

УДК 528.83

DOI: <https://doi.org/10.26456/2226-7719-2018-3-40-56>

МАКЕТ ПОТОКОВОЙ ОБРАБОТКИ ДАННЫХ SENTINEL-2

Л.А. Бочка¹, А.В. Кашницкий²

¹МГУ имени М.В. Ломоносова, Москва

²Институт космических исследований РАН, Москва

В данной работе проведен анализ продуктов Sentinel-2 различных уровней. Европейское космическое агентство предоставляет программное обеспечение для создания по распространяемым данным продуктов уровня L1C продукты L2A - Sen2Cor. В работе рассмотрена структура плагина, какие являются входными данными, какие данные на выходе. Результатом работы является программно-аппаратный комплекс потоковой обработки данных со спутников Sentinel-2, который позволяет в автоматическом режиме создавать информационные продукты уровня обработки L2A и внедрять данный продукт в действующие геоинформационные системы.

Ключевые слова: дистанционное зондирование Земли, данные Sentinel-2, обработка спутниковых данных, технологии работы с данными, атмосферная коррекция.

Мониторинг окружающей среды является важной составляющей современного мира. Он обеспечивает безопасное и комфортное проживание людей, позволяет понять климатические изменения и какие факторы им способствовали.

В рамках европейской программы глобального мониторинга окружающей среды и безопасности были разработаны спутники Sentinel-2, которые обеспечивают оптическими снимками высокого разрешения (10 – 20 м).

Миссия Sentinel-2 предлагает комбинацию следующих возможностей:

- 1) Систематический глобальный охват земной поверхности: от 56 южной широты и 84 северной широты, прибрежные воды и Средиземное море
- 2) Высокая частота: каждые 5 дней на экваторе и каждые 2-3 дня в средних широтах
- 3) Высокое пространственное разрешение: 10 м, 20 м и 60 м
- 4) Мультиспектральная информация с 13-ю каналами в видимой, ближней инфракрасной и коротковолновой инфракрасной части спектра

5) Широкий угол обзора: 290 км

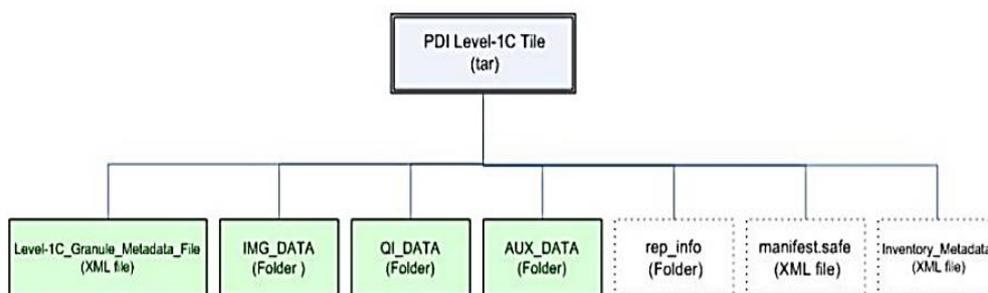
Оператором Sentinel-2 является Европейское космическое агентство (ESA), которое предоставляет данные по снимкам в открытом доступе для любого пользователя. Продукты Sentinel-2 предоставляются в виде гранул фиксированного размера. Размер гранул зависит от уровня обработки. Европейское космическое агентство обеспечивает различными уровнями. В данной работе будут рассмотрены уровни L1C и L2A. Уровень L1C – орто-трансформированный продукт, т.е проекция изображения получается с помощью системы DEM для исправления геометрических искажений, включает географические и радиометрические коэффициенты калибровки, а также привязку по времени и дополнительной информацией. Уровень обработки L2A – включает в себя атмосферную коррекцию, классификацию сцен и получение коэффициентов с нижних границ атмосферы. Дальше в работе будут подробнее освещены все типы продуктов.

На данный момент в действующих геоинформационных системах России в открытом доступе открыто распространяются данные по любой территории Северной Евразии и отдельные регионы мира информационные продукты уровня обработки, в то время, как ESA предоставляет в открытом доступе программное обеспечение для создания по распространяемым данным L1C продуктов уровня обработки L2A.

Цель работы – создание макета потоковой обработки спутниковых данных Sentinel-2. Все данные получаемые аппаратурой MSI систематически обрабатываются к уровню 1C центром по обработки получаемых данных (PDGS). Пользователям выпускаются только продукты уровня 1C и уровня 2A. Продукты уровня 2A генерируются либо PDGS с помощью плагина Sen2Cor, либо на стороне пользователя через Sentinel-2 Toolbox. [2]

Продукт Level-1C состоит из плиток 100x100 км² (орто-изображений в проекции UTM/WGS84). Продукт Level-1C является результатом использования цифровой модели рельефа (ЦМР) для проецирования изображения в картографической геометрии. Продукты Level-1C пересчитаны с постоянным земным расстоянием замера (GSD) 10, 20 и 60 м в зависимости от родного разрешения различных спектральных диапазонов. Продукты Level-1C включают данные с коэффициентами отражения на верхней границе атмосферы в 13 каналах для разного разрешения, а также дополнительно включают маски суши и воды, облачные маски и данные ECMWF (общий столбец озона, общий столбец водяного пара и среднее давление на уровне моря) [4]

Структура продукта отражена на рис. 1.



Р и с. 1. Структура L1C продукта [2]

PDI плитка уровня 1C содержит следующую информацию:

1. Level-1C_Tile_Metadata_File: файл метаданных XML, содержащий запрошенный уровень информации и ссылающийся на все элементы продукта, составляющие плитку;
2. IMG_DATA: папка, содержащая файлы данных изображений, сжатые с помощью алгоритма JPEG2000, по одному файлу на канал;
3. QI_DATA: папка, содержащая отчеты XML, включая проверки качества, файлы маски GML и файл JP2 PCI;
4. AUX_DATA: папка с ECMWF данные пересчитаны в проекции utm;
5. Inventory_Metadata.xml: файл, содержащий метаданные, необходимые для инвентаризации PDI;
6. manifest.safe: XML - безопасный файл содержащий метаданные
7. rep_info: папка, содержащая доступную схему XSD, которая проверяет компоненты PDI

Продукты уровня 1C регулярно геометрически уточняются. Однако они могут быть обработаны без уточнения, т.е. "по мере необходимости". Эти продукты можно определить по наличию выделенного флага в метаданных. Как правило, продукт этого уровня, включая все каналы составляет файл размером около 500 Мб. Sentinel-2 Level-2A обработка предоставляет продукты (отражение на нижней границе атмосферы) полученные на основе продуктов уровня 1C (отражение на верхней границе атмосферы). Каждый продукт Level-2A составлен плиток 100x100 km² в картографической геометрии (проекции UTM/WGS84). Продукция уровня 2A систематически генерируется в наземном сегменте по всей Европе, но производство будет расширено до глобального уровня к концу 2018 года. Поэтому для доступа к данным других регионов можно генерировать продукты через Sentinel-2 Toolbox, используя в качестве входных данных соответствующий продукт Level-1C (рис. 2).

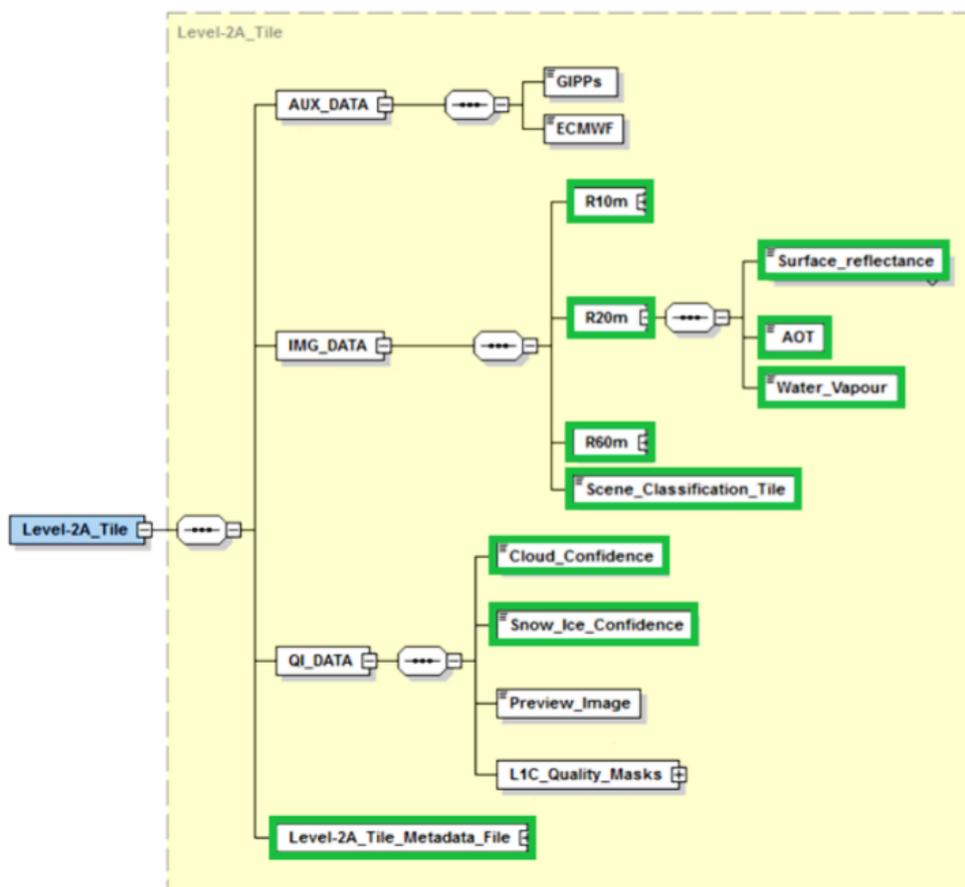


Р и с. 2: слева – уровень обработки 1С и справа – полученное изображение уровня 2А на основе 1С.

Продукты второго уровня представляют собой набор следующих данных для разного разрешения:

- синтез в естественных цветах;
- коэффициенты отражения на нижней границе атмосферы;
- карта аэрозольной оптической толщины;
- карта водяного пара;
- результат классификации (12 классов);
- показатели вероятности снега/льда;
- показатели вероятности облаков.

Структура продукта L2A строго основана на структуре продукта L1C с некоторыми различиями, которые выделены зеленым цветом на рис. 2. Основное отличие состоит в том, что папка IMG_DATA содержит три каталога: по одному для каждого разрешения на 60 м, 20 м и 10 м. Плитка классификация сцены доступна в корне папки IMAGE_DATA на 20 м или 60 м в зависимости от типа обработки, которая была выбрана.

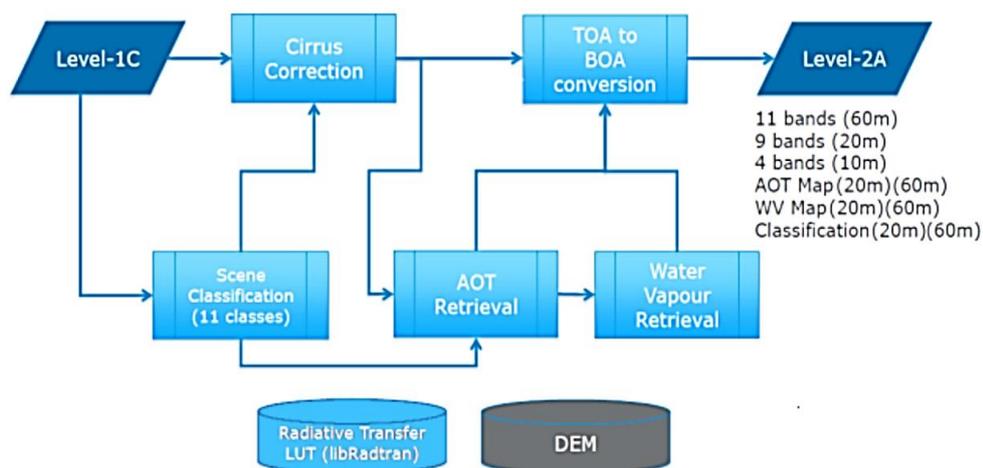


Р и с. 3. Структура продукта L2A

PDI плитка уровня 2A содержит следующую информацию:

1. Level-2A_Tile_Metadata_File: файл метаданных XML, содержащий информацию и ссылающийся на все элементы продукта, составляющие Tile;
2. IMG_DATA: папка, содержащая файлы данных изображений, сжатые с использованием алгоритма JPEG2000;
3. QI_DATA: папка, содержащая отчеты XML, включая проверки качества, файлы маски GML и файл PVI JP2;
4. AUX_DATA: папка, содержащая данные ECMWF, повторно сэмплированные в проекции UTM;
5. Inventory_Metadata.xml: файл, содержащий метаданные, необходимые для инвентаризации PDI.

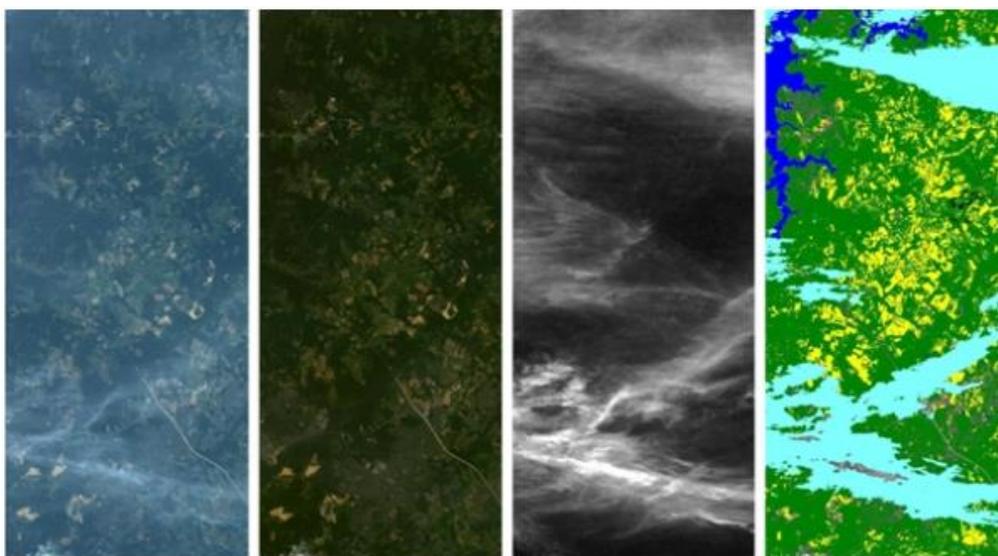
Обработка Level-2A применяется для гранул верхнего уровня атмосферы (TOA) - продуктов отражения ортоизображения уровня 1C.



Р и с. 4. Блок-схема алгоритма преобразования продуктов из одного уровня в другой

Обработка уровня 2A включает классификацию сцен и атмосферную коррекцию, применяемую к продуктам ортоизображения верхнего уровня атмосферы (TOA). На выходе получаем ортофотоплан отражения на нижней части атмосферы с примененными исправлениями продукта. Дополнительно на выходе карта аэрозольной оптической толщины (AOT), карта водяного пара (WV) и карта классификации сцены (SCM) вместе с индикаторами качества (QI) для вероятностей облака и снега на разрешении 60 м. Выходные изображения L2A будут пересчитаны и сгенерированы с одинаковым пространственным разрешением для всех каналов на основе запрошенного разрешения (10 м, 20 м или 60 м). В 10 м разрешении продукт содержит следующие спектральные каналы 2, 3, 4 и 8 и AOT карта. В 20 м продукт содержит 2 - 7, 8A, 11 и 12 каналы, а также AOT и WV карты. Продукт разрешением 60 м содержит все компоненты продукта разрешения 20 м и дополнительно в 1-м и 9-ом каналах. Алгоритм представляет собой комбинацию современных методов выполнения атмосферных коррекций (включая коррекцию перистых облаков [5]), которые были адаптированы к среде SENTINEL-2 вместе с модулем классификации сцен, описанным в [4]. Алгоритм классификации сцен позволяет обнаруживать облака, снег и тени облаков и генерировать классификационную карту, которая состоит из четырех различных классов для облаков (включая перистые), а также шести различных классификаций для теней, теней облаков, растительности, почв / пустынь, воды и снега. Алгоритм основан на серии пороговых тестов, которые используют в качестве входного коэффициента отражение TOA в качестве входного сигнала из спектральных полос SENTINEL-2. Тип аэрозоля и видимость или оптическая толщина атмосферы получены с использованием алгоритма плотной темной

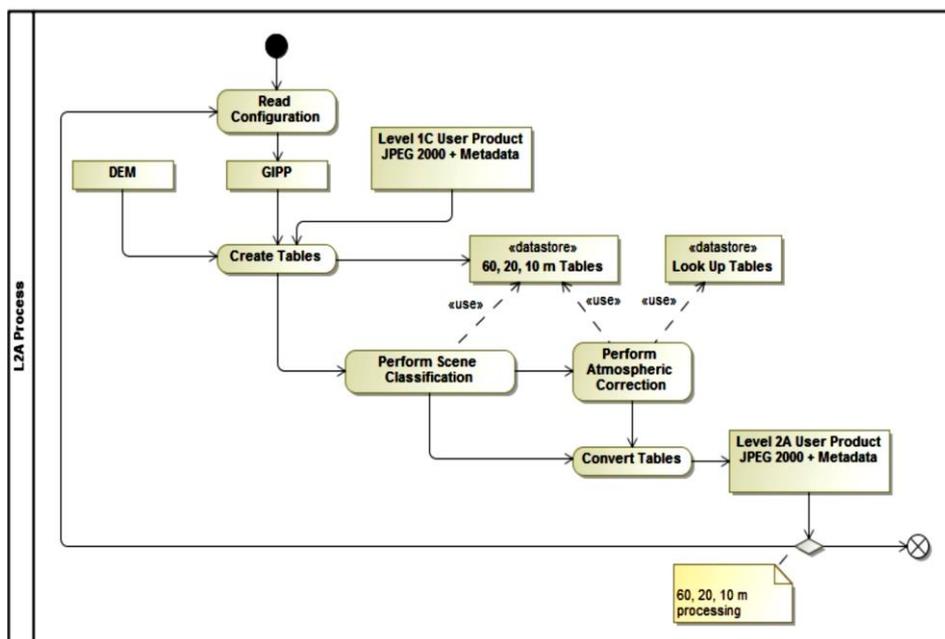
растительности (DDV) [2]. Этот алгоритм требует, чтобы сцена содержала опорные области известного поведения отражения, предпочтительно DDV и водные объекты.



Р и с. 5: слева направо: (1) моделируемый уровень-1С toa reflectance input Image, (2) cirrus и атмосферно скорректированный уровень-2А BOA reflectance image, (3) The cirrus band B10, (4) классификация сцены входного изображения Level-1С (данные AVIRIS, предоставленные НАСА) [10]

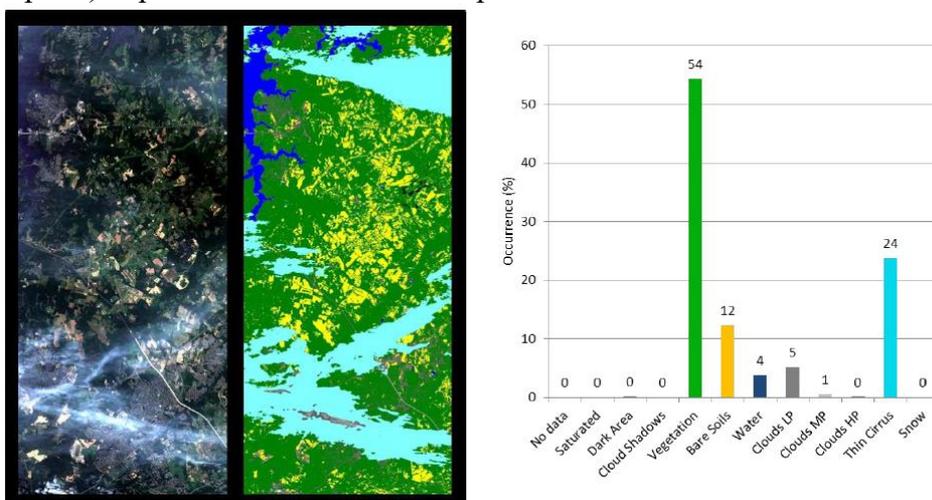
Для получения продуктов второго уровня Европейское космическое агентство использует разработанный плагин – Sen2Cor. Sen2Cor использует 2 главных вспомогательных источника данных: радиационные таблицы и цифровая модель высоты (DEM), которая не встроена в сам программный продукт, но может быть обеспечена пользователем в формате DTED. Пользователь может также положиться на DEM - значения по умолчанию.

Плагин Sen2Cor содержит высокоуровневую архитектуру и состоит из 10 основных модулей. На рисунке 9 приведена графическое представление выполнения модуля L2A_Process, начиная с чтения конфигурационных данных и до получения продукта второго уровня.



Р и с. 9. Общее представление работы плагина Sen2Cor [3]

Алгоритм модуля классификации сцен позволяет обнаруживать облака, снег и тени облаков и генерировать классификационную карту, которая состоит из 4 различных классов для облаков (включая перистые), вместе с шестью различными классификациями для теней, теней облаков, растительности, почв / пустынь, воды и снега. Алгоритм основан на серии пороговых тестов, использующих в качестве входных данных отражение верхней части атмосферы от спектральных полос Sentinel-2. На рис. 10 ниже показаны результаты работы классификации сцен (правая сторона) на основе измененных тестовых данных (левая сторона). Представлено двенадцать различных классов.



Р и с.10. Классификация сцен [3]

Список генерируемых классов можно видеть в табл. 1.

Таблица 1

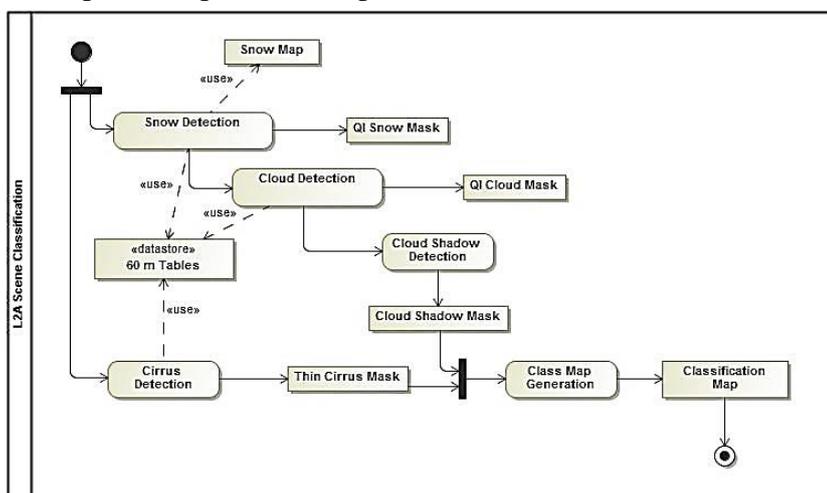
Список сцен для классифицированного изображения

Label	Classification
0	NO_DATA
1	SATURATED_OR_DEFECTIVE
2	DARK_AREA_PIXELS
3	CLOUD_SHADOWS
4	VEGETATION
5	BARE_SOILS
6	WATER
7	CLOUD_LOW_PROBABILITY
8	CLOUD_MEDIUM_PROBABILITY
9	CLOUD_HIGH_PROBABILITY
10	THIN_CIRRUS
11	SNOW

На основе полученных результатов генерируются соответствующие показатели качества вероятности снега и облака. Эти качественные показатели рассчитывают вероятность (0-100%) того, что поверхность земли будет перекрыта облаками или оптически толстым аэрозолем (льдом или снегом).

Обработка SC состоит из шести различных этапов (рис. 11):

1. Обнаружение снега;
2. Обнаружение облака;
3. Обнаружение теней в облаке;
4. Обнаружение перистых облаков;
5. Генерация карты классификации.

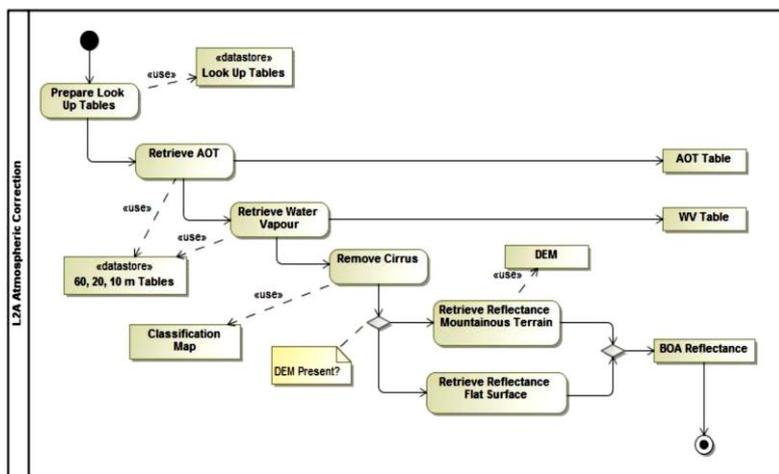


Р и с. 11. Блок-схема классификации сцен [3]

Детальное описание алгоритма классификации описано в [7] и [6].

Модуль атмосферной коррекции состоит из набора из четырех

различных подзадач (AOT, WV и terrain retrieval (опция с коррекцией terrain и cirrus, с тремя различными пользовательскими продуктами в качестве выходных данных: AOT и WV таблицы на уровне пикселей и BOA исправленные изображения для всех измеренных диапазонов. На рис. 11 ниже показан блок-схема для модуля атмосферной коррекции.



Р и с. 11. Блок-схема модуля атмосферной коррекции [3]

Атмосферическая модель Sen2Cor (L2A_AtmCorr) зависит от расчета радиационных передаточных функций для различных сенсорных и солнечных геометрий, высот земли и атмосферных параметров. В следующей таблице представлено 6-мерное пространство параметров и интервал сетки для каждого параметра. Процессор считывает параметр в виде таблиц подстановки (LUTs), относящихся к этому пространству параметров, и при необходимости интерполирует его. LUTs были созданы с помощью libRadtran, библиотеки для расчета солнечного и теплового излучения в атмосфере Земли. [1]

Т а б л и ц а 2

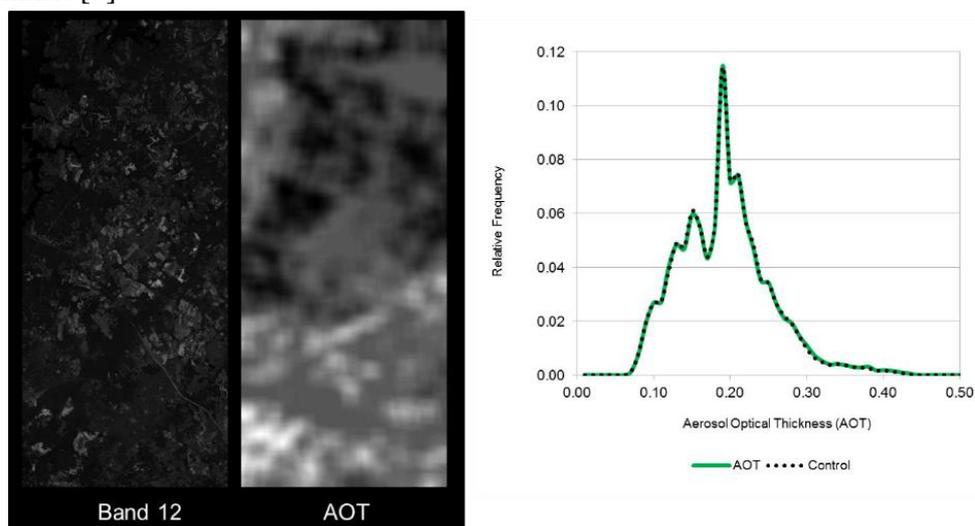
Пространство параметров для атмосферной коррекции

Parameter	Range	Increment / grid points
Solar zenith angle	0 -70°	10°
Sensor view angle	0 -10°	10°
Relative azimuth angle	0 -180°	30°(180°= backscatter)
Ground elevation	0 -2.5 km	0.5 km
Visibility	5 -120 km	5, 7, 10, 15, 23, 40, 80, 120 km
Water vapour, summer	0.4 -5.5 cm	0.4, 1.0, 2.0, 2.9, 4.0, 5.0 cm
Water vapour, winter	0.2 -1.5 cm	0.2, 0.4, 0.8, 1.1 cm

Извлечение атмосферной оптической толщины (AOT) обеспечивает измерение для визуально прозрачной атмосферы. Он

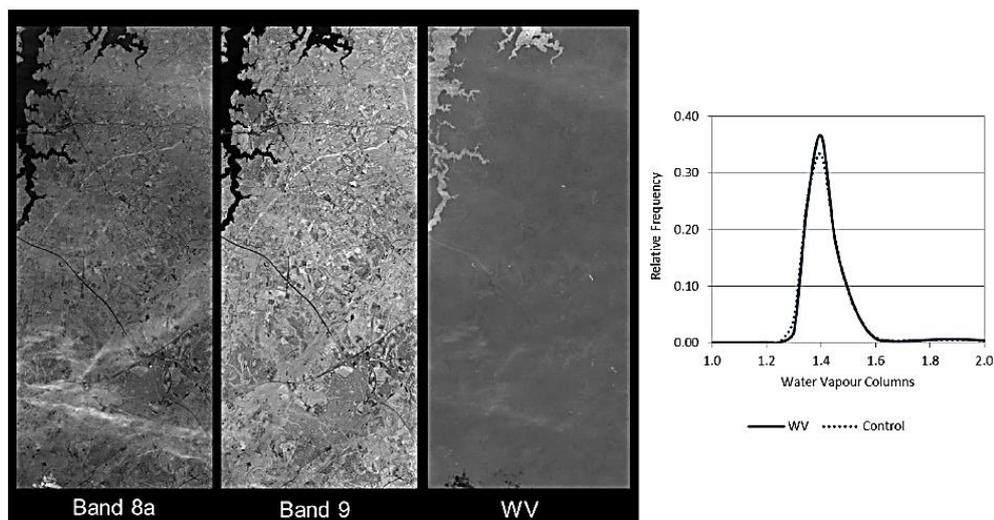
получен с использованием алгоритма DDV (густая Темная растительность) [5], используя коротковолновую инфракрасную (SWIR) полосу 12 и коррелирует ее отражение с полосами 4 (красный) и 2 (синий). Алгоритм требует, чтобы сцена содержала опорные области известного поведения отражения, предпочтительно темной густой растительности (DDV) и / или темной почвы и водных объектов.

Алгоритм предоставляет карту AOT, как показано на рисунке 12 ниже. [8]



Р и с. 12. Карта оптической толщины [3]

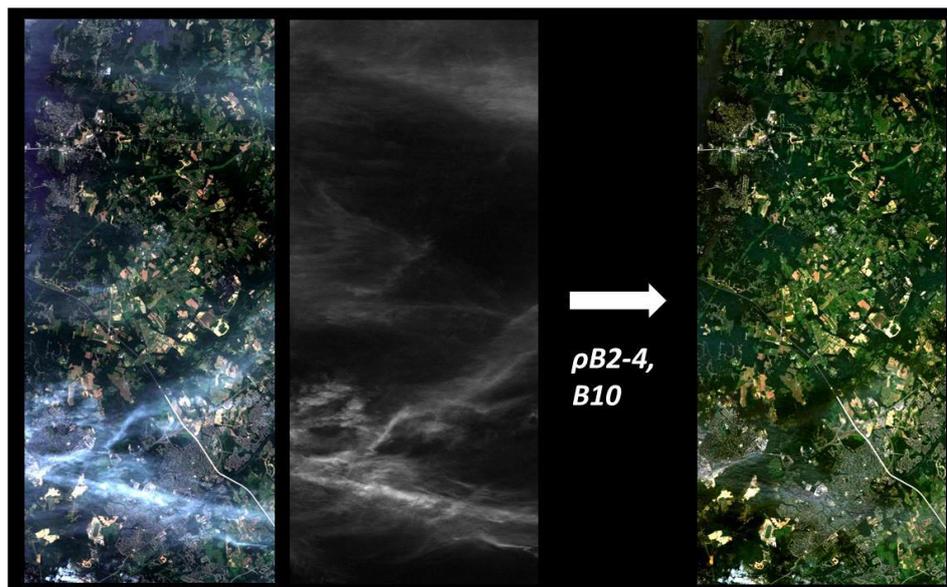
Извлечение Водяного Пара (WV) по суше осуществляется с помощью предварительно скорректированного алгоритма дифференциального поглощения атмосферы (APDA, [6]), который применяется к двум диапазонам Sentinel-2 B8A и B9 (Рис. 13). Диапазон 8A является опорным каналом в области атмосферного окна. Диапазон B9 канал измерения в зоне поглощения. Глубина поглощения оценивается путем вычисления яркости для атмосферы без WV, при условии, что коэффициент отражения поверхности для измерительного канала такой же, как и для опорного канала. Глубина поглощения затем является мерой содержания столбца WV.



Р и с. 13. Карта водяного пара [3]

Алгоритм коррекции перистых облаков использует канал В10. Тонкие перистые облака влияют на видимую, ближнюю и коротковолновую инфракрасную области спектра. Они частично прозрачны и поэтому их трудно обнаружить с помощью широкополосных мультиспектральных датчиков, особенно на пространственно-неоднородных участках суши. Водяной пар, напротив, доминирует в нижней тропосфере 0-5 км. Узкая спектральная полоса в спектральной области очень сильного поглощения WV (полоса 10), таким образом, будет поглощать наземный отраженный сигнал, но будет получать рассеянный сигнал перистых облаков.

Таким образом, отражательная способность перистых облаков диапазона 10 может быть коррелирована с другими полосами в области VNIR и SWIR, и таким образом вклад облаков может быть удален из сигнала сияния для получения сцены с исправлением. Это показано на рис. 14 ниже в качестве качественного результата.



Р и с. 14. Пример коррекции облаков [9]

Для каждого последовательного спектрального диапазона В1 – В12 выполняется поиск отражения поверхности на нижней границе атмосферы. На рисунке 15 ниже показаны входные данные уровня 1 С и соответствующие Выходные данные уровня 2А после атмосферной коррекции со сцены Ла-паса, полученные 28.03.2016.



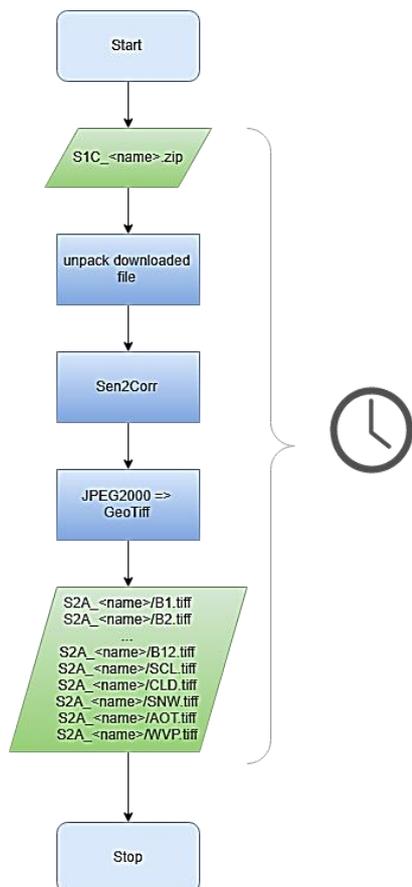
Р и с. 15. Слева – входной продукт L1C в видимом диапазоне; справа – выходной продукт – композиция 2-4 каналов, RGB.

Плагин Sen2Cor можно использовать двумя способами:

1. При помощи пакета Sentinel-2 Toolbox (SNAP)
2. Как утилита в командной строке

Возможность вызова плагина через командную строку – стала основой для разработки макета потоковой обработки данных Sentinel-2 из уровня 1С в 2А.

Плагин Sen2Corr был разработан на языке Python, по этой причине для разработки макета был выбран данный язык программирования. Концептуальная схема работы макета показана ниже.



Входными данными для разработанного программного продукта является набор заархивированных продуктов первого уровня, которые расположены в рабочей директории. После обнаружения архивов, программа распаковывает пакеты и после этого отправляет полученные пакеты на обработку Sen2Corr с теми аргументами, которые пользователю необходимы. После обработку плагина, из полученных данных второго уровня вытягиваются изображения в формате JPEG2000 и при помощи утилиты gdal_translate преобразуются к формату GeoTiff. В результате работы макета из первоначальных архивов данных с продуктами уровня 1С мы получаем наборы данных второго уровня в формате GeoTiff. Разработка макета велась в операционной среде Linux. Полученный макет был добавлен в файл cron, для систематической проверки каждый час на наличие архивов с данными для обработки.

Р и с. 16. Блок-схема разработанного макета.

Обработка атмосферной коррекции для разрешений 10 м использует огромное количество памяти из-за 10.000 x 10.000 пикселей для каждого изображения. Несколько изображений должны храниться через определенные промежутки времени полностью в памяти для выполнения корреляции. Таким образом, для каждой обработки с разрешением 10 м, как правило, должен быть доступен ресурс 8 ГБ памяти на процесс. Таким образом, если у вас нет машины с достаточным объемом памяти во время обработки возникнут ошибки переполнения памяти.

На тестируемой машине со следующими характеристиками: Linux, 16 ГБ оперативной памяти, 2.10 ГГц, Intel i7, обработка плитки для изображения с разными разрешениями, была успешно выполнена во время тестов.

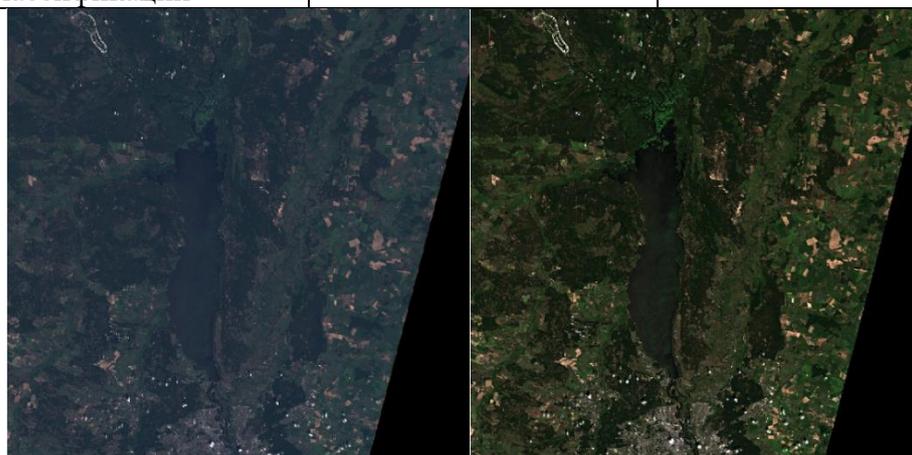
Согласно документации, оценка качества полученных продуктов в формате изображений можно отразить в табл.3.

Т а б л и ц а 3

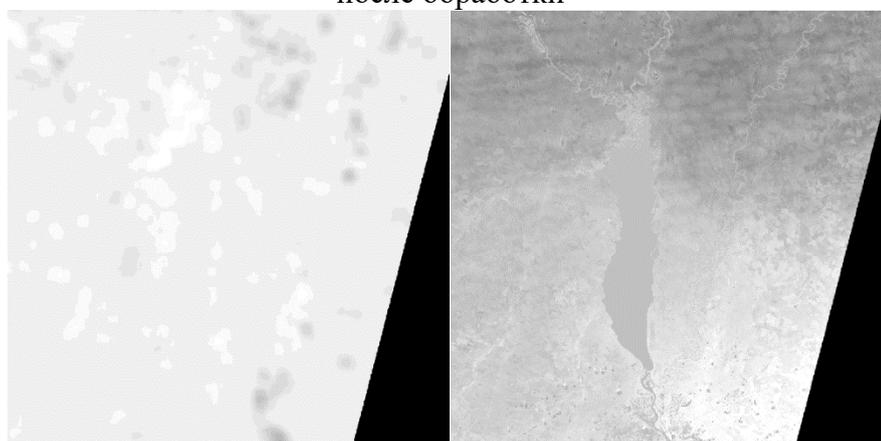
Таблица продуктов и оценке точности, заявленной ESA [9–10]

В ходе тестов были получены следующие продукты:

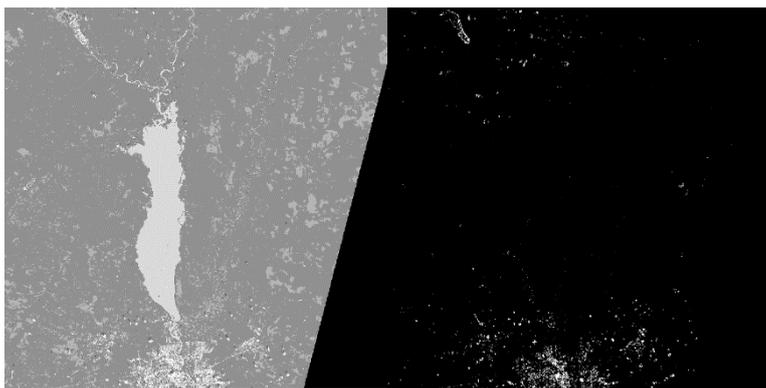
Наименование продукта	Алгоритм получения	Точность
Коэффициенты отражения на нижней границе атмосферы	L2A_AtmosCorr	5 %
Карта аэрозольной оптической толщины	L2A_AtmosCorr	5 – 10 %
Карта водяного пара	L2A_AtmosCorr	5 – 10 %
Результат классификации	L2A_SceneClass	(80 ± 5) %



Р и с. 17. Синтез в естественных цветах. Слева до обработки, справа после обработки



Р и с. 18. Слева карта аэрозольной оптической толщины, справа – карта водяного пара



Р и с. 19. Слева результат классификации, справа - карта вероятности облаков

В ходе работы также составлены таблицы с расширенной информацией по продуктам, получаемым после проведения атмосферной коррекции и классификации сцен.

В рамках данной работы были получены следующие результаты:

1. Проанализированы продукты и способы получения Sentinel-2 уровня обработки 1С
2. Проанализированы продукты Sentinel-2 уровня обработки 2В
3. Изучена архитектура плагина Sen2Cor для проведения атмосферной коррекции и получения продуктов из уровня 1С в уровень 2В.
4. Спроектирована архитектура макета для дальнейшей разработки
5. Реализован макет потоковой обработки данных
6. Протестирован макет на разных сценах

Реализованный макет потоковой обработки данных Sentinel-2 может быть внедрен в геоинформационную систему для получения продуктов из уровня обработки L1C.

Список литературы

1. Kaufman, Y., et al. (1997). The MODIS 2.1 μm channel – correlation with visible reflectance for use in remote sensing of aerosol. *IEEE TGRS*, **35**, 1286 –1298.
2. Kaufman, Y., Sendra, C. Algorithm for automatic atmospheric corrections to visible and near-IR satellite imagery, *International Journal of Remote Sensing*, Volume 9, Issue 8, 1357-1381 (1988).
3. Louis J., Debaecker V., Pflug B., Magdalena Main-Knorn, Bieniarz J., Uwe Mueller-Wilm, Cadau E., Ferran Gascon - SENTINEL-2 SEN2COR: L2A PROCESSOR FOR USERS.
4. Louis J., A. Charantonis & B. Berthelot, "Cloud Detection for Sentinel-2", *Proceedings of ESA Living Planet Symposium* (2010).

5. Richter R., Wang, X., Bachmann, M., and Schlaepfer, D., "Correction of cirrus effects in Sentinel-2 type of imagery", *Int. J. Remote Sensing*, Vol.32, 2931-2941 (2011).
6. Schläpfer, D. et al., "Atmospheric precorrected differential absorption technique to retrieve columnar water vapour", *Remote Sens. Environ.*, Vol. 65, 353366 (1998).
7. Salomonson V.V., I. Appel I. (2004). Estimating fractional snow cover from MODIS using the normalized difference snow index. *Remote Sensing of Environment* 89, 351–360.
8. Vane G., Green R., Chrien T., Enmark H., Hansen E., and Porter W. (1993). The airborne visible / infrared imaging spectrometer (AVIRIS). *Remote Sens. Environ.* 44, 127–143.
9. Sentinel Toolboxes. ESA - [Электронный ресурс]: URL: <https://sentinel.esa.int/web/sentinel/toolboxes> (дата обращения 20.05.2017)
10. Level-1C. ESA - [Электронный ресурс]: URL: <https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/product-types/level-1c> (дата обращения 20.05.2017)

THE LAYOUT OF THE STREAM DATA PROCESSING SENTINEL-2

L.A Bochka¹, A.V. Kashnitsky²

¹Lomonosov Moscow state University, Moscow

²Institute of space research RAS, Moscow

In this paper, the analysis of Sentinel-2 products of different levels. The European space Agency provides the software to create a distributed data products level products L1C L2A - Sen2Cor. The structure of the plug-in, which are the input data, which data is the output. The result of the work is a hardware-software system of streaming data processing from Sentinel-2 satellites, which allows to automatically create information products of L2A processing level and implement this product into the existing geographic information systems.

Keywords: *remote sensing of the Earth, Sentinel-2 data, satellite data processing, data processing technology, atmospheric correction.*

Об авторах:

БОЧКА Любовь Андреевна – студент 1-го курса магистратуры факультета космических исследований МГУ имени М.В.Ломоносова, e-mail: persifik@mail.ru

КАШНИЦКИЙ Александр Витальевич – кандидат технических наук, научных сотрудник отдела «Технологии спутникового мониторинга» Института космических исследований Российской Академии Наук, e-mail: kashnizky@gmail.com.