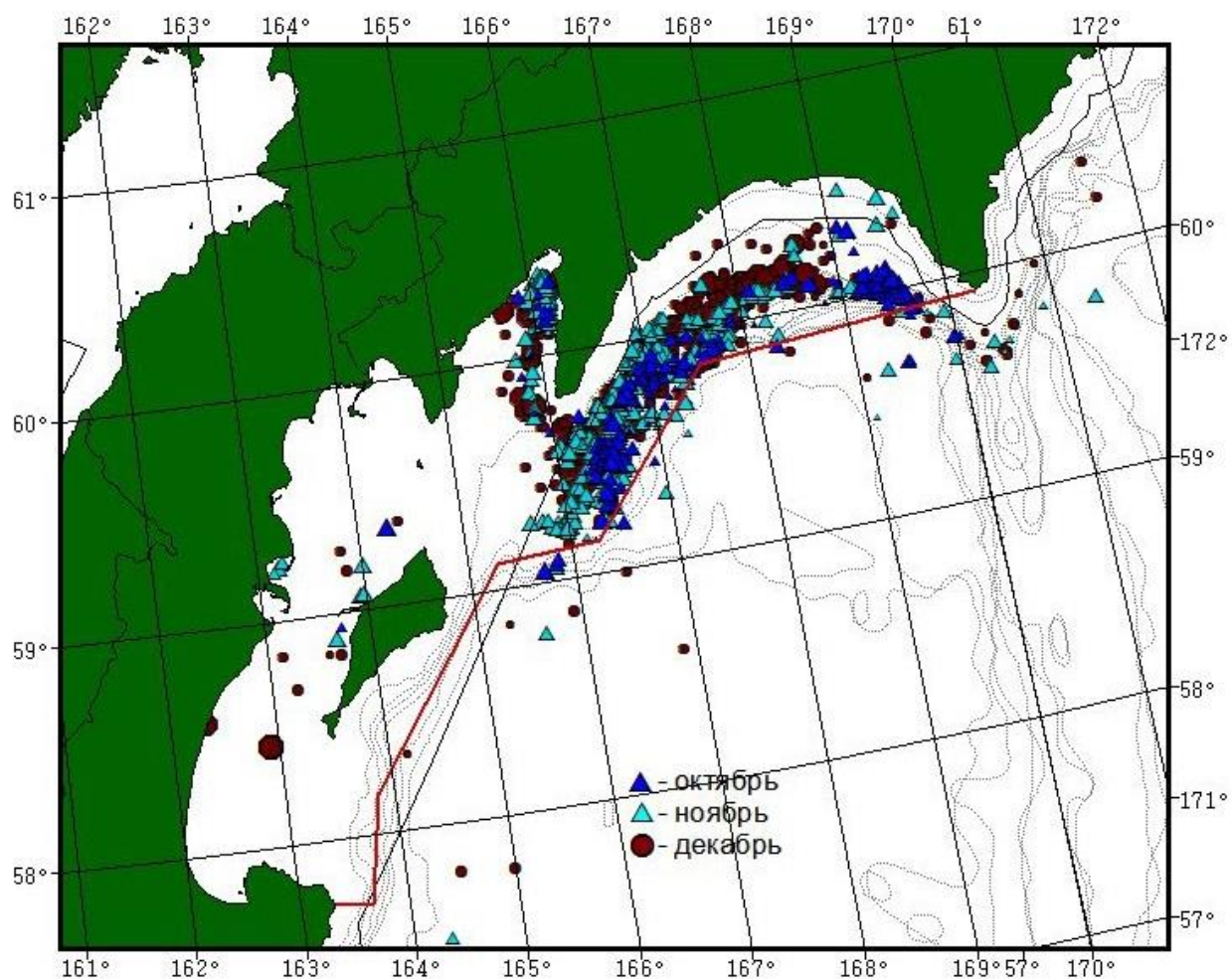


Рисунок 5.9 – Карта распределения флота, занятого специализированным траловым промыслом корфо-карагинской сельди в западной части Берингова моря в путину 2012 г (с сайта камчатНИРО<sup>1</sup>)

<sup>1</sup> Промысел в прикамчатских водах. Предварительные итоги промысла корфо-карагинской сельди в путину 2012 года. КамчатНИРО. [http://www.kamniro.ru/presscenter/obzory\\_promysla/promysel/predvaritelnye\\_itogi\\_promysla\\_korfokaraginskoj\\_seldi\\_v\\_putinu\\_2012\\_goda](http://www.kamniro.ru/presscenter/obzory_promysla/promysel/predvaritelnye_itogi_promysla_korfokaraginskoj_seldi_v_putinu_2012_goda)



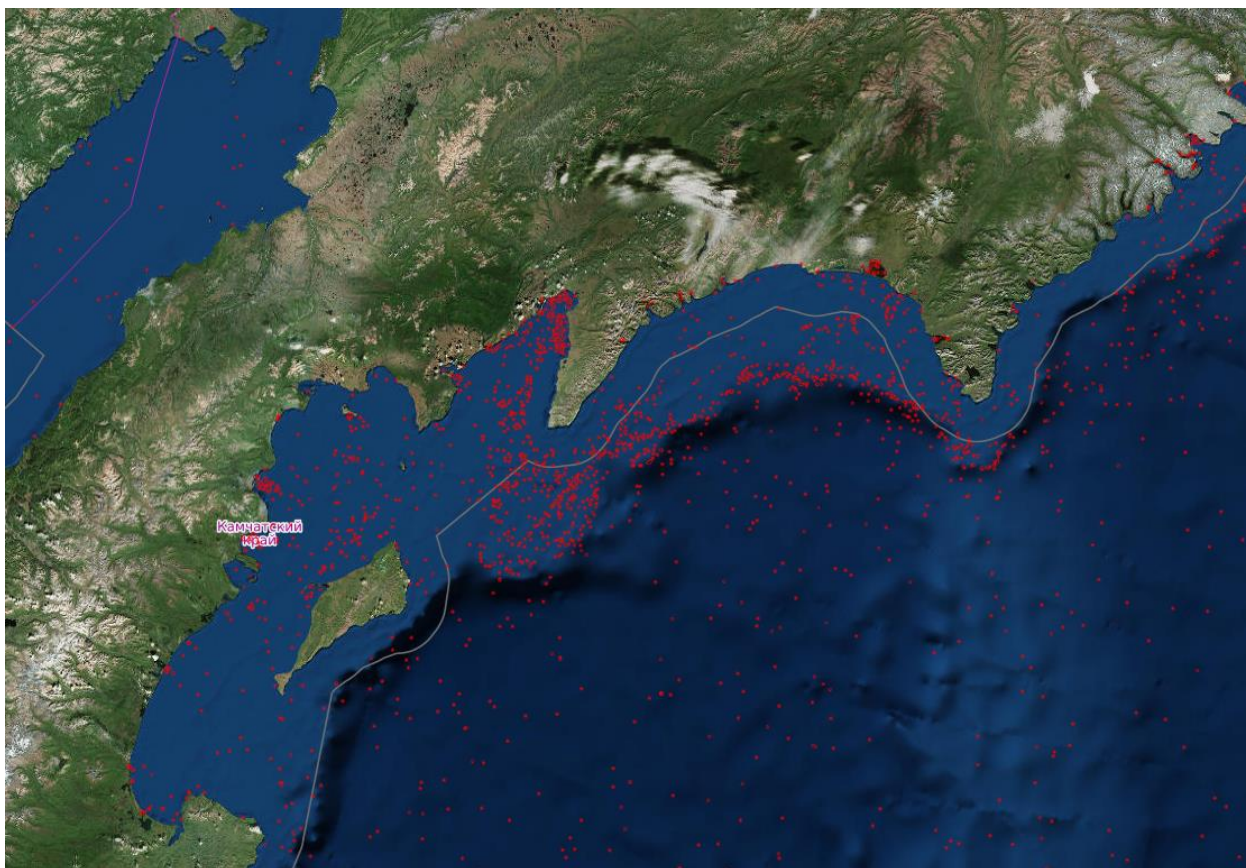


Рисунок 5.10 – Распределение VBD детекций в районе лова корфо-карагинской сельди с января 2017 г. по февраль 2018 г.

Интересно также проследить за сезонными изменениями распределения судов. На рисунках 5.11 и 5.12 представлено распределение VBD детекций в том же регионе, но в разные периоды времени. Ни рисунке 5.11 детекции взяты за период с марта по август 2017 года (6 месяцев), а на рисунке 5.12 - с сентября по декабрь 2017 года (4 месяца). Даже глазом видно, что за 4 месяца в конце года выявляется гораздо больше детекций, чем за 6 предшествующих месяцев. Это связано с тем, что период ловли сельди приходится на конец года.

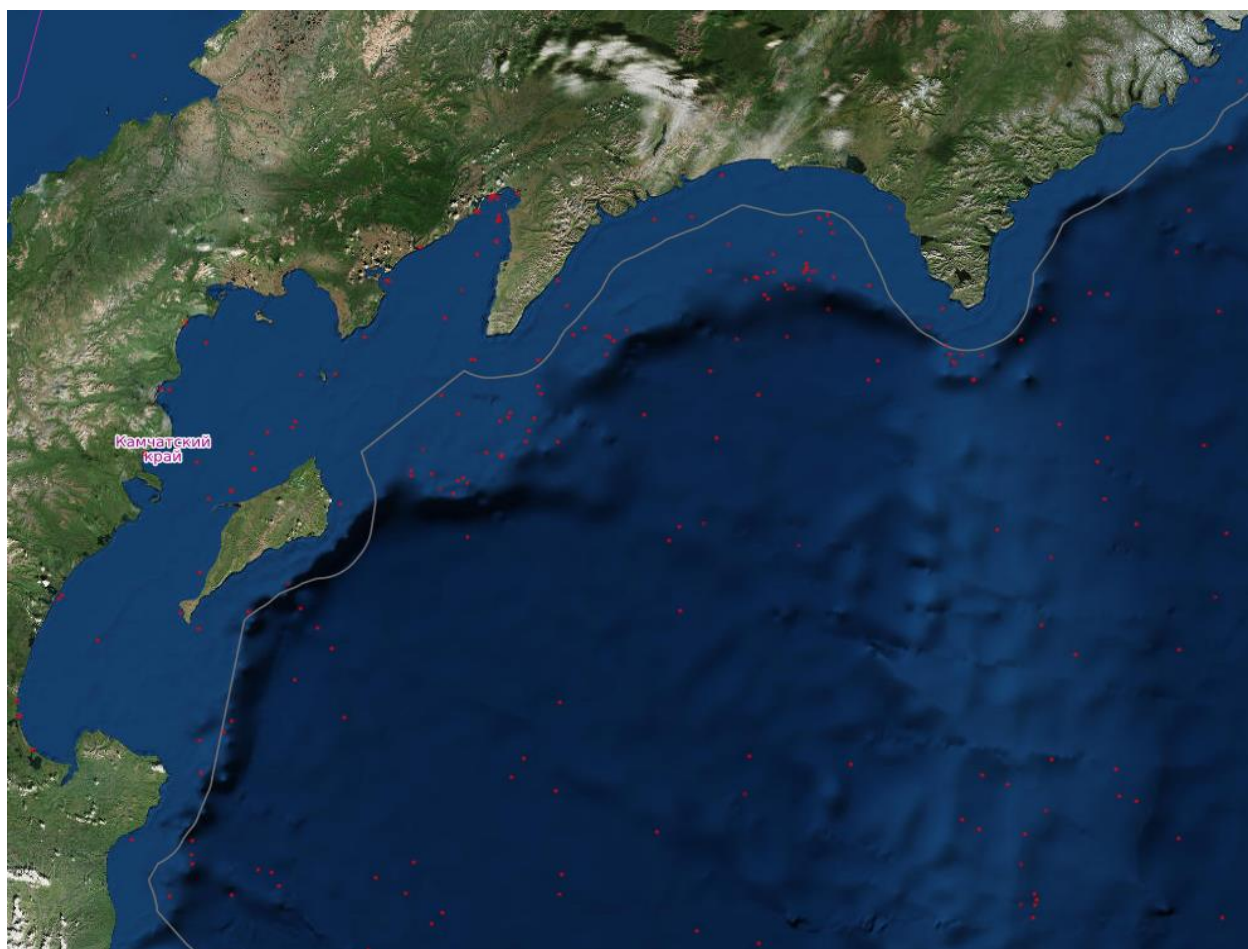


Рисунок 5.11 – Распределение VBD детекций в районе лова корфо-карагинской сельди с марта по август 2017 г.

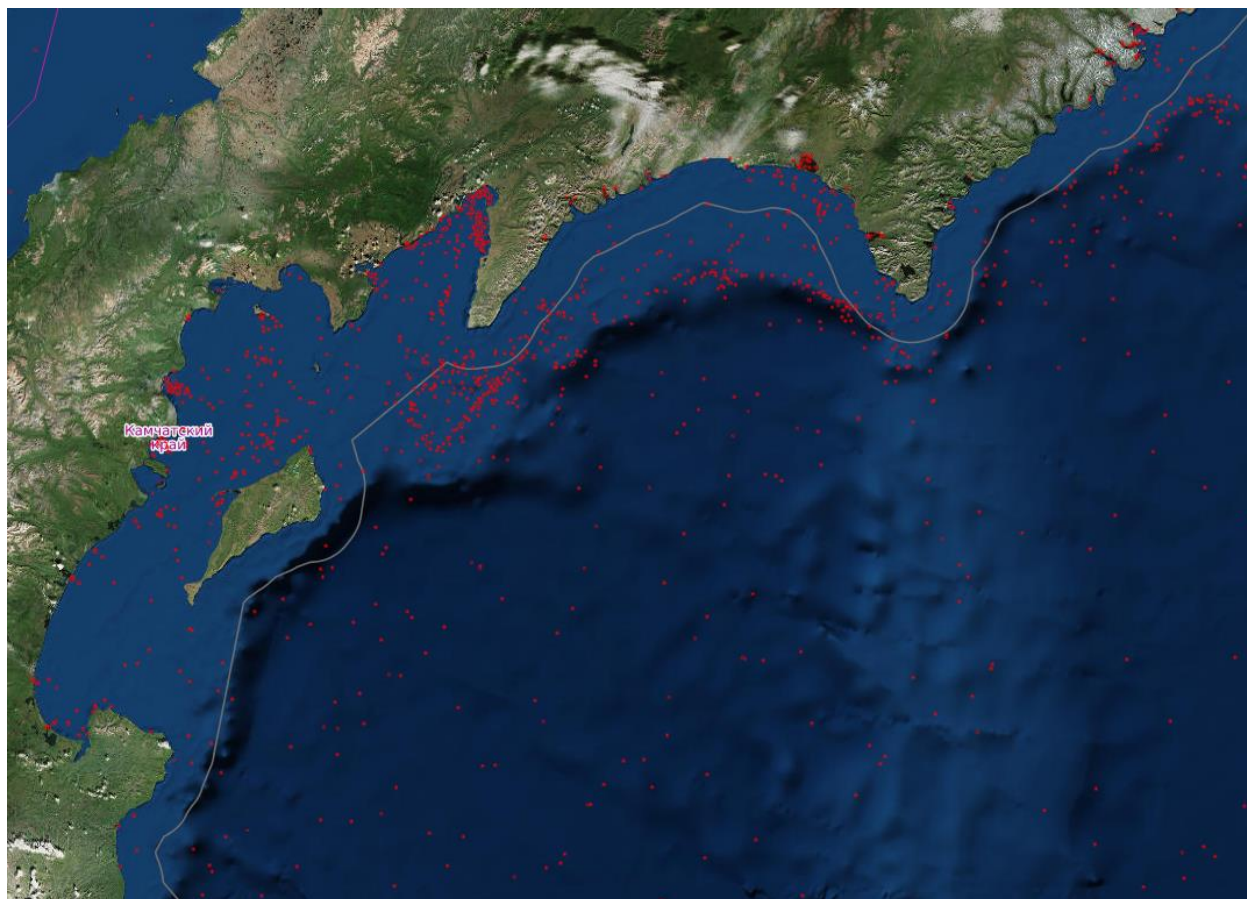


Рисунок 5.12 – Распределение VBD детекций в районе лова корфо-карагинской сельди с сентября по декабрь 2017 г.

### 5.3 Суда в портах и на рейдах

На рисунке 5.13 представлена карта VBD детекций, увеличенная в районе Петропавловска-Камчатского. Видно скопление детекций в бухте “Авачинская губа” на внутреннем рейде судов. Также прослеживаются два направления, ведущие из бухты на север и на юг вдоль побережья.



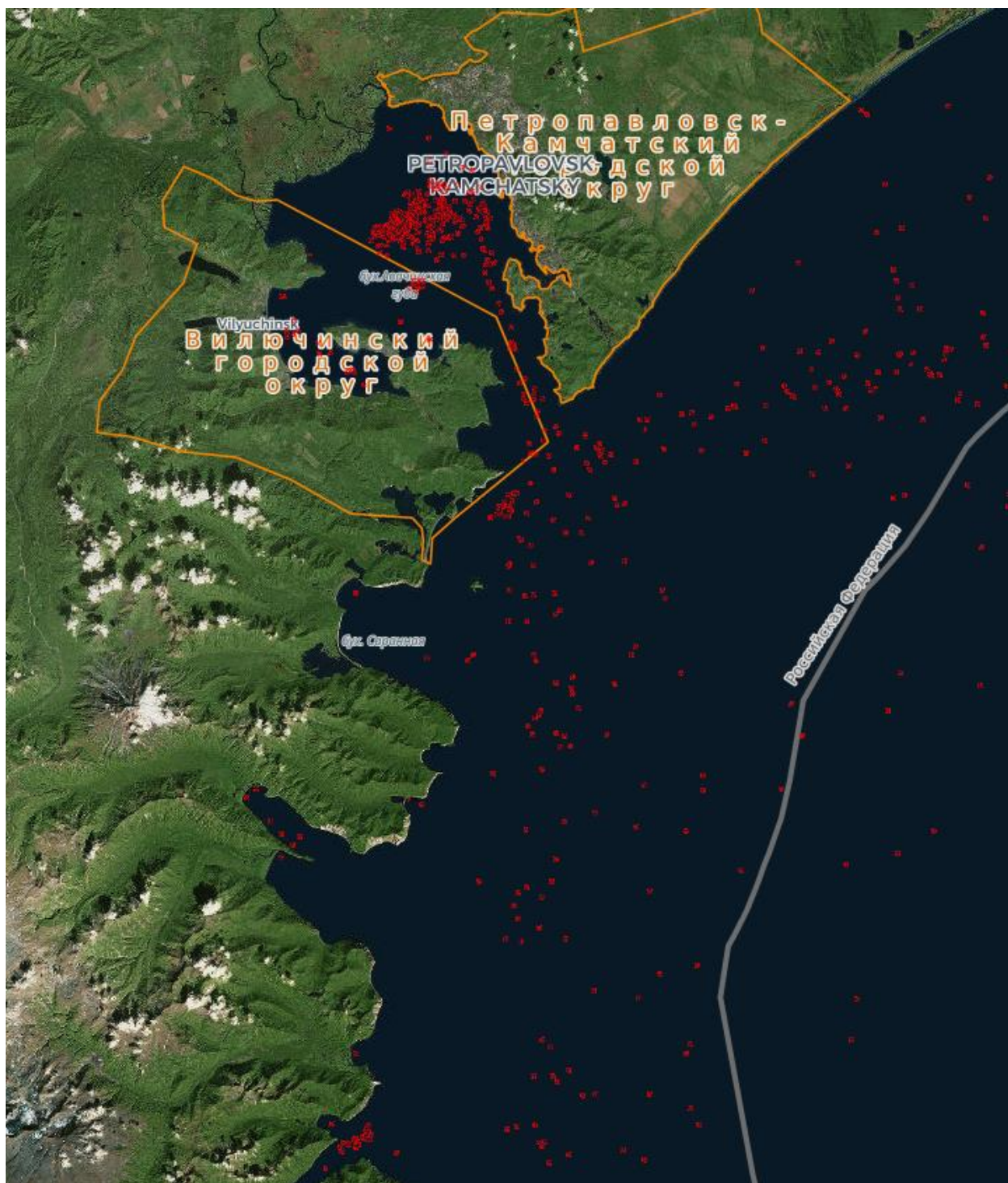


Рисунок 5.13 – Карта VBD детекций с января 2017 г. по февраль 2018 г. в районе порта города Петропавловск-Камчатский

На рисунке 5.14 показана карта VBD детекций в районе порта Магадана.



Рисунок 5.14 – Карта VBD детекций с января 2017 г по февраль 2018 г. в районе порта города Магадан

На рисунке 5.15 показана карта VBD детекций в районе порта Владивостока. Видны детекции в районах восточного и западного рейдов, но нет детекций на внутреннем рейде и в районе самого порта (см. схему на рисунке 5.16).





Рисунок 5.15 – Распределение VBD детекций за период с января 2017 г. по февраль 2018 г. в районе порта г. Владивосток



Рисунок 5.16. Схема области в районе порта г. Владивосток (с сайта РОСМОРПОРТ<sup>1</sup>)

#### 5.4 Детекции воздушных линий

На рисунке 5.17 представлена карта VBD детекций в районе полуострова Камчатка, Командорских и Алеутских островов. На рисунке прослеживаются 2 линии, образованных точками детекций и идущих параллельно со стороны Японии в сторону Аляски между Командорскими и Алеутскими островами.

<sup>1</sup> Схема рейда и порта города Владивосток на сайте РОСМОРПОРТ.  
[http://www.rosmorport.ru/filials/vlf\\_serv\\_loc/](http://www.rosmorport.ru/filials/vlf_serv_loc/)

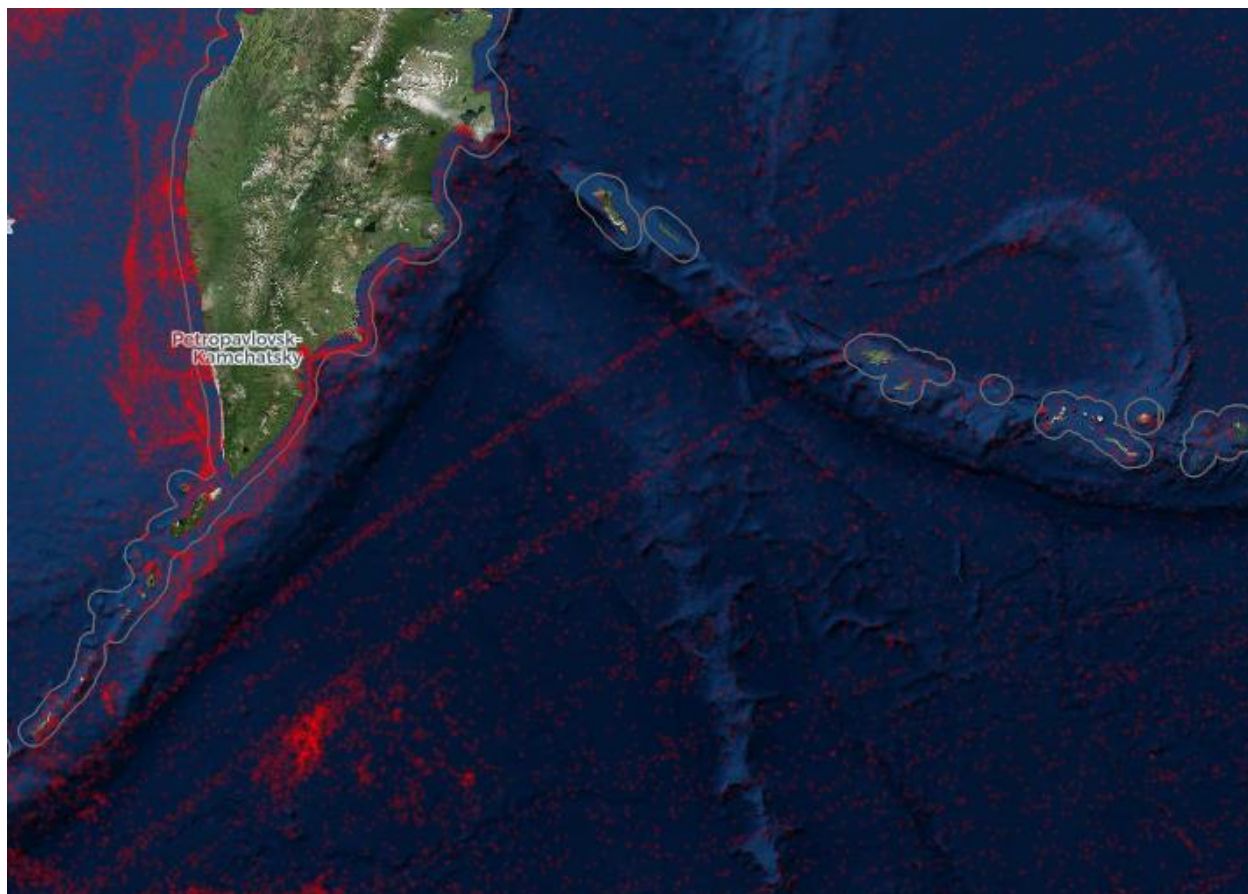


Рисунок 5.17 – Карта VBD детекций в районе полуострова Камчатка, Командорских и Алеутских островов

Скорее всего, выявленные линии соответствуют детекциям самолетов, летящих по фиксированным маршрутам. Множество таких детекций, наблюдаемых из ночи в ночь сливаются в единую линию. Детекций самолетов получается много еще и по той причине, что они летят над облаками, поэтому облачный покров не является препятствием для их обнаружения.

## ГЛАВА 6. ИНДЕКС ЭФФЕКТИВНОСТИ ОГРАНИЧЕНИЙ РЫБНОГО ПРОМЫСЛА

### 6.1 Определение индекса эффективности ограничений рыбного промысла

Были разработаны два индекса оценки эффективности ограничений рыбного промысла, использующие результаты алгоритма VBD (сокращенно будем называть их VCI-индексы, от английского “VBD closure indices”): табличный и grid-ориентированный. Оба индекса используют в качестве входных данных временные ряды наблюдений, заданные в узлах регулярной сетки с шагом по пространству равным 15 угловым секундам. Различие индексов заключается в том, что табличный VCI-индекс вычисляется с использованием данных VBD, агрегированных по времени и по



пространству по всей исследуемой зоне, в то время как grid-ориентированный VCI-индекс использует результаты VBD, агрегированные только по времени (агрегация по пространству есть, но она локальная: например, 30 угловых секунд).

Основная идея VCI-индексов заключается в сравнении числа детекций судов, зафиксированных в период действия ограничения рыболовства, со средним числом детекций судов, зафиксированных в период отсутствия ограничений. Наша первая попытка разработки VCI-индекса заключалась в том, чтобы просто вычислять отношение числа детекций за период ограничения рыбной ловли к среднему числу детекций за те же месяцы вне периода ограничений с последующим вычитанием полученного результата из 1 и умножением на 100 (формула 6.1). Согласованность по месяцам мы использовали, чтобы нивелировать эффект сезонных изменений.

Данный подход к вычислению VCI-индекса дает правдоподобные результаты, если в период ограничения рыбной ловли число судов (и соответственно VBD детекций) действительно сократилось, как и должно быть (для этого и вводятся ограничения). Например, полученное по данной формуле значение, равное 90% показывает, что число судов в исследуемой зоне во время ограничения рыбной ловли сократилось на 90% по сравнению с периодом, когда ограничение не действовало. Отрицательные значения индекса, наоборот, показывают увеличение числа детекций во время периода действия ограничений. Но данная формула не работает для случаев, когда число детекций в период отсутствия ограничений равно нулю (в этом случае мы получаем деление на ноль). Кроме того, в случае, когда число детекций в период ограничений сильно превышает число детекций вне действия ограничений, формула может давать очень большие отрицательные значения.

$$\text{Simple ratio VCI} = 100 * (1 - (\text{VBD\_cl} / \text{VBD\_ref})) \quad (6.1)$$

Чтобы устранить указанные негативные эффекты, мы разработали нормированную версию VCI-индекса. Чтобы гарантировать, что VCI-индекс имеет диапазон от -100 до 100, мы берем в отношении в качестве значения знаменателя максимум из двух величин: числа детекций во время ограничений и числа детекций вне периода ограничений. В случаях, когда количество детекций, зафиксированных в течение периода действия ограничений меньше числа детекций контрольного периода, значение VCI-индекса будет положительным, в противном случае значение VCI-индекса будет отрицательным, но не меньше -100. При такой формулировке VCI-индекс, равный 50% показывает, что количество детектируемых лодок сократилось наполовину по сравнению с контрольным периодом. И наоборот, значение VCI-индекса, равное -50% указывает, что количество детекций увеличилось на 50% по сравнению с контрольным периодом. Значение VCI-индекса, равное -100, соответствует ячейкам сетки, в которых VBD-детекции присутствуют в период действия ограничений, но отсутствуют в контрольный период. И наоборот, значение VCI-индекса, равное 100, соответствует ячейкам, в которых VBD-детекции присутствуют в контрольный период, но отсутствуют в исследуемый период действия ограничений.

$$VCI = 100(VBD\_cl - VBD\_ref) / \max(VBD\_ref, VBD\_cl), \text{ где:} \quad (6.2)$$

VBDcl - суммарное число VBD-детекций в период действия ограничений;

VBDref - среднее число VBD-детекций в период отсутствия ограничений.

Grid-ориентированный VCI-индекс вычисляется аналогичным образом, только вместо общего числа детекций по всему региону сравнивается количество детекций в каждой ячейке географической сетки. Grid-ориентированный VCI-индекс позволяет определить эффективность введения ограничений не только по отношению ко всей зоне запрета в целом, но и выявить пространственные вариации эффективности внутри зоны запрета.

## 6.2 Экспериментальные исследования и апробация индекса эффективности ограничений рыбного промысла

### Регионы, на которых проводились экспериментальные исследования

Мы провели анализ с использованием VCI-индекса для трех случаев. В первом случае речь идет об ограничении рыбной ловли, связанном с промышленной аварией, повлекшей выброс в море токсичных отходов. Второй случай — это сезонное ограничение рыбной ловли, которое действует в течение трех месяцев в году. Третий случай — это запрет коммерческой рыбной ловли в прибрежных водах. В таблице 6.1 приведены краткие сведения о периоде действия ограничений.

Таблица 6.1 – Названия ограничений рыбной ловли и периоды их действия

Ограничение рыбного промысла	Дата начала	Дата конца	Примечание
Муниципальные воды г. Кауаян	1 апреля 2014г.	по наст. время	
Запрет на вылов пятнистой ставриды у о. Палаван	1 ноября 2015 г.	31 января 2016 г.	Было повторено в сезоне 2016-2017гг.
Авария у берегов Вьетнама	4 мая 2016 г.	30 сентября 2016 г.	Авария вызвала выброс в море токсичных отходов в начале апреля 2016

*Авария у берегов Вьетнама в 2016 году*

Авария у берегов Вьетнама, произошедшая в 2016 году, была связана с выбросом токсичных отходов на сталелитейном заводе в Формозе, расположенном в экономической зоне Вунг-Анг. Точная дата выброса не была раскрыта, но большие объемы мертвой рыбы начало выбрасывать на берег уже 6 апреля 2016 года. Четыре провинции пострадали от аварии: Ха Тинь, Куанг Бинь, Куанг Три и Туа Тхиен Хюэ (рис. 6.1). К началу мая погибло около 100 тонн рыбы. В ответ на опасения, связанные с возможным воздействием на здоровье людей загрязненных морепродуктов, правительство ввело санитарный запрет на рыбную ловлю в 20 мильной прибрежной зоне 4 мая 2016 года. Запрет на рыбную ловлю был снят в конце сентября (PhysOrg, 2017).

#### *Запрет на вылов пятнистой ставриды у острова Палаван*

Это сезонное ограничение, установленное для защиты численности популяции пятнистой ставриды. Пятнистая ставрида, или «galunggong», считается рыбой бедняков на Филиппинах из-за его низкой стоимости. В водах северной части острова Палаван добывается более 90% пятнистой ставриды, взятой от общего объема ее добычи на Филиппинах. Период ограничения на вылов ставриды составляет 3 месяца - с 1 ноября по 31 января - и приходится на период нереста. Три типа рыболовных снастей запрещено к использованию в это время: кошельковый невод, кольцевой невод и односадковый ставной невод. Лодки, использующие эти типы снастей на Филиппинах, как правило используют электрические огни для привлечения рыбы. Впервые ограничение на лов пятнистой ставриды было введено в 2015 году (совместный административный приказ DA-DILG, № 1, 2015) и далее повторялось каждый год. В 2015 году ограничение носило лишь рекомендательный характер до его подписания в середине декабря. Трехмесячное ограничение было также повторено в 2016 и 2017 годах (рисунок 6.2).

#### *Муниципальные воды г. Кауаян*

На Филиппинах действует запрет на коммерческий промысел рыбы в прибрежных водах, простирающихся на 15 км от берега. Они известны как «муниципальные воды» и управляются подразделениями местного самоуправления (LGU). Цель запрета заключается в обеспечении стабильности запасов морских ресурсов и улова коренным жителям. Эти ограничения действуют на государственном уровне с 1998 года (десятый Конгресс Филиппин, 1998 год). Несмотря на этот запрет, коммерческий рыбный промысел довольно распространен во многих водах LGU. Известно, что у властей города Кауаян в провинции Западный Негрос постоянно возникают проблемы с браконьерами. В конце марта 2014 года Провинциальная организация по управлению окружающей средой (ПЕМО) и другие организации начали кампанию по обеспечению запрета на коммерческие ловушки, работающие в муниципальных водах г. Кауаян. Данные VBD были использованы для контроля эффективности регионального запрета на вылов рыбы (рисунок 6.2).

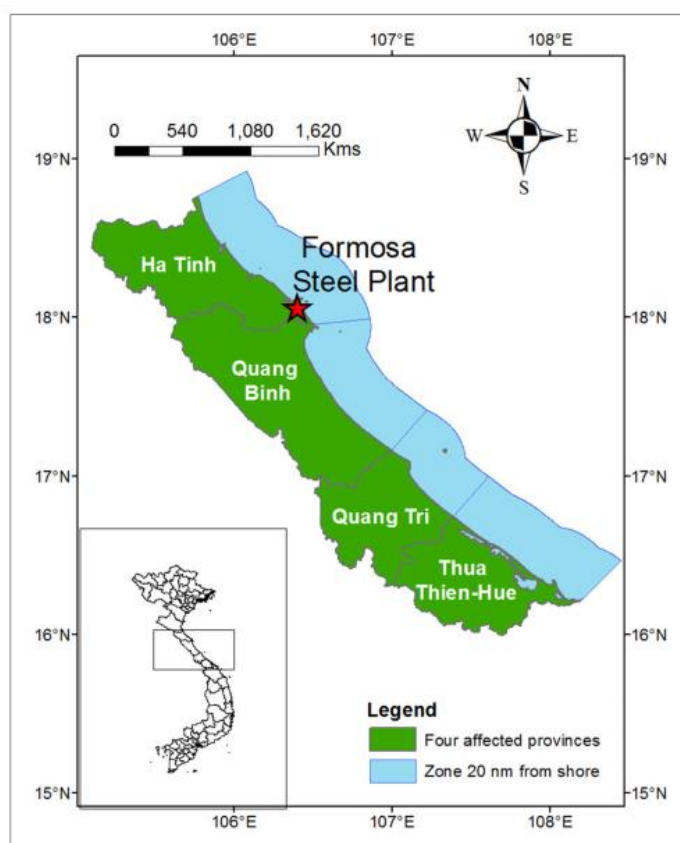


Рисунок 6.1 – Провинции Вьетнама, пострадавшие от аварии 2016 года

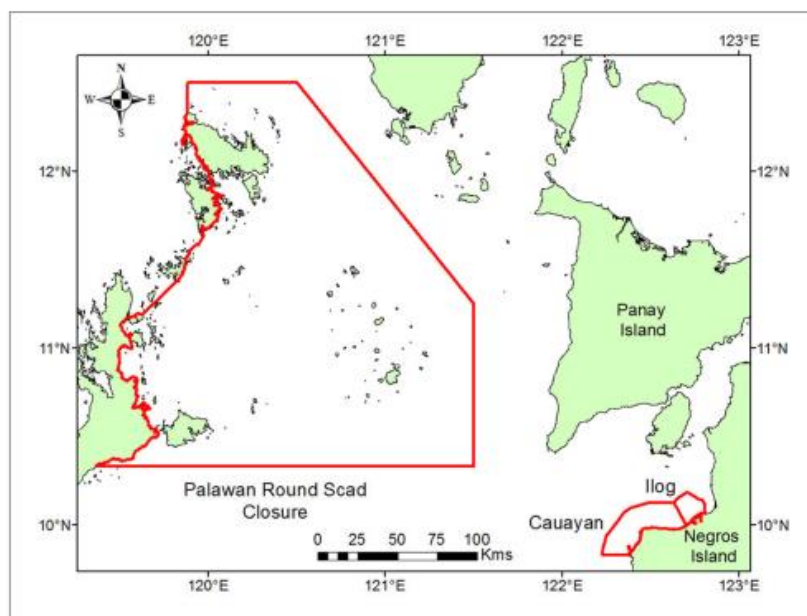


Рисунок 6.2 – Карта вод Филиппин, на которой выделено 2 региона ограничения рыбной ловли: в районе острова Палаван и в муниципальных вода Кауаяна



## Дополнительные методы исследования

### *Анализ типов судов, обнаруженных при помощи VBD*

Чтобы определить, какие типы рыболовных судов обнаруживаются при помощи алгоритма VBD, было проведено сопоставление VBD детекций и VMS данных для 3550 Индонезийских рыболовных судов, для которых известен их тип. Сопоставление проводилось с использованием спутниковых данных VIIRS, полученных в период с 2014 по сентябрь 2016 годов (таблица 6.2). Как правило, каждое судно подает VMS-сигнал о своем положении примерно 1 раз в час. Все сигналы были распределены по трем категориям, в зависимости от их активности: посадка-высадка, транзит, рыбная ловля. Причисление к конкретной категории осуществляется на основе ряда характеристик: удаленность судна от портов, направление движения, скорость (Lee et al., 2010). Ключевой момент сопоставления VMS-данных и VBD-детекций — это предсказание положения каждого судна в момент пролета спутника. Если в радиусе 1 км. от предсказанного на основе VMS-данных положения судна будет обнаружена VBD-детекция, то они считаются соответствующими друг другу. Процент успеха обнаружения судна записывается в результирующую таблицу как вероятность обнаружения. Вероятность обнаружения вычисляется отдельно для категории рыбной ловли и категории транзита судна.

Таблица 6.2. Результаты сопоставления данных VBD и WMS

Gear type bahasa	English	Number of vessels	Fishing match rate (%)	In-transit match rate (%)
Pancing Ulur	Pole and line	3	83.5	13.4
Bouke ami	Squid lift net	535	82.3	7.0
Pancing Cumi	Squid Jigging	190	76.9	6.9
Hand Line Tuna	Hand line tuna	75	72.2	4.3
Pukat Cincin Pelagis Kecil	Purse seine small pelagics	1,149	59.5	4.9
Huhate	Skipjack pole and line	77	9.9	0.4
Rawai Tuna	Longline tuna	433	8.7	2.2
Pukat Ikan	Trawler	216	6.8	3.4
Pengangkut	Transporter	402	5.1	0.8
Jaring insang oseanik	Oceanic gill net	185	4.6	1.8
Jaring liong bun	Gillnet cutcut	144	1.0	0.2
Pukat cincin grup pelagis kecil	Purse seine small pelagics—multiboat	15	0.6	1.6
Rawai dasar	Longline	126	0.3	0.1

### *Статистическое тестирование*

В зонах рыбной ловли, для которых выявляется большое число VBD-детекций соблюдение ограничений должно приводить к более высоким значениям VCI-индекса, рассчитанному для периода действия ограничений, по сравнению с VCI-индексом, рассчитанным для тех же зон, но для периода отсутствия ограничений. Действительно, в этом случае число VBD-детекций, полученных в период действия ограничений, должно

быть меньше, чем в период их отсутствия, в то время как число детекций в контрольный период (т.е. знаменатель в формуле 6.2) останутся неизменными. Чтобы проверить статистическую значимость отличий значений VCI-индексов в период действия ограничений и вне их, мы выполняем t-тесты с использованием нулевых и альтернативных гипотез. Нулевая гипотеза состоит в том, что нет разницы в средних значениях VCI-индексов, рассчитанных для периода действия ограничений и предшествующего периода. Альтернативная гипотеза состоит в том, что средний VCI-индекс, рассчитанный для периода действия ограничений, имеет более высокое значение. Тестирование проводится с помощью методов, описанных Фрейндом и Уилсоном (Freund and Wilson, 2003). Когда нулевая гипотеза отвергается, то принимается гипотеза, что среднее значение VCI-индекса, рассчитанного для периода действия ограничений, больше, чем значение VCI-индекса, рассчитанного для периода, предшествующего ограничениям.

## **Результаты экспериментальных исследований**

### *Какие типы лодок обнаруживает VBD*

Алгоритм VBD позволяет детектировать суда с яркими судовыми огнями, излучающими в небо достаточно сильный сигнал, чтобы он мог быть зафиксирован DNB-детектором. Поэтому не следует ожидать, что VBD будет регулярно обнаруживать любое рыболовное судно. Хорошо детектируются только те рыболовные суда, которые используют световые источники для привлечения улова. Чтобы исследовать, какие именно типы рыболовных судов хорошо детектируются при помощи VBD, мы проанализировали вероятность обнаружения судов заданных типов. Анализируемая вероятность обнаружения судов была вычислена на основании сравнения результатов работы VBD с данными VMS, полученными за 33-месячный период с 1 января 2014 года по 30 сентября 2016 года (таблица 6.2). Как видно из таблицы, наибольшую вероятность обнаружения во время рыбной ловли (>59%) имеют суда, ведущие удебный лов (специфический вид ловли с открытых площадок), лов и переработку кальмара, ручную ловлю тунца, а также сейнеры, ведущие кошельковый лов мелкой океанической рыбы. Суда, ведущие удебный лов полосатого тунца, ярусный лов тунца и траулеры дают вероятность обнаружения при рыбной ловле порядка 6-10%. Суда для ловли жаберной сетью и суда для ярусной ловли дают вероятность обнаружения менее 5%.

Из 3350 рыболовных судов, VMS данные которых были использованы в данном эксперименте, 1548 судов (46,6%) детектируются при помощи VBD в 10% случаев или чаще, когда они находятся в режиме ловли рыбы. Сейнеры, ведущие кошельковый лов мелкой океанической рыбы, дали 57% всех детекций. Сравнение вероятностей детектирования лодок при помощи VBD в режимах рыбной ловли и транзита показывает, что если судно задетектировано при помощи VBD, то 90%, что оно находится в режиме рыбной ловли. Еще одним важным достижением является то, что за 33-месячный период исследования при помощи VBD было выявлено 1 448 986 детекций. Из них только 15,6% детекций соответствовали судам, оснащенным VMS-системами.

Таким образом, 84% VBD детекций в Индонезии было получено от судов, не оснащенных VMS-системами.

### *Агрегация по времени*

Существуют три причины, делающими актуальной агрегацию VBD данных по времени: зависимость количества детекций от фазы лунного цикла, сезонные циклы и плотность облачного покрова, влияющего на картину детекций. Чтобы исследовать эффект, который оказывает фаза лунного цикла на число детекций, мы проанализировали результаты работы алгоритма VBD на спутниковых изображениях, полученных в районе Индонезии. Для каждого дня мы вычисляли яркость лунной засветки и среднее число VBD детекций (рисунок 6.3). Из рисунка видно, что количество детекций и яркость лунной засветки являются обратнозависимыми величинами. При полной луне количество детекций сокращается примерно на 95% по сравнению с новолунием. Сокращение может быть связано с повышением порога детектирования в условиях яркой лунной засветки в сочетании с тем фактом, что рыбная ловля с использованием электрических огней для приманки рыбы не столь эффективна при полной луне.

Чтобы оценить эффективность ограничений, мы предлагаем сравнить количество VBD детекций, выявленных на спутниковых данных, полученных в период действия ограничений, с количеством детекций, выявленных на спутниковых данных, полученных в период отсутствия ограничений. Но уместно ли сравнивать данные за одну ночь (если не делать агрегацию по времени) с данными, полученными в эту же дату, но в другие года? Например, мы можем рассмотреть ограничение на ловлю рыбы, объявленное 1 ноября 2016 года. Корректно ли сравнивать количество VBD детекций, полученных по данным за 1 ноября 2016 года с количеством VBD детекций, полученных, например, по данным за 1 ноября 2012-2015 годов? Существует две причины, по которым такое сравнение может привести к ложным результатам. Во-первых, на количество VBD-детекций влияют лунные циклы. Лунный цикл составляет 27,3 дня, что не соответствует месячным интервалам. В результате существует вариативность лунной освещенности (LI) для одних и тех же дат, но взятых в разные года. Во-вторых, величина и конфигурация облачного покрова наверняка будут различаться для одной и той же даты, но в разные года.

Цель временной агрегации - смягчить воздействие лунных, сезонных и облачных эффектов. Лучшая временная схема агрегации должна решать все три эти проблемы. Так как длина лунного цикла превышает неделю, то недельные временные агрегаты также подвержены влиянию лунного цикла (рис. 6.4), что ставит вопрос о необходимости более длительного временного агрегирования. Один из вариантов - использовать длину лунного цикла (27,3 дня) в качестве окна временного агрегирования. Этот вариант хорошо удаляет эффекты лунного цикла (рис. 6.4). Использование длины лунного цикла также хорошо решает задачу сглаживания облачных эффектов, так как облака перемещаются из ночи в ночь. В течение лунного цикла свободные от облачности данные накопятся по анализируемому региону. Сложность использования длины

лунного цикла в качестве окна временного агрегирования заключается в том, что длина лунного цикла не кратна месячным интервалам. Если, например, сопоставить даты третьего лунного цикла за разные годы, то они будут различаться между собой. Это приведет к неопределенностям и при выборе подходящего для анализа лунного цикла, а также к возможным погрешностям, связанным с разным влиянием сезонных эффектов на один и тот же цикл, приходящийся на разные даты.

Агрегация с месячным окном учитывает каждый из трех критериев: эффект лунного цикла, сезонные колебания и вариацию облачного покрова. Месячные агрегаты VBD детекций почти такие же, как и агрегаты, соответствующие лунному циклу (рисунок 6.4). Месячное агрегирование устраняет эффекты лунного цикла и предоставляет достаточный временной интервал для накопления необходимого числа свободных от облачности наблюдений. Месячная агрегация упрощает выбор контрольных периодов, используемых для расчета VCI-индекса, и как правило хорошо соответствует продолжительности сезонных ограничений, которая обычно кратна полному месяцу. При месячной агрегации нет неопределенности при выборе данных для расчета VCI-индекса. Исходя из этих соображений, для расчета VCI-индекса мы используем месячные агрегаты. Следует отметить, что ежемесячная временная агрегация широко используется в экономических анализах для устранения сезонных воздействий на экономические и финансовые данные (Burkhauser et al., 2000; Neidell, 2004; Garza Gil et al., 2005).

#### *Результаты апробации индекса эффективности ограничения рыбного промысла*

Результаты апробации включают таблицы с графиками и картографические изображения. На графиках в виде вертикальных колонок показаны месячные агрегаты VBD детекций, рассчитанные за период с апреля 2012 года по январь 2018 года, в то время как значения VCI-индекса отображены в виде точек. Ось и шкала VBD детекций находится слева, ось и шкала VCI-индекса - справа. Для статистической проверки значения VCI-индекса вычисляются не только для периода действия ограничений рыбной ловли, но и для периода, предшествующего введению ограничений (и они также приведены на графике). Средние значения VCI-индекса для периода действия ограничений и предшествующего периода отмечены горизонтальными линиями. Картографические изображения, отображающие значения grid-ориентированного VCI-индекса, показывают области с высокой эффективностью анализируемого ограничения (VCI-индекс 90-100) - зеленым цветом, и области с низкой эффективностью ограничения (отрицательные значения VCI-индекса) - красным цветом.

##### **1. Авария у берегов Вьетнама**

На рисунке 6.5 приведены результаты анализа для Вьетнама. Мы рассчитали VCI-индекс для каждой из 4 пострадавших провинций в течение периода ограничения рыбной ловли и в течение контрольного периода, в качестве которого мы выбрали те же месяцы на протяжении 4 предшествующих лет (2012-2015 гг.). В VBD данных можно

видеть явно выраженный сезонный цикл с большим количеством детекций в сухой сезон, который длится с апреля по август, и малым числом детекций в дождливый сезон, продолжающийся как правило с октября по январь. Такое сокращение детекций в дождливые месяцы связано, вероятно, с двумя факторами: более плотным облачным покровом, затрудняющим спутниковый мониторинг, и штормами в это время года. В то же время, VCI-индекс разработан так, чтобы минимизировать эффекты сезонности за счет использования в качестве контрольного периода тех же месяцев, но за другие годы. Два самых высоких значения VCI-индекса в период ограничения пришлись на первый месяц действия запрета для двух провинций: Thua Thien-Hue (VCI-индекс 70%) и Quang Binh (VCI-индекс 44%). Два месяца VCI-индекс принимал отрицательные значения для провинций Ха-Тинь и Куанг-Три. Несмотря на четыре этих отрицательных значения, среднее значение VCI-индекса по всем 4 провинциям за все время действия ограничений составил 18%. При этом среднее значение VCI-индекса в течение контрольного периода составило 3,6%. Статистическое тестирование показало, что среднее значение VCI-индекса в течение периода действия ограничений было выше, чем в течение контрольного периода (таблица 6.3). Таким образом, проведенный при помощи VCI-индекса статистический анализ показал, что ограничения, введенные Вьетнамом, привели к ожидаемому снижению объема рыбного промысла, но это снижение было весьма умеренным (менее 40% почти по всем провинциям).

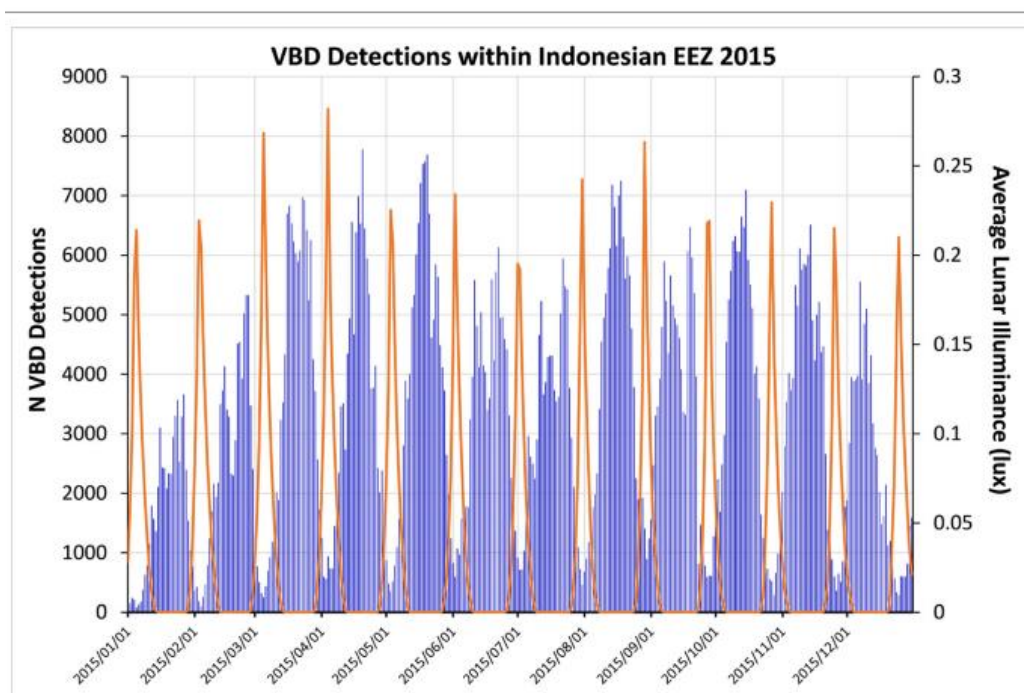


Рисунок 6.3 – Динамика во времени яркости лунной засветки и среднего числа VBD детекций в районе Индонезии

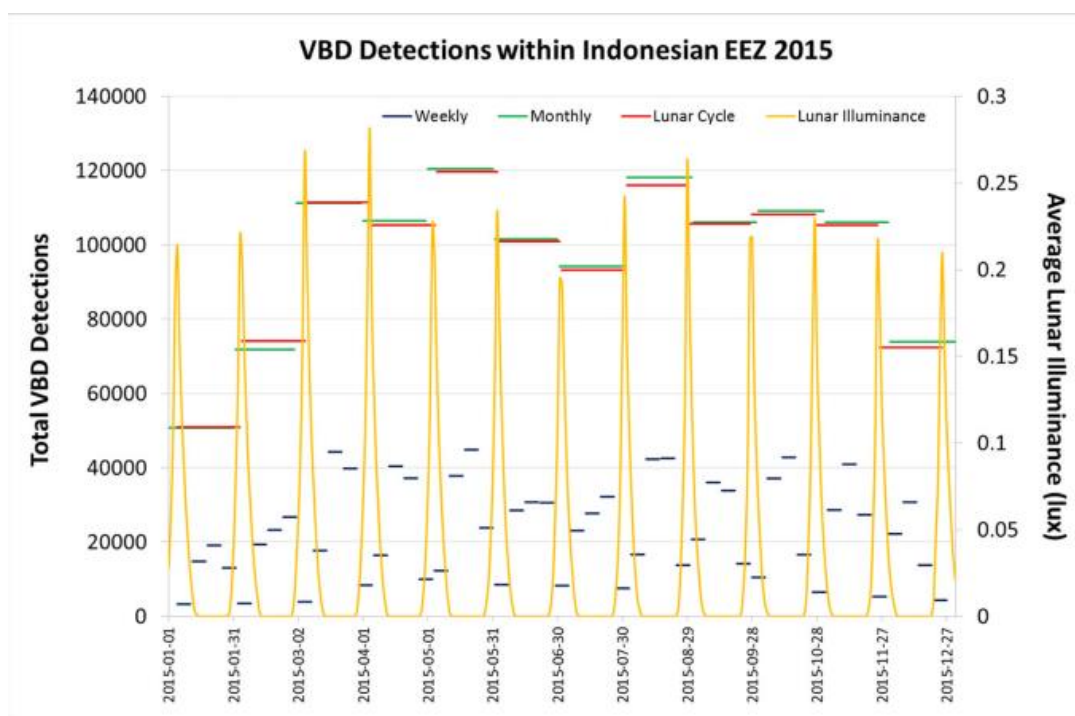


Рисунок 6.4 – Результат агрегирования числа VBD детекций по времени с различным размером окна агрегации

## 2. Запрет на вылов пятнистой ставриды у острова Палаван

Как и в случае с Вьетнамом, табличный VCI-индекс, рассчитанный для региона Палавана, демонстрирует четкие сезонные колебания: в сухое время года (с февраля по апрель) детектируется значительно больше рыболовных судов, чем в сезон дождей, особенно в его завершении (с октября по январь) (рис. 6.6). Ограничение вылова рыбы было введено в 2015 году и продолжалось с ноября по январь. После этого, оно приняло форму ежегодного и было повторено в 2016-2017 годах. Значения VCI-индекса были относительно постоянными в сезоне 2015 года (39-49%). Наиболее сильные перепады VCI-индекса среди всех трех лет пришлось на 2016 год. На этот же год пришлось самое низкое значение VCI-индекса в течение всех трех лет действия ограничения (27% в ноябре). Самое высокое значение VCI-индекса за время действия ограничений (94%) пришлось на декабрь 2017 года. За период с 2015 по 2017 годы прослеживается тенденция к повышению эффективности действия ограничений. Самое низкое значение VCI-индекса (41%) было зафиксировано в ноябре 2016 года, а самое высокое - в следующем месяце - в декабре и составило 79%. Среднее значение VCI-индекса в период действия ограничений составило 58%, в то время как среднее значение VCI-индекса в контрольный период равно 3,36%. Статистическое тестирование методом  $t_2$  показало, что значения VCI-индекса в период действия ограничений превышают аналогичные значения, полученные для контрольного периода (таблица 6.3). Все это подтверждает статистически значимый эффект от введения ограничений в районе острова Палаван.

Grid-ориентированный VCI-индекс, рассчитанный для ограничения рыбной ловли у острова Палаван, представлен на рисунке 6.7. Перед вычислением VCI-индекса мы провели пространственную агрегацию данных в скользящем окне размером  $12 \times 12$  ячеек. На рисунке зеленым цветом выделены регионы, где VCI-индекс дает значения 90% и более, а регионы, в которых VCI-индекс принимает отрицательные значения, выделены красным цветом. Из рисунка видно, что в центральной и южной частях анализируемого региона введение ограничений показало высокую эффективность. На севере в 2015-2016 годах отчетливо прослеживалась устойчивая зона отрицательных значений VCI-индекса, расположенная к северу от острова Бусуанга (Филиппины). Это привело к тому, что Бюро рыболовства и водных ресурсов (BFAR) активизировало борьбу с браконьерством в этой зоне в 2017 году. И это дало результат: значение VCI-индекса увеличилось в этой зоне в 2017 году.

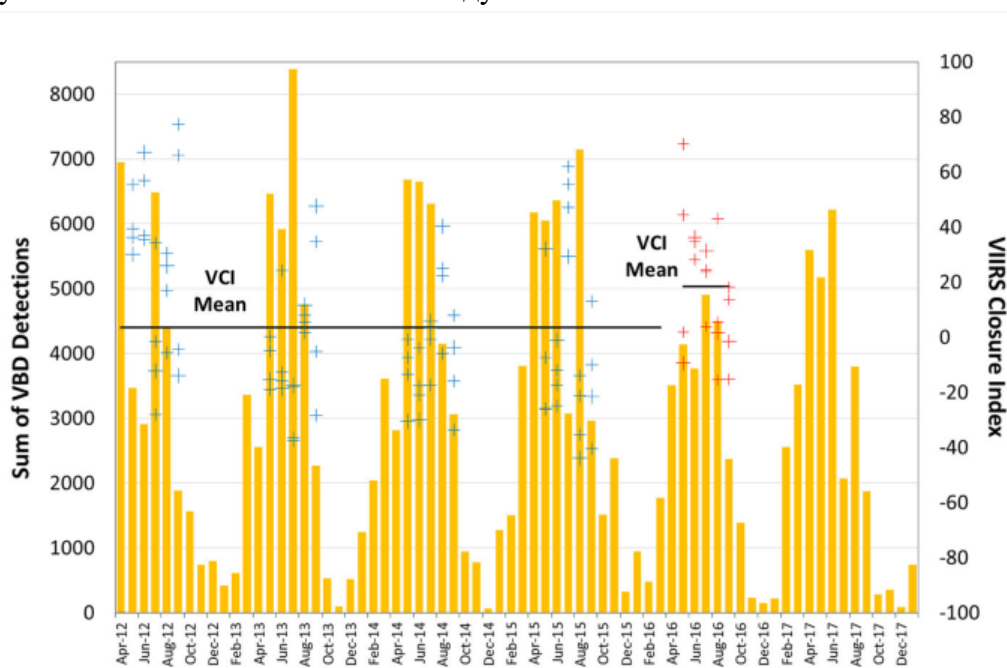


Рисунок 6.5 – Результаты анализа VCI-индекса для ограничения рыбной ловли у берегов Вьетнама в 2016 году

Таблица 6.3. Статистический анализ VCI-индекса в районах ограничения рыбного промысла

Vietnam	Reference	Closure	Palawan	Reference	Closure	Cauayan	Reference	Closure
Samples	80	20	Samples	9	9	Samples	24	46
Mean	3.58	18.35	Mean	3.36	58.62	Mean	0.98	74.24
Std. Dev.	29.28	22.21	Std. Dev.	29.65	22.06	Std. Dev.	35.27	26.02
Std. Error	4.97		Std. Error	7.35		Std. Error	3.84	
Test statistic		2.97	Test statistic		7.52	Test statistic		19.08
Result	Reject null hypothesis		Result	Reject null hypothesis		Result	Reject null hypothesis	
Conclusion	VCI higher during closure		Conclusion	VCI higher during closure		Conclusion	VCI higher during closure	
Confidence level		99.8%	Confidence level		99.9%	Confidence level		99.9%



### 3. Муниципальные воды г. Кауаян (Филиппины)

На рисунке 6.8 приведена гистограмма, показывающая ежемесячную сумму VBD детекций в районе действия Кауаянского ограничения на лов рыбы. Мы рассчитали значения VCI-индекса независимо для отдельных месяцев (на графике представлены точками, соединенными красной и зеленой ломаными линиями), а также их среднее значение, рассчитанное до начала кампании по борьбе с браконьерством, стартовавшей в 2014 году. В качестве контрольного периода выбраны 24 месяца до начала кампании (с 2012 по 2014 год). Анализируемый период усиленной охраны водных ресурсов составил 46 месяцев. Среднее значение VCI-индекса в период проведения кампании составил 74%, в то время как в течение контрольного периода он был равен 1%. Статистическое тестирование показало, что среднее значение VCI-индекса в период усиленной охраны рыбных ресурсов превышает аналогичное значение, рассчитанное для предшествующего периода (таблица 6.3). Это говорит о том, что VCI-индекс подтвердил эффективность введенных ограничений, причем для большинства анализируемых месяцев значение VCI-индекса составило более 75%. Эффективность падала на короткие периоды (2-3 месяца в году): июнь-август 2014 года, июль-сентябрь 2015 года, июль-август 2016 года.

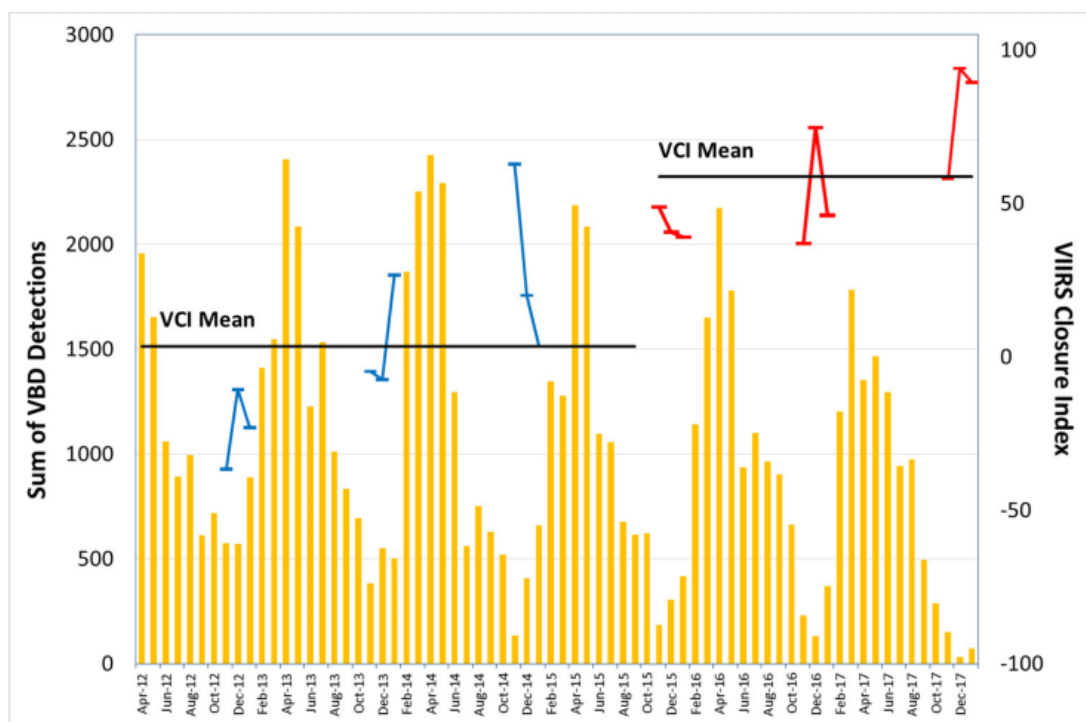


Рисунок 6.6 – Результаты анализа VCI-индекса для ограничения рыбной ловли у берегов острова Палаван в 2015-2017 годах

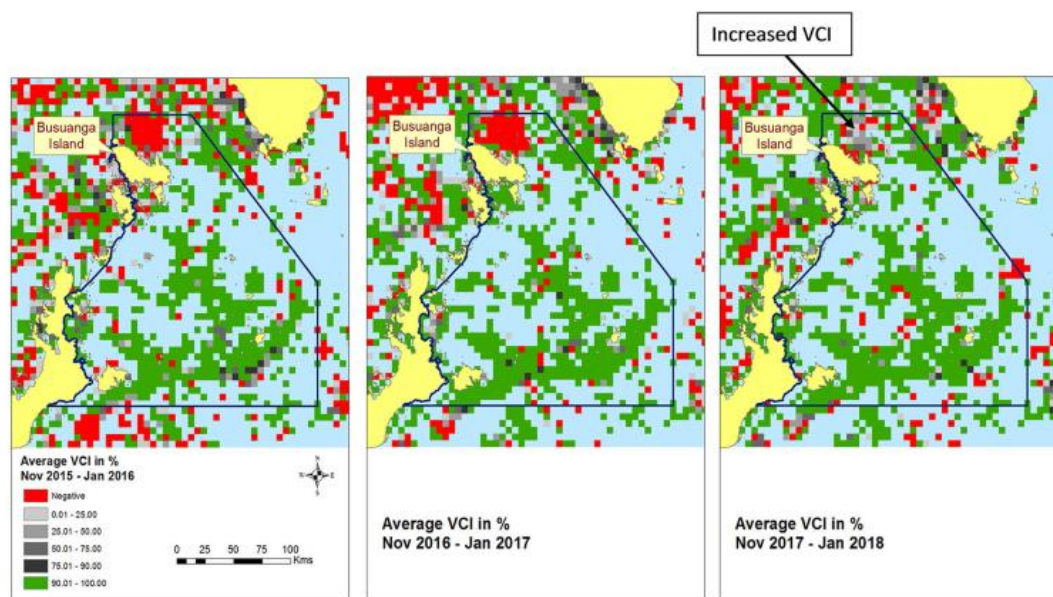


Рисунок 6.7 – Годовые данные по grid-ориентированному VCI индексу у острова Палаван за 2015-2017 года

На рисунке 6.9 приведен годовой grid-ориентированный VCI-индекс, рассчитанный для Кауаянского ограничения. Перед вычислением VCI-индекса мы провели пространственную агрегацию данных в скользящем окне размером  $2 \times 2$  ячейки. В ходе анализа мы также охватили муниципальные воды муниципального округа Илог, который не принимал участия в кампании по усилению контроля за соблюдением ограничения. VCI-индекс, рассчитанный за 2014 год, охватывает период с апреля по декабрь. Как видно из рисунка, введенные меры продемонстрировали высокую эффективность для большинства рыболовных зон, принадлежащих Кауаянскому округу (90-100%). Однако, в каждом году в период кампании появляются полосы более низких значений VCI-индекса. Эта зона относительно широкая в 2014 году, а в 2015-2017 годах ее ширина составляет 1-2 ячейки сетки. По данным VCI-анализа для территории Илога в 2014 и 2015 годах, была выявлена высокая эффективность введения ограничений, особенно во внутренних частях рыболовной зоны. Только небольшой участок около внешней границы зоны выпадает из этой закономерности. И такая высокая эффективность продолжила сохраняться в 2016-2017 годах. Это означает, что кампания по усилению контроля за соблюдением запрета, проводимая властями Кауяна, положительно сказалась не только на самом Кауаянском муниципальном округе, но и на соседнем Илоге, даже несмотря на то, что сам Илог никаких мер не предпринимал. Возможное объяснение этому факту в следующем. Наибольший вылов рыбы в Илоге приходится на небольшой промысловый участок, расположенный на пересечении границ муниципальных вод Кауяна и Илог. Причем выловом на этом участке занимается небольшое число компаний. Вероятно, столкнувшись с усилением контроля на данном участке со стороны Куаяна, компании уменьшили свой вылов и на той его территории, которая подпадает под юрисдикцию округа Илог.

На рисунке 6.8 мы видим существенные ежегодный провалы значений VCI-индекса в районе августа. Т.е. в августе ограничение нарушается чаще, чем в другие месяцы. Чтобы исследовать этот факт мы проанализировали распределение VBD-детекций, полученных по августовским данным. Оказалось, что подавляющее большинство “детекций-нарушителей” запрета располагаются вдоль внешней границы охраняемой зоны. Но мы знаем, что пиксель спутникового изображения в DNB канале, который используется алгоритмом VBD, имеет форму квадрата, который в надире покрывает участок поверхности Земли размером  $742 \times 742$  м. При этом всему пикселю ставится в соответствие географические координаты той точки, которая попадает в центр пикселя. Именно эту координату использует алгоритм VBD, когда присваивает детекции географические координаты. Таким образом, реальное положение детектируемого объекта может быть любым внутри той зоны, которая покрывается одним пикселем, - детекции все равно будут присвоены географические координаты центральной точки пикселя. В связи с этим требуется проверить еще одну гипотезу, связанную с августовскими всплесками детекций внутри зоны запрета: возможно суда находились за пределами границы (т.е. не нарушали запрет), но их детекциям были поставлены в соответствие координаты центров пикселей, которые находились уже внутри охраняемой зоны? Чтобы проверить это мы взяли отдельные VBD детекции, полученные в пикселях, центры которых проецируются вблизи внешней границы охраняемой зоны. Мы нанесли на карты центры этих пикселей и их границы. Результат представлен на рисунке 6.10. По рисунку видно, что большинство пикселей, соответствующих “детекциям-нарушителям” целиком проецируются на зону действия запрета, а значит суда, соответствующие данным детекциям, действительно, нарушили границы. Возможно, в августе складываются такие условия лова, что некоторые рыболовы готовы рискнуть и нарушить введенный запрет.

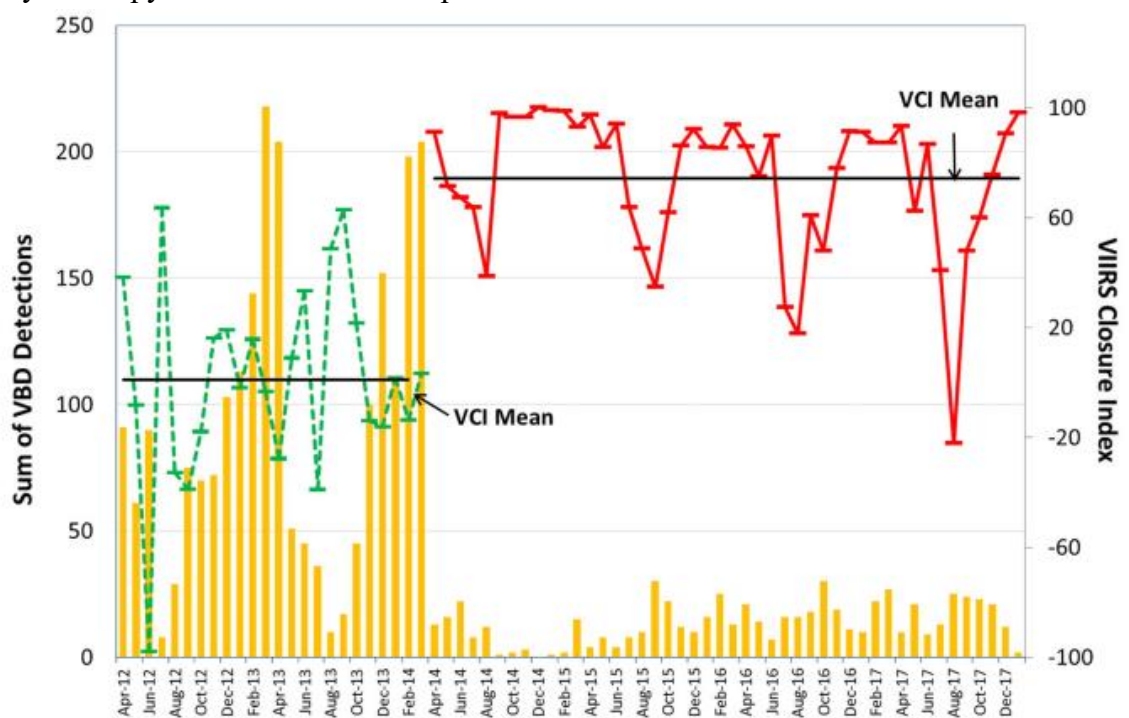


Рисунок 6.8 – Результаты анализа VCI-индекса для ограничения рыбной ловли в муниципальных водах округа Кауаян

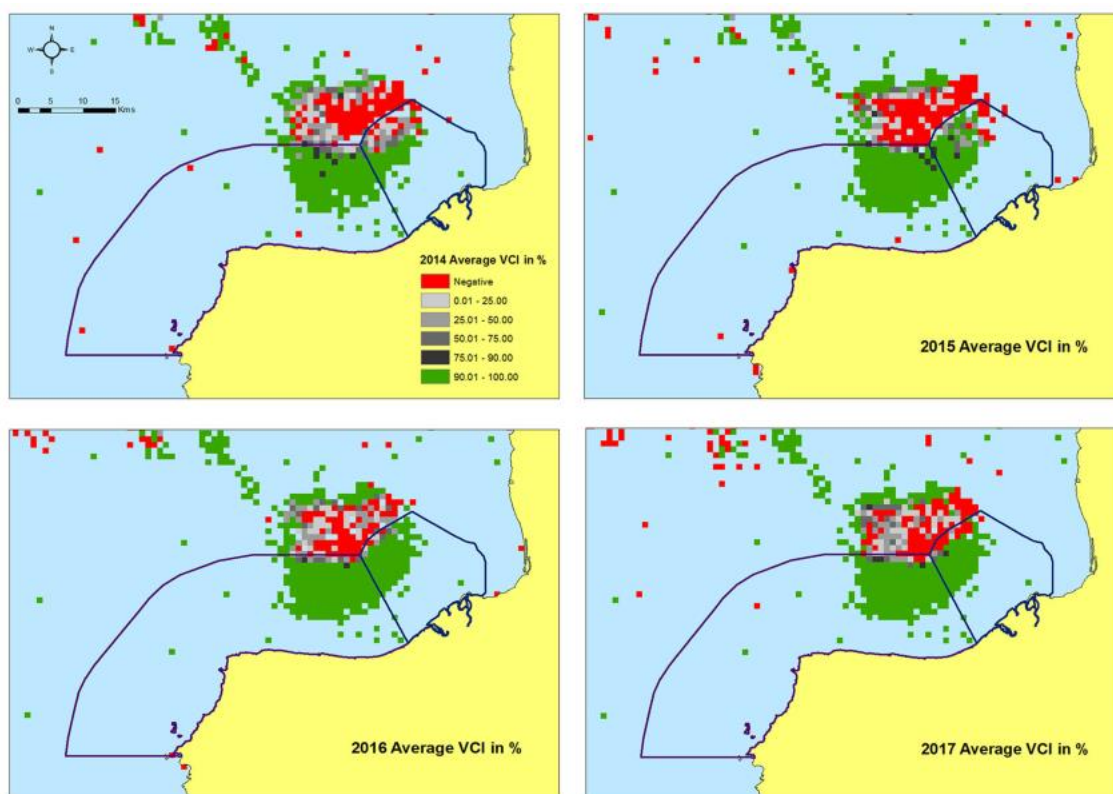


Рисунок 6.9 – Grid-ориентированный VCI-индекс, рассчитанный для муниципальных вод Кауаяна и Илога за 2014-2017 года

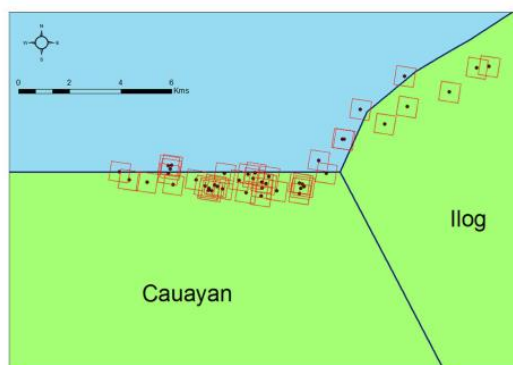


Рисунок 6.10 – Центры и границы пикселей DNB канала VIIRS, лежащие рядом с границей муниципальных вод округов Кауаян и Илог, в которых были зарегистрированы VBD детекции

## **ГЛАВА 7. ИНСТРУМЕНТ ДЛЯ АНАЛИЗА ДАННЫХ «GOOGLE EARTH ENGINE ASSESSMENT TOOL»**

### **7.1 Технические аспекты**

Основываясь на отзывах пользователей в Индонезии, Таиланде и Филиппинах, мы разработали веб-приложение для визуализации и анализа базы данных VBD, которое назвали “Google Earth Engine Assessment Tool”.

В основу разработанного приложения положено программное средство Google Earth Engine (GEE). GEE — это бета-версия ГИС-сервиса, предоставляемого компанией Google, который позволяет пользователям использовать язык Java Script для удобного взаимодействия с большими объемами данных. Пользователям не нужно задумываться об объеме данных или технических характеристиках локальных вычислительных машин, так как все операции GEE выполняются в облаке через обычный веб-браузер. На текущий момент связь между базой данных VBD и Google Earth Engine является статической. Чтобы использовать GEE для визуализации, требуется произвести ряд манипуляций:

1. извлечь подмножество VBD детекций для исследуемого интервала времени (обычно один месяц);
2. перевести извлеченные данные в формат растрового изображения;
3. загрузить полученный растр в Google Earth Engine;
4. написать сценарий анализа данных на языке JavaScript;
5. опубликовать веб-ссылки на сценарий и растровые изображения, которые будут использоваться внешними пользователями.

Альтернативный подход, использующий разработанный специальный тайл-сервер, который напрямую связан с базой данных VBD, описан в главе 4. Он имеет ограниченное число инструментов для ГИС-анализа, но его можно использовать для высокоразрешающей визуализации, совмещенной с картами OpenStreet Maps и дневными спутниковыми снимками высокого разрешения DigitalGlobe.

Далее описаны функциональные возможности и примеры использования разработанного веб-приложения для визуализации статистики, полученной на более чем пятилетнем интервале VBD-данных.

### **7.2 Обзор приложения**

Перед созданием приложения, мы зафиксировали набор ключевых функций, которые необходимы для статистического анализа детекций судов в заданных административных регионах и которые должно поддерживать разрабатываемое приложение:

1. Расчет и визуализация числа VBD детекций на квадратный километр для выбранного региона.
2. Расчет и визуализация пространственного распределения VBD детекций.
3. Визуализация количества пролетов спутника по итогам которых получены детекции.

4. Визуализация распределения значений яркости пикселей в DNB диапазоне

5. Построение графиков временных рядов значений произвольных атрибутов по заданному региону.

6. Построение гистограмм сравнения значений произвольных атрибутов для выбранного региона

Как результат, перечисленные выше функции были интегрированы в виде пяти различных режимов работы приложения:

1. Режим “Coverage”. В данном режиме отображается карта качества покрытия исследуемой территории спутниковыми снимками. Цветовая палитра показывает число наблюдений из космоса для каждой точки растра. Векторные границы, предварительно выбранные пользователем, отображаются на карте для географической привязки. Гистограммы показывают суммарное число космоснимков для каждой выделенной области.

2. Режим “Grid Radiance”. В данном режиме отображается цветовая карта. Цветовая палитра соответствует среднему значению яркости VBD детекций, зафиксированных в данной ячейке. Соотношение интегрального показателя яркости по всем регионам представляется в виде гистограммы.

3. Режим “Grid n\_detect”. В данном режиме для каждой ячейки регулярной сетки вычисляется суммарное число VBD детекций, зафиксированных в соответствующем регионе. Полученное значение наносится на карту в виде цветовой точки, цвет которой зависит от величины полученного значения. Соотношение между суммарным числом детекций в выбранных пользователем регионах представлено в виде гистограммы. Кроме того, для каждого из выбранных регионов строится график временного ряда суммарного числа детекций.

4. Режим “Regional Density n\_detect”. В этом режиме отображается цветовая карта, отражающая плотность VBD детекций в выбранных пользователем регионах из расчета числа детекций на квадратный километр. Соотношение средней плотности детекций между регионами представлено в виде гистограммы. Дополнительно визуализируется график динамики суммарного числа детекций во времени для выбранного региона.

5. Режим “Regional Sum Radiance”. В данном режиме представлена цветовая карта, где цвет региона зависит от суммарной яркостью пикселей в DNB диапазоне по этому региону. Суммарная яркость пикселей по каждому региону также представлена в виде гистограммы, а ее изменение во времени - в виде графиков временных рядов.

6. Режим “Regional Sum n\_detect”. В данном режиме представлена цветовая карта, в которой цвет каждого региона зависит от суммарного числа детекций, зафиксированных в этом регионе. Эти же данные по каждому региону представлены на гистограмме, а их динамика во времени - на графике временного ряда.

### 7.3 Работа с приложением

На рисунке 7.1 показана начальная страница приложения.



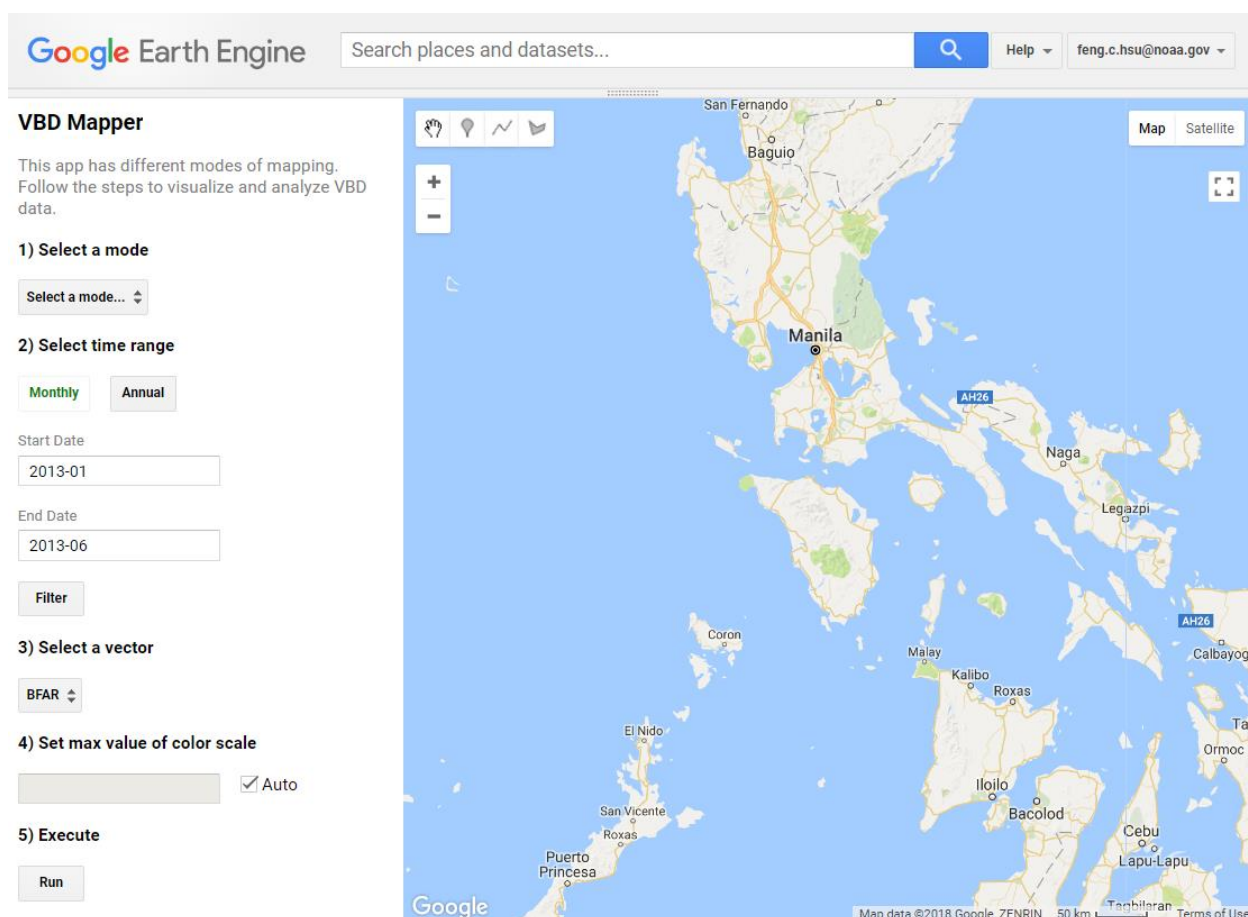


Рисунок 7.1 – Начальная страница разработанного приложения “Google Earth Engine Assessment Tool”

В левой панели отображены параметры и настройки, которые пользователь может менять.

1. Выбор режима. Может быть выбран один из 6 режимов, описанных ранее в этом разделе (рисунок 7.2).

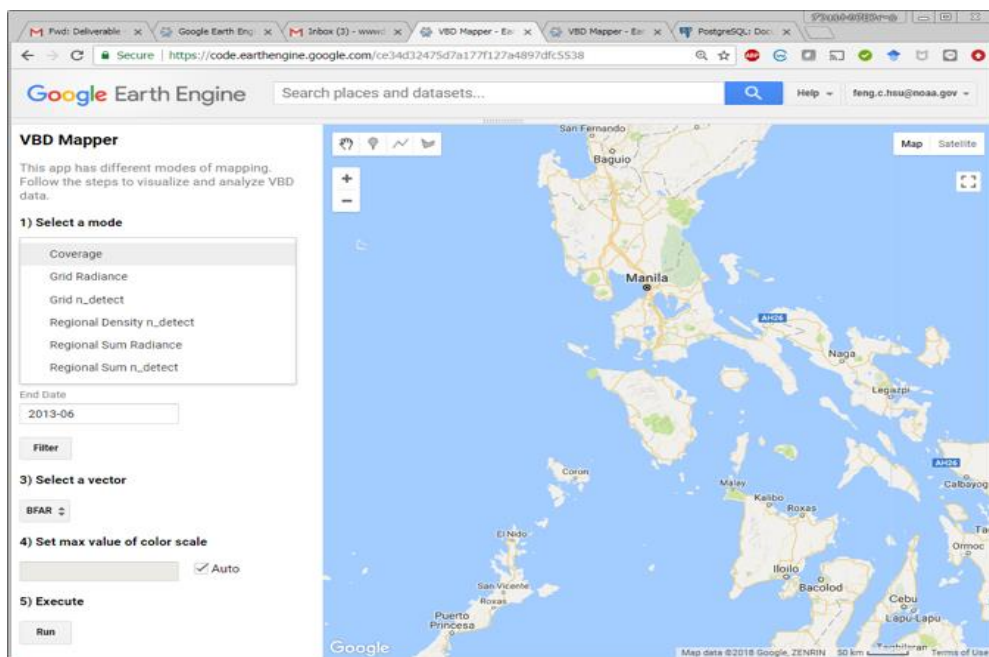


Рисунок 7.2. Выбор режима функционирования приложения

2. Выбор анализируемого временного интервала. В этом же разделе настроек следует указать временной шаг, который может быть как месячным, так и годовым. Пользователь может нажать на кнопку “Filter” и посмотреть набор выбранных для обработки файлов детекций (рисунок 7.3).

## 2) Select time range

Monthly

Annual

Filter

Start Date

2013-01

Found 6 files

Hide List

End Date

2013-06

Filter

Found 6 files

Show List

[1] VBD\_npp\_20130101-20130131\_phl\_pc\_v23\_c201707151800\_cvg

[2] VBD\_npp\_20130201-20130228\_phl\_pc\_v23\_c201707151800\_cvg

[3] VBD\_npp\_20130301-20130331\_phl\_pc\_v23\_c201707151800\_cvg

[4] VBD\_npp\_20130401-20130430\_phl\_pc\_v23\_c201707151800\_cvg

[5] VBD\_npp\_20130501-20130531\_phl\_pc\_v23\_c201707151800\_cvg

[6] VBD\_npp\_20130601-20130630\_phl\_pc\_v23\_c201707151800\_cvg

Рисунок 7.3. Выбор анализируемого временного интервала

3. Выбор региона. Например, для Филиппин доступно 6 наборов регионов для выбора:

А. CLOSURES: ограничения рыбной ловли (в наборе задано 4 региона).

- В. LGU-BFAR: Независимые муниципальные воды (LGU), выбираемые по обозначениям BFAR.
- С. LGU-BFAR (объединенный вариант): объединение всех LGU для каждого BFAR.
- Д. LGU-Province: LGU, выбираемые по обозначениям провинций. LGU-Province.
- Е. LGU-Province (объединенный вариант): объединение всех LGU для каждой провинции.
- Ф. NIPAS MPA: Морская охраняемая территория, обозначенная как “NIPAS”.

4. Определение числового значения максимума цветовой палитры. По умолчанию числовое значение максимума цветовой палитры определяется автоматически. В качестве такого значения берется максимальное среди всех значений, по которым строится цветовая карта. Однако это не всегда удобно. Отклонение в единичном пикселе может при таком подходе сильно сместить шкалу в область высоких значений. В результате все значения (за исключением точки выброса) попадут примерно в одну цветовую категорию. Чтобы этого избежать, пользователь может снять флажок “Auto” и вручную задать значение, которое будет на цветовой шкале соответствовать цвету с максимальным значением. При этом все значения, превышающие заданное, также будут отображаться этим же цветом.

5. Выполнение. Кнопка “Run” запускает процесс расчета и визуализации результатов.

## **ЗАКЛЮЧЕНИЕ**

В результате проделанной работы были получены следующие результаты:

— Создана программная реализация алгоритма детектирования судовых огней. Система реализована на языках программирования Java, MATLAB и IDL, C++, которые оркестрируются скриптом на языке программирования Python в соответствии с парадигмой MapReduce.

— Программная реализация оптимизирована для ускорения работы наиболее ресурсоемких алгоритмов участков кода, в том числе за счет их адаптации для параллельной работы на многоядерных вычислителях, например графических процессорах. Оптимизация позволила ускорить работу программы в десятки раз.

— Разработана база данных детекций судовых огней. База данных разработана на основе СУБД с открытым исходным кодом PostgreSQL и ее пространственного расширения PostGIS и оптимизирована для выполнения поиска по географическим координатам и пространственным ограничениям.

— Разработано веб-приложение для визуализации и статистического анализа распределения судовых огней на основе tile-сервера. Разработанное приложение позволяет создавать интерактивные карты детекций при помощи специально

разработанного tile-сервера, а также выполнять интерактивный статистический анализ и строить графики при помощи программного инструмента Jupyter Notebook.

— Разработано веб-приложение для визуализации и анализа базы данных VBD, на основе Google Earth Engine (GEE). Разработанное приложение поддерживает широкий набор инструментов для регионального анализа распределения детекций судов и картографической визуализации его результатов.

— Разработан индекс эффективности введения ограничений на рыбный промысел. При помощи разработанного индекса проведен анализ трех регионов, в которых были введены ограничения на лов рыбы.

— Проведен анализ распределения VBD детекций на морской территории, граничащей с Дальним востоком (Камчатский край, Магаданская и Сахалинская область, Хабаровский край). Анализ показал, что основной объем детекций на море в исследуемом районе распределяется между: огнями рыболовных судов, ведущих рыбный промысел; бортовыми огнями судов в районах портов и рейдов; морскими платформами и обслуживающими их судами; воздушными авиалиниями. Анализ показал, что VBD детекции, соответствующие бортовым огням рыболовных судов, хорошо согласуются с основными промысловыми зонами Охотского моря и Курильских островов как с точки зрения их пространственного расположения, так и с точки зрения сезонности (на примере зоны промысловой добычи сельди).

## ЛИТЕРАТУРА

- Bastiaans, M. On the sliding-window representation in digital signal processing. 1985. IEEE Transactions on Acoustics, Speech, and Signal Processing. Vol. 33. Issue 4.
- Burkhauser, R. V., Couch, K. A., and Wittenburg, D. C. (2000). Who minimum wage increases bite: an analysis using monthly data from the SIPP and the CPS. *South. Econ. J.* 67, 16–40. doi: 10.2307/1061611
- Campello, Ricardo JGB, Davoud Moulavi, and Joerg Sander. 2013. “Density-Based Clustering Based on Hierarchical Density Estimates.” In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 160–72. Springer. doi:10.1007/978-3-642-37456-2\_14
- Cho, K.; Ito, R.; Shimoda, H.; Sakata, T. Fishing fleet lights and sea surface temperature distribution observed by DMSP/OLS sensor. *Int. J. Remote Sens.* 1999, Vol. 20, pp. 3–9.
- Cox, C.; Munk, W. Measurement of the Roughness of the Sea Surface from Photographs of the Sun’s Glitter, *J. Opt. Soc. Am.* 1954, pp. 838–850
- Croft, T.A. Nighttime Images of the Earth from Space. *Sci. Am.* 1978, Vol. 239, pp. 86–98.
- Croft, T.A.; Colvocoresses, A.P. The Brightness of Lights on Earth at Night, Digitally Recorded by DMSP Satellite; USGS Open File Report. U.S. Geological Survey: Reston, VA, USA, 1979.
- Delmas, P., and Georgy Gimel’farb. Image Matching: Correlation. COMPSCI 773 S1C. URL: <https://www.cs.auckland.ac.nz/courses/compsci773s1c/lectures/CS773S1C-CorrelationMatching.pdf>
- Elvidge, C.D.; Baugh, K.E.; Zhizhin, M.; Hsu, F.-C. Why VIIRS data are superior to DMSP for mapping nighttime lights. *Proc. Asia-Pac. Adv. Netw.* 2013, 35, 62–69.
- Freund, R., and Wilson, W. (2003). *Statistical Methods*, 2nd Edn. Cambridge, MA: Academic Press.
- Garza Gil, M. D., Prada-Blanco, M., and Vazquez-Rodriguez, M. X. (2005). Estimating the short-term economic damages from the Prestige oil spill in the Galician fisheries and tourism. *Ecol. Econ.* 58, 842–849. doi: 10.1016/j.ecolecon.2005.09.009
- Gonzales, Rafael C. and Richard E. Woods. *Digital Image Processing*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.
- Janiczek, P. M., and J. A. Deyoung. 1987. Computer Programs for Sun and Moon Illuminance with Contingent Tables and Diagrams. U.S. Naval Observatory Circular 171. Doi: accessed on 12th April 2017. Washington, DC: U.S. Naval Observatory. Available online at: [http://aa.usno.navy.mil/publications/docs/Circular\\_171.pdf](http://aa.usno.navy.mil/publications/docs/Circular_171.pdf)
- Kiyofuji, H.; Saitoh, S. Use of nighttime visible images to detect Japanese common squid *Todarodes pacificus* fishing areas and potential migration routes in the Sea of Japan. *Mar. Ecol. Prog. Ser.* 2004, Vol. 276, pp. 173–186.
- Lee, J., South, A. B., and Jennings, S. (2010). Developing reliable, repeatable, and accessible methods to provide high-resolution estimates of fishing-effort distributions from vessel monitoring system (VMS) data. *Int. Counc. Explor. Sea* 67, 1260–1271. doi: 10.1093/icesjms/fsq010



- Lim J.S. Two-Dimensional Signal and Image Processing: Englewood Cliffs, NJ: Prentice Hall, 1990.
- Maxwell, M.R.; Henry, A.; Elvidge, C.D.; Safran, J.; Hobson, V.R.; Nelson, I.; Tuttle, B.T.; Dietz, J.B.; Hunter, J.R. Fishery Dynamics of the California market squid (*Loligo opalescens*), as measured by satellite remote sensing. *Fish. Bull.* 2004, Vol. 102, pp. 661–670.
- Miller, S.D.; Mills, S.P.; Elvidge, C.D.; Lindsey, D.T.; Lee, T.F.; Hawkins, J.D. Suomi satellite brings to light a unique frontier of nighttime environmental sensing capabilities. *Proc. Natl. Acad. Sci. USA* 2012, 109, 15706–15711.
- Miller, S.D.; Straka, W., III; Mills, S.P.; Elvidge, C.D.; Lee, T.F.; Solbrig, J.; Walther, A.; Heidinger, A.K.; Weiss, S.C. Illuminating the Capabilities of the Suomi National Polar-Orbiting Partnership (NPP) Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band. *Remote Sens.* 2013, Vol. 5, pp. 6717–6766.
- Neidell, M. J. (2004). Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma. *J. Health Econ.* 23, 1209–1236. doi: 10.1016/j.jhealeco.2004.05.002
- PhysOrg (2017). Vietnam Maintains Ban on Deep-Water Fishing in Polluted Area. Available online at: <https://phys.org/news/2017-05-vietnam-deep-water-fishing-polluted-area.html> (Accessed February 21, 2018).
- Rodhouse, P.G.; Elvidge, C.D.; Trathan, P.N. Remote sensing of the global light-fishing fleet: An analysis of interactions with oceanography, other fisheries and predators. *Adv. Mar. Biol.* 2001, Vol. 39, pp. 261–303.
- Schueler, C.F.; Lee, T.F.; Miller, S.D. VIIRS constant spatial-resolution advantages. *Int. J. Remote Sens.* 2013, Vol. 34, pp. 5761–5777.
- Straka, W.; Seaman, C.; Baugh, K.; Cole, K.; Stevens, E.; Miller, S.D. Utilization of the Suomi national polar-orbiting partnership (NPP) visible infrared imaging radiometer suite (VIIRS) day/night band for arctic ship tracking and fisheries management. *Remote Sens.* 2015, 7, 971–989.
- Terraserver.com. URL: <https://www.terraserver.com/> (дата обращения: 24.05.2018);
- Vu, C.T.; Phan, T.D.; Chandler, D.M. S3: A spectral and spatial measure of local perceived sharpness in natural images. *IEEE Trans. Image Process.* 2012, Vol. 21, pp. 934–945.
- Waluda, C.M.; Trathan, P.N.; Elvidge, C.D.; Hobson, V.R.; Rodhouse, P.G. Throwing light on straddling stocks of *Ilex argentinus*: Assessing fishing intensity with satellite imagery. *Can. J. Fish. Aquat. Sci.* 2002, Vol. 59, pp. 592–596.
- Waluda, C.M.; Yamashiro, C.; Elvidge, C.D.; Hobson, V.R.; Rodhouse, P.G. Quantifying light-fishing for *Dosidicus gigas* in the eastern Pacific using satellite remote sensing. *Remote Sens. Environ.* 2004, Vol. 91, pp. 129–133.
- Zhang, Q.; Schaaf, C.; Seto, K.C. The vegetation adjusted NTL urban index: A new approach to reduce saturation and increase variation in nighttime luminosity. *Remote Sens. Environ.* 2013, 129, 32–41.