

## Создание инструментов для удаленной обработки спутниковых данных в современных информационных системах

А.В. Кашницкий, И.В. Балашов, Е.А. Лупян, В.А. Толпин, И.А. Уваров

*Институт космических исследований РАН  
Москва, 117997, Россия  
E-mail: smis@smis.iki.rssi.ru*

В последние годы все актуальней становится задача разработки новых технологий работы с данными дистанционных наблюдений, которые должны позволить организовывать эффективную работу со сверхбольшими, постоянно пополняющимися архивами. «Традиционные» подходы предполагают сначала выбор необходимых наборов данных из архивов, создание их локальных копий у пользователя (исследователя), а потом проведение обработки и анализа. Использование такого подхода ограничивается целым рядом факторов, наиболее существенными из которых являются взрывной рост данных, поступающих от систем спутниковых наблюдений Земли, необходимость выборки, передачи и хранения больших объемов информации и наличие существенных вычислительных ресурсов для анализа данных, которые не всегда имеются у конкретных исследователей. В настоящей работе представлена разработанная в Институте космических исследований Российской академии наук (ИКИ РАН) технология организации распределенной обработки спутниковых данных. Технология ориентирована на то, чтобы обеспечить возможности не только выбора и просмотра различной спутниковой информации, но и проведение ее обработки и анализа полученных результатов с использованием вычислительных средств, расположенных в центрах архивации и представления данных. Технология позволяет исследователям использовать распределенные информационные и вычислительные ресурсы с той же легкостью, что и локальные, обеспечивает возможность эффективной работы со сверхбольшими распределенными архивами спутниковых данных и результатами их обработки. В работе описывается предлагаемая архитектура построения систем распределенной обработки данных, ее основные возможности. Приведены примеры реализации процедур обработки и их использование в различных системах дистанционного мониторинга для организации обработки и анализа спутниковых данных, в том числе гиперспектральных.

**Ключевые слова:** дистанционное зондирование, информационные системы, распределенная обработка данных, обработка спутниковых данных, web-интерфейсы, сверхбольшие архивы данных, спутниковые данные.

### Введение

Данные спутниковых систем наблюдения Земли являются одним из основных, а иногда единственным, источником информации о различных процессах в атмосфере и на поверхности планеты. Это является причиной того, что информация дистанционного зондирования все активнее используется в самых разных областях, от чисто научных задач до коммерческих проектов. Постоянно возрастающая потребность в данных и развитие технологий в последние десятилетия привели к появлению большого количества высококачественных систем дистанционного зондирования Земли из космоса, обеспечивающих стабильное и оперативное получение информации по территории всей поверхности планеты (Лупян, Саворский и др., 2012). Это требует создания новых информационных технологий и подходов к организации работы с данными дистанционного зондирования.

Отметим, что «традиционные» подходы предполагают сначала выбор необходимых наборов данных из архивов, создание их локальных копий у конкретного пользователя (исследователя), а потом проведение их обработки и анализа. Как показано в работах (Лупян, Саворский, 2012; Лупян, Мазуров и др., 2011), такой путь в условиях быстро развивающихся спутниковых систем наблюдения становится все менее эффективным, а в перспективе и нереализуемым. Одна из основных причин этого заключается в значительном количественном и качественном увеличении архивов спутниковой информации и фактически их

взрывном росте. Даже при решении достаточно узких региональных задач может возникать потребность одновременной работы с десятками и сотнями терабайт информации. Другим немаловажным фактором является необходимость в больших вычислительных мощностях у пользователей и наличии у них достаточно сложного и ресурсоемкого программного обеспечения. Поэтому в последние годы стали развиваться более эффективные подходы к организации процессов работы с данными дистанционного мониторинга. В том числе стали создаваться системы, предоставляющие возможности доступа к архивам спутниковых данных и результатам их обработки одновременно со средствами, обеспечивающими проведение их обработки и анализа (Лупян, Саворский, 2012). Данные в таких системах могут находиться в территориально распределенных архивах, а их обработка и анализ, которыми управляет пользователь, обычно производится на вычислительных средствах, расположенных в тех же центрах, в которых физически осуществляется хранение данных. Это позволяет избежать передачи больших объемов информации и обеспечивает возможность эффективной работы со сверхбольшими объемами данных. При этом с развитием информационных технологий в последние годы стало возможным создание достаточно функциональных web-интерфейсов, обеспечивающих не только доступ к данным и их визуализацию, но и управления процессами их обработки. Во многих случаях такие системы по функциональности могут не уступать настольным приложениям, ориентированным на обработку спутниковой информации, обеспечивая при этом возможность непосредственной работы с огромными распределенными массивами информации с любого компьютера, подсоединенного к сети Интернет, и с использованием только web-браузера.

Создание таких систем позволяет принципиально поменять подходы к работе со спутниковой информацией, существенно упростить ее анализ и обработку, и, как следствие, расширить область применения данных дистанционного зондирования. В настоящей статье представлена разработанная в Институте космических исследований Российской академии наук (ИКИ РАН) технология, ориентированная на создание подобных систем и приведены примеры ее использования.

### **Текущее состояние направления**

Работы по созданию технологий и инструментов, позволяющих осуществлять обработку спутниковых данных на распределенных вычислительных ресурсах, ведутся достаточно давно, однако существенные успехи в данном направлении были достигнуты лишь в последние годы. Связано это со сложностью решаемой задачи, необходимостью использования достаточно мощных вычислительных ресурсов и развитием технологий построения web-интерфейсов. В нашей стране работы в этом направлении активно велись в ВЦ СО РАН (Шокин и др., 2012), ИАПУ ДВО РАН (Бабяк и др., 2014), ИКИ РАН (Балашов, Бурцев и др., 2010), подробное исследование таких систем приводится в (Куссуль, Шелестов, 2008). Однако указанные работы в основном были нацелены на предоставление

пользователям возможности работы с распределенными вычислительными ресурсами для организации off-line обработки спутниковых данных. Отдельно стоит отметить, что большинство разработанных подходов в этих работах были ориентированы на специалистов, имеющих достаточно большой опыт программирования и разрабатывающих собственное программное обеспечение для обработки данных дистанционных наблюдений.

В последние годы стали активно вестись разработки, направленные на максимальную интеграцию систем, обеспечивающих ведение сверхбольших распределенных архивов данных, и систем on-line анализа данных. Причем такие системы ориентированы на самый широкий круг специалистов, как использующих для работы с данными дистанционного зондирования различные специализированные программные комплексы и не имеющих глубоких навыков и знаний программирования, так и имеющих таковые для создания собственных процедур обработки. Наиболее яркой и быстроразвивающейся системой такого типа, на наш взгляд, является Google Earth Engine (Moore, Hansen, 2011; <https://earthengine.google.org>). Доверенные пользователи этой системы могут запускать любые обработки над всем архивом данных, доступных в системе, на кластере компании Google. В Google Earth Engine доступен архив спутниковых снимков за последние 40 лет, основу которого составляют данные спутников группировки Landsat и MODIS. Ее пользователям доступны как инструменты анализа данных (например, классификация), так и возможность запуска своего программного кода в вычислительной сети компании Google для обработки любых данных, имеющихся в архивах Google Earth Engine.

Другим примером системы, предоставляющей возможности доступа к архивам спутниковых данных и результатам их обработки одновременно со средствами, обеспечивающими проведение их обработки и анализа, может служить NASA GIOVANNI (Acker, Leptoukh, 2007). Кроме организации доступа к распределенным архивам данных, эта система предоставляет средства для дополнительной постобработки спутниковых данных и их анализа, при этом набор доступных преобразований сводится к заранее predetermined набору алгоритмов выборки и агрегации данных.

### **Постановка задачи и основа для ее решения**

В последние годы в ИКИ РАН проводится достаточно много работ по созданию, внедрению и поддержке различных информационных систем дистанционного мониторинга (Лупян, Мазуров и др., 2011). Для их реализации в институте созданы автоматизированные технологии, обеспечивающие возможность осуществления сбора спутниковой информации, ее обработки, ведения распределенных архивов спутниковых данных и распределенное представление спутниковой информации и результатов ее обработки. Следует отметить, что совместно с партнерами ИКИ РАН создает достаточно сложные информационные системы, ориентированные на решение как прикладных, так и научных задач (Лупян, Савин и др., 2011; Баргалева, Ершов и др., 2010; Бурцев и др., 2011; Лупян, Матвеев и др., 2012; Ефремов

и др., 2012). В таких системах, в первую очередь научных, постоянно возникает необходимость иметь возможность не только просмотра и визуального анализа информации, но также проведения удобной и эффективной обработки достаточно больших объемов информации. Поэтому обсуждаемая выше задача («разработка доступных и удобных в использовании инструментов анализа данных») стала особенно важной для развития возможностей создаваемых систем дистанционного мониторинга. При этом следует учитывать, что большинство пользователей таких систем, являясь специалистами в конкретных предметных областях, заинтересованы в получении и использовании инструментов анализа данных, не требующих дополнительного сложного программирования.

Таким образом, перед нами встала задача разработать технологию, которая позволяла бы достаточно легко добавлять в создаваемые информационные системы специализированные процедуры, управление которыми могли бы достаточно просто осуществлять пользователи этих систем. При этом управление обработкой должно осуществляться с использованием web-интерфейсов, а сама обработка данных проводится на вычислительных ресурсах, расположенных в центрах архивации. Схема реализации таких процедур должна быть вполне универсальной и легко интегрируемой в существующие информационные системы дистанционного мониторинга и центры архивации и предоставления данных. Она также должна обеспечивать возможности работы с долговременными архивами спутниковых данных, созданными и поддерживаемыми ИКИ РАН. В связи с этим созданная технология должна основываться на разработках института по созданию систем сбора, архивации, обработки и предоставления спутниковых данных. Принципы построения этих систем, архивов данных и доступа к ним в информационных системах, разработанных в ИКИ РАН, описаны в статьях (Балашов, Бурцев др., 2008; Бурцев и др., 2012; Балашов, Халикова и др., 2013; Балашов, Ефремов и др., 2009; Антонов и др., 2010), особенности аппаратной реализации систем обработки и хранения данных описаны в работе (Миклашевич и др., 2012).

Разработка проводилась с учетом того, что одним из основных средств работы со спутниковой информацией и результатами ее обработки в развиваемых системах дистанционного мониторинга являются специализированные картографические web-интерфейсы, построенные по универсальной технологии GEOSMIS (Толпин и др., 2011). Данная технология рассчитана на создание многофункциональных web-интерфейсов для работы с большими распределенными многомерными архивами спутниковых данных, в том числе и на создание инструментов анализа и управления данными. Поэтому она является хорошей основой для разработки интерфейсов управления различными процедурами распределенной обработки данных.

### **Описание разработанной технологии**

Создаваемая технология должна была обеспечить возможность разработки специализированных блоков информационных систем для реализации следующих основных этапов, характерных для обработки спутниковых данных:

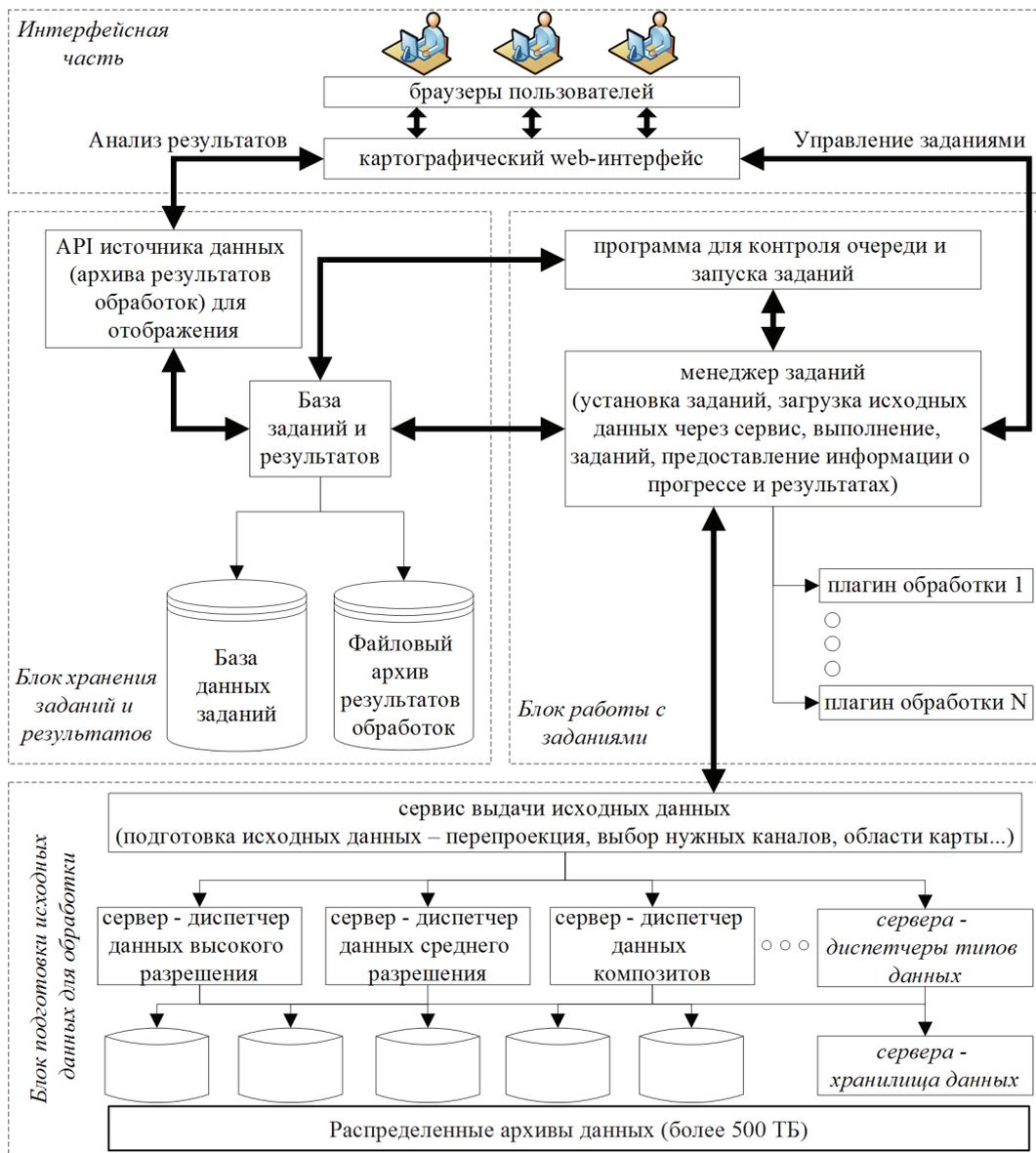


Рис. 1. Принципиальная схема построения технологии

- выбор данных для обработки;
- управление и настройку операций обработки;
- выполнение операций обработки (подготовка исходных данных, обработка, сохранение и представление результата);
- проведение анализа результатов.

Исходя из этого, была разработана структура создаваемой технологии, представленная на рис. 1. Технология обеспечивается реализацией следующих основных блоков:

- *Интерфейсная часть*. Обеспечивает возможность удаленного управления процедурами обработки (выбор данных, задание параметров, контроль исполнения и т.д.) и проведение анализа результатов. Реализуется в рамках картографических web-интерфейсов, созданных на основе технологии GEOSMIS. Для выбора наборов данных для обработки в интерфейсах используются стандартные средства поиска и выбора информации из архивов спутниковых данных (Толпин и др., 2011).

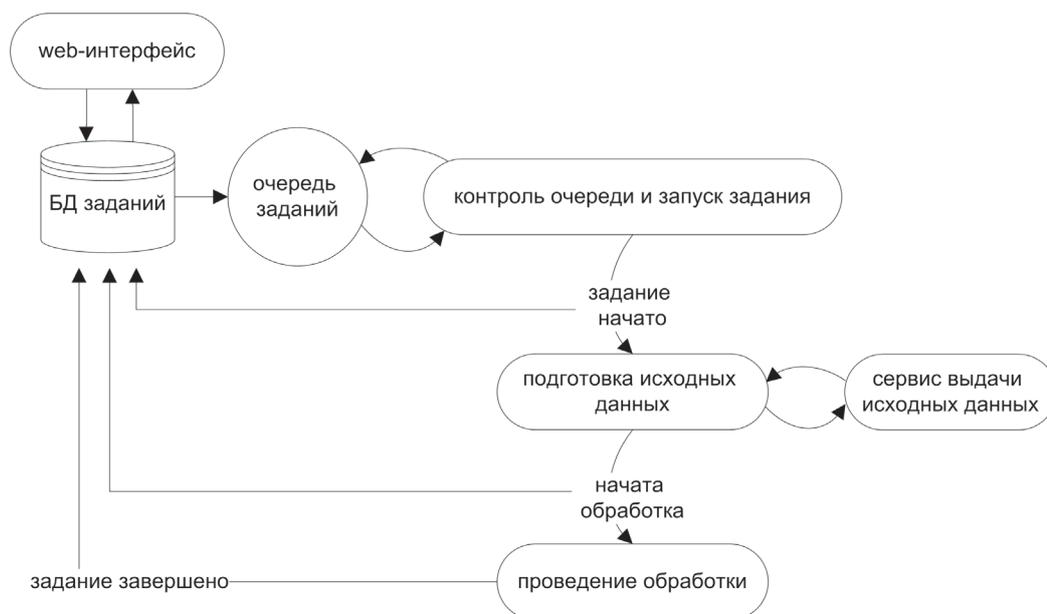


Рис. 2. Жизненный цикл запроса на обработку

При разработке технологии создан ряд универсальных блоков интерфейсов, которые могут использоваться для настройки параметров различных процедур обработки данных.

- *Блок хранения заданий и результатов.* Состоит из базы данных заданий (БДЗ), файлового архива результатов проведенных обработок, библиотеки доступа к ним и стандартизованных модулей интерфейса источника данных для сервиса отображения smiswms (Толпин и др., 2011). В БДЗ хранятся все задания на обработку с полной информацией о прогрессе и результатах выполнения.
- *Блок работы с заданиями.* Функционально состоит из менеджера заданий с отдельным плагином по каждому типу обработки и программы для контроля очереди и запуска заданий.
- *Блок подготовки данных для обработки.* Обеспечивает выбор наборов данных из архивов в соответствии с установками пользователей и передачу их в процедуры обработки.

Все блоки, созданные в рамках реализации технологии, являются универсальными и рассчитаны на использование в различных информационных системах и центрах данных дистанционного мониторинга. В соответствии с описанной базовой структурой однотипно для различных процедур обработки данных организуется цикл обработки, представленный на рис. 2. Через web-интерфейс пользователь в картографическом интерфейсе устанавливает интересующие его параметры заданий. Так как выполнение отдельных обработок может занимать значительное время, то интерфейс поддерживает возможность контроля прохождения обработки. После получения от пользователя задания на обработку происходит попадание задания в базу данных, а пользователю возвращается информация о том, что его задание поставлено в очередь. Каждый этап обработки со всеми параметрами фиксируется в базе данных заданий. Если пользователь ожидает завершения задания, то web-интерфейс

с некоторой периодичностью опрашивает базу данных через менеджер заданий и информирует пользователя о статусе его выполнения.

Работу с очередью осуществляет отдельная программа для контроля и запуска заданий. Очередь формируется по времени поступления, одновременно обрабатывается не более  $M$  заданий, где значение  $M$  подбирается в соответствии с техническими возможностями и загруженностью конкретного вычислительного ресурса, используемого в центре архивации данных для проведения обработки по заказам пользователей. Если в момент проверки очереди выполняется менее  $M$  обработок, программа для контроля и запуска через менеджер заданий инициализирует подготовку данных для следующей обработки.

Подготовка данных для одной обработки проводится параллельно для всех указанных пользователем источников данных конкретного задания. Со стороны менеджера заданий, работающего на сервере обработки, этот этап представляет собой обращение к серверу-диспетчеру исходных данных и загрузки через него данных. Процесс подготовки исходных данных связан с существующей архитектурой систем архивации и доступа к данным (Балашов, Бурцев и др., 2008; Бурцев и др., 2012). На данный момент можно выделить по виду их хранения три основных типа данных: сцен высокого и среднего разрешения, а также временных композитов. Для каждого типа данных выделен отдельный сервер-диспетчер, задача которого состоит в нахождении серверов, хранящих необходимые данные и перенаправление на него запроса. Далее непосредственно на сервере-хранилище над исходными данными проводятся все необходимые предварительные операции: перепроецирование (если необходимо), получение запрашиваемых для обработки каналов, вырезание данных только по требуемой географической области, объединение и формирование единого файла, если исходные данные на запрашиваемую область разбиты на несколько частей. Как результат, на выходе сервер подготовки выдает файл в формате GeoTIFF с исходными данными, который загружается на сервер обработки. Программный комплекс, выполняющий операции по созданию этого файла, получил название «сервис выдачи исходных данных». Важным моментом является тот факт, что, независимо от типа исходных данных, местоположения и принципа их хранения, выходной файл сервиса всегда стандартизован, что позволяет единообразно использовать исходные данные в разных обработках.

После получения исходных данных и занесения информации об этом в базу заданий на сервере инициализируется сама обработка. Все общие функции вынесены в менеджер заданий, а специфические для конкретной обработки – в соответствующий плагин. Такая схема позволяет легко расширять возможности системы новыми типами обработок, написав для каждого типа свой плагин.

После успешного завершения задания файлы результатов обработок могут передаваться на специально выделенные для этих целей сервера или храниться непосредственно на серверах, занимающихся обработкой. В любом случае информация о местонахождении результатов отображается в базе данных заданий и используется для доступа к результатам через `api` источника данных к сервису `smiswms` (Толпин и др., 2011). Отдельно стоит отметить, что по данной технологии можно проводить обработки с использованием ранее

полученных результатов. В таком случае вместо обращения к сервису исходных данных в качестве исходных могут браться файлы результатов ранее проведенных обработок.

Сервера, обеспечивающие работу различных блоков и функций созданной технологии (обработок, диспетчеризации, хранения и т.д.), могут быть территориально распределены. Взаимодействие между ними организовано через протокол http. Характеристики вычислительных средств, используемых для технической реализации технологии в центре хранения ИКИ РАН, описаны в (Миклашевич и др., 2012). Проведение обработки данных по разработанной технологии может масштабироваться на любое количество серверов. В таком случае контроль и распределение заданий по разным серверам обработки осуществляется программой для контроля очереди и запуска заданий через единую базу данных. Количество серверов, выполняющих обработки, подбирается в зависимости от количества запросов пользователей на выполнение заданий. Все сервера функционируют под управлением операционной системы FreeBSD, в качестве СУБД используется сервер баз данных MySQL. Программное обеспечение реализовано преимущественно на языке Perl с использованием библиотеки GDAL (Warmerdam, 2008) и модулей из открытой геоинформационной системы GRASS GIS (Neteler et al., 2012; <http://grass.osgeo.org/>).

Таким образом, в рамках технологии были разработаны все базовые элементы, которые позволяют создавать инструменты обработки данных в различных информационных сервисах и системах. Несколько примеров реализации таких возможностей приведены в следующем разделе.

### **Примеры реализации инструментов анализа данных, построенных на основе созданной технологии**

На основе разработанной технологии распределенной обработки данных были разработаны и внедрены в различные информационные системы инструменты, позволяющие непосредственно через картографические web-интерфейсы проводить классификации и различные операции (вычисления) с выбранными данными. Расчеты и классификации могут проводиться над всеми доступными в информационных системах архивами данными, в том числе гиперспектральными, причем не только над разными каналами одной сцены, но и над каналами разновременных сцен любых приборов любых спутников.

Инструмент для вычислений позволяет проводить арифметические, логические операции и различные математические преобразования над данными. Значения в отдельных каналах спутниковых снимков могут быть преобразованы по формуле, задаваемой пользователем непосредственно в интерфейсе. Результатом обработки являются новые растровые изображения, которые созданы из существующих растровых слоев с помощью арифметических и логических выражений с участием целых и дробных чисел, нормировок и математических функций. С помощью этого инструмента в картографических web-интерфейсах возможен расчет «на лету» различных индексов, проведение экспериментов по созданию

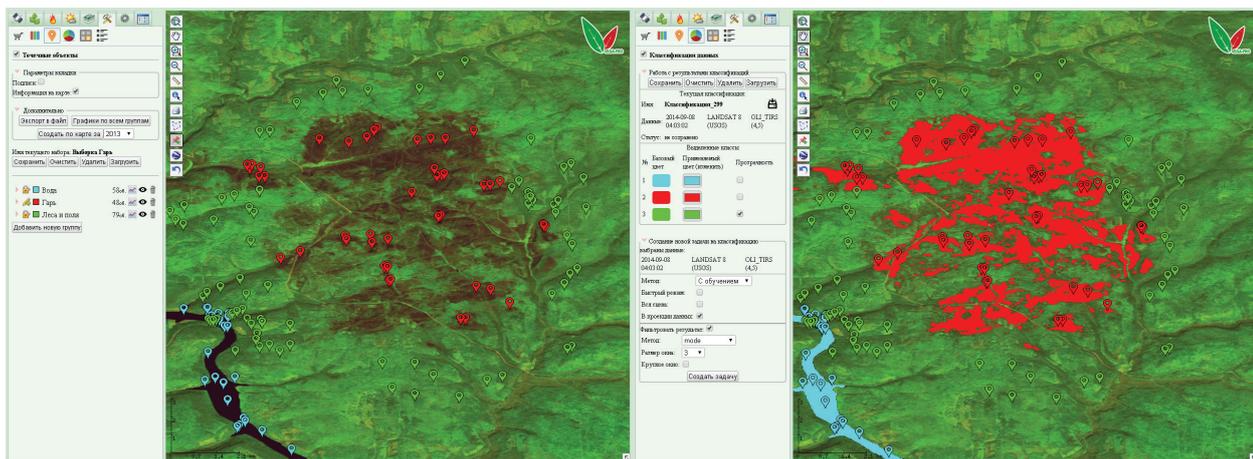


Рис. 3. Классификация лесной гари по каналам 4,5 сцены спутника Landsat 8 с обучением на вручную созданной выборке

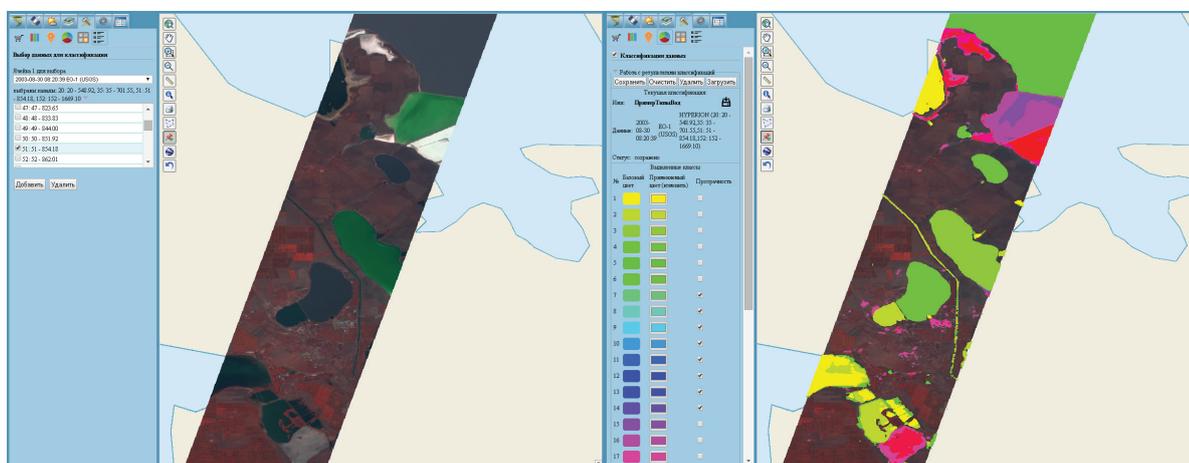


Рис. 4. Пример классификации без обучения по данным гиперспектрометра Hyperion спутника EO1

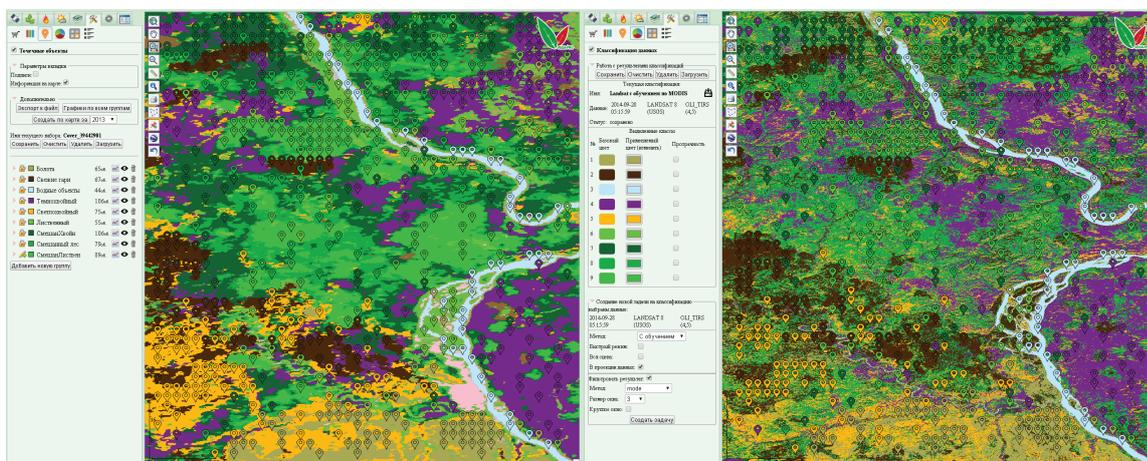


Рис. 5. Классификация данных Landsat с обучением на автоматически созданной выборке по карте растительности по данным MODIS

новых тематических продуктов. Например, инструмент позволяет исследователям строить на основе зон поглощения и излучения в спектре разностные индексы для изучения минералогического и химического состава подстилающей поверхности.

Инструмент проведения классификаций позволяет проводить достаточно сложную обработку данных для разделения спутниковых изображений на отдельные классы по определенным параметрам. Это может быть использовано для выделения различных особенностей изображения и построения тематических карт. Доступны методы с обучением и без обучения. Реализованные алгоритмы классификации взяты из модулей GRASS GIS (Neteler et al., 2012; <http://grass.osgeo.org/>). Метод без обучения разбит на два этапа: расчет сигнатур и сама классификация методом максимального правдоподобия. Базируется на автоматической кластеризации объектов по признаку схожести спектральных характеристик. Все пиксели, имеющие сходные спектральные характеристики, объединяются в один класс. Число выделяемых классов задается пользователем.

В случае метода с обучением для задания обучающих выборок используется отдельная система управления векторными объектами. Эта система позволяет задавать выборку вручную или создавать автоматически. Например, для автоматического создания обучающей выборки для проведения локальной классификации растительного покрова по данным Landsat за основу могут быть взяты карты растительного покрова, полученные по данным MODIS (Барталев и др., 2011), а сама обучающая выборка (информация о значениях яркостей в точках, входящих в обучающую выборку) будет уже сформирована из данных, которые в дальнейшем подлежат классификации. При этом каждый пиксель на карте растительности, окруженный пикселями того же класса, считается доверенным и может быть выбран для обучения. Такой подход позволяет обучаться на ранее созданных и проверенных картах растительности по данным MODIS для проведения классификации данных более высокого пространственного разрешения. Независимо от способа получения обучающей выборки, с использованием рассчитанных по ней сигнатур проводится сегментация и контекстная классификация изображения методом оценки с помощью апостериорного максимума (SMAP) (Bouman, Shapiro, 1994) или методом максимального правдоподобия.

На рисунках приведены примеры проведения различных вариантов классификации в спутниковых сервисах Vega (Лупян, Савин и др., 2011) и See the Sea (Лупян, Матвеев и др., 2012), выполненных инструментами, реализованными на основе созданной технологии. На *рис. 3* слева показано создание обучающей выборки для выделения лесной гари по снимку Landsat 8 и справа – результат проведенной по созданной выборке классификации каналов 4 и 5, причем выделенный класс «леса и поля» скрыт при отображении для более наглядной оценки результата. На *рис. 4* приведены примеры выделения типов загрязнений вод на Перекопском перешейке с помощью классификации без обучения по данным гиперспектрометра Nuregion спутника EO1: слева – вид интерфейса для выбора каналов для проведения классификации, справа – результат классификации, на котором некоторые классы скрыты для наглядности. На *рис. 5* слева показано автоматическое создание обучающей выборки по карте растительного покрова, полученной по данным MODIS, а справа – результат проведения классификации данных более высокого разрешения (Landsat 8) по этой выборке.

На данный момент инструменты обработки и анализа данных, построенные по описываемой в статье технологии распределенной обработки, интегрированы в следующие информационные системы:

- спутниковый сервис ВЕГА, предназначенный для решения задач дистанционного мониторинга растительного покрова (Лупян, Савин и др., 2011);
- информационная система дистанционного мониторинга Федерального агентства лесного хозяйства (ИСДМ Рослесхоз) (Ефремов и др., 2011; Барталев и др., 2010);
- объединенная система работы со спутниковыми данными центров приема и обработки спутниковых данных НИЦ «Планета» Росгидромета (Бурцев и др., 2011; Бурцев и др., 2012);
- спутниковый сервис See the Sea, ориентированный на решение задач, связанных с исследованием различных процессов на поверхности океана (Лупян, Матвеев и др., 2012);
- информационный сервис VolSatView «Дистанционный мониторинг активности вулканов Камчатки и Курил» (Ефремов и др., 2012).

Работы по интеграции инструментов обработки данных в различные научные информационные системы проводились при поддержке РФФИ 13-07-12017 офи\_м, 13-07-12180 офи\_м, 15-07-05564 А.

Реализованные инструменты позволяют в настоящее время работать с информацией, имеющейся в архивах этих спутниковых систем. Реализация и использование созданных инструментов в этих системах для решения различных задач показали работоспособность созданной технологии и удобство ее применения при разработке различного функционала распределенной обработки и анализа данных в системах дистанционного мониторинга.

### **Заключение**

Таким образом, нам удалось создать работоспособную технологию, позволяющую создавать достаточно сложные процедуры распределенной обработки и анализа данных дистанционного зондирования и интегрировать ее в интерфейсы, обеспечивающие удаленную работу с данными в информационных системах дистанционного мониторинга. Данная технология позволяет реализовывать принципиально новые подходы к работе со сверхбольшими распределенными архивами данных спутниковых наблюдений Земли. В частности она:

- позволяет избежать необходимости формирования у пользователей крупных архивов спутниковых данных, при этом расширяя возможности по использованию информации, находящейся в крупных центрах архивации и представления данных;
- позволяет пользователям осуществлять работу со спутниковыми данными из любой удобной для них точки и в любое удобное время;
- позволяет в ряде случаев избежать для проведения анализа спутниковых данных необходимости использования дорогостоящих специализированных настольных приложений;

- позволяет достаточно простым и прозрачным образом организовать использование высокопроизводительных вычислительных ресурсов.

В тоже время следует отметить, что современные web-технологии в настоящий момент обладают некоторыми ограничениями, которые пока затрудняют реализацию в них сложного функционала, который может присутствовать в современных специализированных настольных приложениях (в основном связанного с разными интерфейсными функциями). Однако, судя по развитию этих технологий в последние годы, в ближайшей перспективе эти ограничения во многом могут быть сняты.

Работа выполнена также при поддержке гранта РФФИ 13-07-12116 офи\_м.

## Литература

1. Антонов А.В., Бурцев М.А., Ефремов В.Ю., Калашников А.В., Крамарева Л.С., Крашенинникова Ю.С., Лупян Е.А., Матвеев А.М., Прошин А.А., Флитман Е.В. Построение объединенного каталога распределенных архивов спутниковых данных различных центров // Современные проблемы дистанционного зондирования Земли из космоса, 2010. Т. 7. № 2. С. 84-89.
2. Бабяк П. В., Недолужко И. В., Фомин Е. В. Подход к предоставлению услуг по обработке спутниковых данных в Центре коллективного пользования регионального спутникового мониторинга окружающей среды ДВО РАН // Материалы конф. «Интернет и современное общество», 2014. С. 27-32.
3. Балашов И.В., Бурцев М.А., Галеев А.А., Ефремов В.Ю., Крашенинникова Ю.С., Лупян Е.А., Мазуров А.А., Московский А.А., Прошин А.А., Сергеева К.О., Флитман Е.В. Разработка эффективных высокопроизводительных решений для создания систем динамической обработки спутниковых данных и результатов их анализа // Третья международная научная конференция «Суперкомпьютерные системы и их применения» (SSA' 2010). 25-27 мая 2010. Минск. Сборник докладов, ОИПИ НАН Беларуси, 2010. Т. 1. С. 30-34.
4. Балашов И.В., Бурцев М.А., Ефремов В.Ю., Лупян Е.А., Прошин А.А., Толпин В.А. Построение архивов результатов обработки спутниковых данных для систем динамического формирования производных информационных продуктов // Современные проблемы дистанционного зондирования Земли из космоса, 2008. Выпуск 5. Т. 1. С. 26-32.
5. Балашов И. В., Ефремов В.Ю., Лупян Е.А., Прошин А.А., Толпин В.А. Построение систем, обеспечивающих динамическое формирование комплексных информационных продуктов на основе данных дистанционного зондирования // Современные проблемы дистанционного зондирования Земли из космоса, 2009. Выпуск 6. Т. 2. С. 513-520.
6. Балашов И.В., Халикова О.А., Бурцев М.А., Лупян Е.А., Матвеев А.М. Организация автоматического получения наборов информационных продуктов из центров архивации и распространения спутниковых и метеоданных // Современные проблемы дистанционного зондирования Земли из космоса, 2013. Т. 10. № 3. С. 9-20.
7. Барталев С.А., Егоров В.А., Еришов Д.В., Исаев А.С., Лупян Е.А., Плотников Д.Е., Уваров И.А. Спутниковое картографирование растительного покрова России по данным спектрорадиометра MODIS // Современные проблемы дистанционного зондирования Земли из космоса, 2011. Т. 8. № 4. С. 285-302.
8. Барталев С.А., Еришов Д.В., Коровин Г.Н., Котельников Р.В., Лупян Е.А., Щетинский В.Е. Основные возможности и структура информационной системы дистанционного мониторинга лесных пожаров Федерального агентства лесного хозяйства (ИСДМ Рослесхоз) // Современные проблемы дистанционного зондирования Земли из космоса, 2010. Т. 7. № 2. С. 97-105.
9. Бурцев М.А., Антонов В.Н., Ефремов В.Ю., Кашицкий А.В., Крамарева Л.С., Лупян Е.А., Мазуров А.А., Матвеев А.М., Милехин О.Е., Прошин А.А., Соловьев В.И. Система работы с распределенными архивами результатов обработки спутниковых данных центров приема НИЦ «Планета» // Современные проблемы дистанционного зондирования Земли из космоса, 2012. Т. 9. № 5. С. 55-76.
10. Бурцев М.А., Ефремов В.Ю., Балашов И.В., Мазуров А.А., Прошин А.А., Лупян Е.А., Милехин О.Е. Система доступа к данным Европейского, Сибирского и Дальневосточного центров приема Росгидромета. // Современные проблемы дистанционного зондирования Земли из космоса, 2011. Т. 8. № 3. С. 113-119.
11. Ефремов В.Ю., Балашов И.В., Котельников Р.В., Лупян Е.А., Мазуров А.А., Прошин А.А., Толпин В.А., Уваров И.А., Флитман Е.В. Объединенный картографический интерфейс для работы с данными ИСДМ Рослесхоз // Современные проблемы дистанционного зондирования Земли из космоса, 2011. Т. 8. № 3. С. 129-139.
12. Ефремов В.Ю., Гирина О.А., Крамарева Л.С., Лупян Е.А., Маневич А.Г., Мельников Д.В., Матвеев А.М., Прошин А.А., Сорокин А.А., Флитман Е.В. Создание информационного сервиса «Дистанционный мониторинг активности вулканов Камчатки и Курил» // Современные проблемы дистанционного зондирования Земли из космоса, 2012. Т. 9. № 5. С. 155-170.
13. Куссуль Н. Н., Шелестов А. Ю. Grid-системы для задач исследования Земли // К.: Наукова думка, 2008. 452 с.
14. Лупян Е.А., Мазуров А.А., Назиров Р.Р., Прошин А.А., Флитман Е.В., Крашенинникова Ю.С. Технологии построения информационных систем дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса, 2011. Т. 8. № 1. С. 26-43.

15. Лупян Е.А., Матвеев А.А., Уваров И.А., Бочарова Т.Ю., Лаврова О.Ю., Митягина М.И. Спутниковый сервис See the Sea – инструмент для изучения различных явлений на поверхности океана // Современные проблемы дистанционного зондирования Земли из космоса, 2012. Т. 9. № 2. С. 251-262.
16. Лупян Е.А., Савин И.Ю., Барталев С.А., Толпин В.А., Балашов И.В., Плотников Д.Е. Спутниковый сервис мониторинга состояния растительности («ВЕГА») // Современные проблемы дистанционного зондирования Земли из космоса, 2011. Т. 8. № 1. С. 190-198.
17. Лупян Е.А., Саворский В.П., Шокин Ю.И., Алексанин А.И., Назиров Р.Р., Недолужко И.В., Панова О.Ю. Современные подходы и технологии организации работы с данными дистанционного зондирования Земли для решения научных задач // Современные проблемы дистанционного зондирования Земли из космоса, 2012. Т. 9. № 5. С. 21-44.
18. Миклашевич С.Э., Балашов И.В., Бурцев М.А., Ефремов В.Ю., Мазуров А.А., Матвеев А.М., Прошин А.А., Радченко М.В., Флитман Е.В. Программно-аппаратный комплекс для сбора, обработки, архивации и распространения спутниковых данных и продуктов их тематической обработки // Современные проблемы дистанционного зондирования Земли из космоса, 2012. Т. 9. № 4. С. 47-56.
19. Толпин В.А., Балашов И.В., Ефремов В.Ю., Лупян Е.А., Прошин А.А., Уваров И.А., Флитман Е.В. Создание интерфейсов для работы с данными современных систем дистанционного мониторинга (система GEOSMIS) // Современные проблемы дистанционного зондирования Земли из космоса, 2011. Т. 8. № 3. С. 93-108.
20. Шокин Ю.И., Антонов В.Н., Добрецов Н.Н., Кихтенко В.А., Лагутин А.А., Смирнов В.В., Чубаров Д.Л., Чубаров Л.Б. Распределенная система приема и обработки спутниковых данных Сибири и Дальнего Востока. Текущее состояние и перспективы развития // Современные проблемы дистанционного зондирования Земли из космоса, 2012. Т. 9. № 5. С. 45-54.
21. Acker J. G., Lepoutouk G. Online analysis enhances use of NASA earth science data // Eos, Transactions American Geophysical Union, 2007. Vol. 88. No. 2. P. 14-17.
22. Bouman C. A., Shapiro M. A multiscale random field model for Bayesian image segmentation // IEEE Trans. Image Process. 1994. Vol. 3. No. 2. P. 162-177.
23. Moore R. T., Hansen M. C. Google Earth Engine: a new cloud-computing platform for global-scale earth observation data and analysis // AGU Fall Meeting Abstracts, 2011. Vol. 1. P. 2.
24. Neteler M., Bowman M.H., Landa M., Metz M. GRASS GIS: A multi-purpose open source GIS // Environmental Modelling & Software, 2012. Vol. 31. P. 124-130.
25. Warmerdam F. The geospatial data abstraction library // Open Source Approaches in Spatial Data Handling. Springer Berlin Heidelberg, 2008. P. 87-104.

## Development of software tools for satellite data remote processing in contemporary information systems

A.V. Kashnitskiy, I.V. Balashov, E.A. Loupian, V.A. Tolpin, I.A. Uvarov

*Space Research Institute RAS  
Moscow 117997, Russia  
E-mail: smis@smis.iki.rssi.ru*

The problem of development of new technologies for satellite observation data processing capable for the efficient extra large continuously updated archive management has become more and more urgent in recent years. “Traditional” approaches propose selecting matching input datasets from the archives, making their local copy for the user or researcher at the first step, and performing data processing and analysis after that. Using of this kind of approach is limited due to a number of factors. The most significant ones are the explosive growth of the data volume acquired from Earth observation satellite systems, the need for selection, transmission and storage of huge data volumes, the existence of powerful datacenter-based computational resources, which typically do not belong to end users. This paper presents the technology of distributed satellite data processing management development in Space Research Institute of the Russian Academy of Sciences (IKI RAS). The technology is aimed at extension of functionality from simple selection and displaying of satellite based information to providing the means of data processing and analysis of data products with use of computational resources located in data archiving and representation centers. The technology enables the researchers to use the distributed information and computational resources as easily as if they were local. It provides capabilities of efficient access to extra large distributed satellite imagery and data products archives. The paper describes the proposed distributed data processing systems architecture and its principal functional capabilities. One can find examples of data processing procedures implementation and use in a number of remote monitoring systems for arrangement of satellite data processing and analysis, including hyperspectral data.

**Keywords:** remote sensing, information systems, distributed data processing, satellite data processing, web interfaces, extra large data archives, satellite data.

## References

1. Antonov A.V., Bourtsev M.A., Efremov V.Yu., Kalashnikov A.V., Kramareva L.S., Krasheninnikova Yu.S., Loupian E.A., Matveev A.M., Proshin A.A., Flitman E.V. Postroenie ob"edinennogo kataloga raspredelennykh arkhivov sputnikovyykh dannykh razlichnykh tsentrov (Design of the union catalogue of distributed satellite data archives of various centers), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2010, Vol. 7, No. 2, pp. 84-89.
2. Babyak P.V., Nedoluzhko I.V., Fomin E.V. Podkhod k predostavleniyu uslug po obrabotke sputnikovyykh dannykh v Tsentre kollektivnogo pol'zovaniya regional'nogo sputnikovogo monitoringa okruzhayushchei sredy DVO RAN (Approach to providing the satellite data processing service in framework of regional environmental monitoring Collective Users' Center of the RAS Far East branch), *Internet i sovremennoe obshchestvo* (Internet and the modern society), Proc. Conf., 2014, pp. 27-32.
3. Balashov I.V., Bourtsev M.A., Galeev A.A., Efremov V.Yu., Krasheninnikova Yu.S., Loupian E.A., Mazurov A.A., Moskovskii A.A., Proshin A.A., Sergeeva K.O., Flitman E.V. Razrabotka effektivnykh vysokoproizvoditel'nykh reshenii dlya sozdaniya sistem dinamicheskoi obrabotki sputnikovyykh dannykh i rezul'tatov ikh analiza (Development of effective high performance solutions for making of satellite imagery and data products dynamic processing systems), *3rd Int. Sci. Conf. Supercomputer systems and applications (SSA' 2010)*, May 25-27, 2010, Minsk, Proc. Conf., 2010, Vol. 1, pp. 30-34.
4. Balashov I.V., Bourtsev M.A., Efremov V.Yu., Loupian E.A., Proshin A.A., Tolpin V.A. Postroenie arkhivov rezul'tatov obrabotki sputnikovyykh dannykh dlya sistem dinamicheskogo formirovaniya proizvodnykh informatsionnykh produktov (Development of satellite data processing products archives for dynamic derived information products production systems), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2008, Issue 5, No. 1, pp. 26-32.
5. Balashov I.V., Efremov V.Yu., Loupian E.A., Proshin A.A., Tolpin V.A. Postroenie sistem, obespechivayushchikh dinamicheskoe formirovanie kompleksnykh informatsionnykh produktov na osnove dannykh distantsionnogo zondirovaniya (Development of Remote Sensing Complex Information Products Dynamic Creation Systems), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2009, Issue 6, No. 2, pp. 513-520.
6. Balashov I.V., Khalikova O.A., Bourtsev M.A., Loupian E.A., Matveev A.M. Organizatsiya avtomaticheskogo polucheniya naborov informatsionnykh produktov iz tsentrov arkhivatsii i rasprostraneniya sputnikovyykh i meteodannykh (Organization of automatic data acquisition from satellite and meteorological data and distribution centers), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2013, Vol. 10, No. 3, pp. 9-20.
7. Bartalev S.A., Egorov V.A., Ershov D.V., Isaev A.S., Loupian E.A., Plotnikov D.E., Uvarov I.A. Sputnikovoe kartografirovaniye rastitel'nogo pokrova Rossii po dannym spektrometra MODIS (Mapping of Russia's vegetation cover using MODIS satellite spectroradiometer data), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 4, pp. 285-302.
8. Bartalev S.A., Ershov D.V., Korovin G.N., Kotelnikov R.V., Loupian E.A., Shchetinskii V.E. Osnovnye vozmozhnosti i struktura informatsionnoi sistemy distantsionnogo monitoringa lesnykh pozharov Federal'nogo agentstva lesnogo khozyaistva (ISDM Rosleskhoz) (The Main Functionalities and Structure of the Forest Fire Satellite Monitoring Information System of Russian Federal Forestry Agency (ISDM-Rosleskhoz)), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2010, Vol. 7, No. 2, pp. 97-105.
9. Bourtsev M.A., Antonov V.N., Efremov V.Yu., Kashnitskiy A.V., Kramareva L.S., Loupian E.A., Mazurov A.A., Matveev A.M., Milekhin O.E., Proshin A.A., Solov'ev V.I. Sistema raboty s raspredelennymi arkhivami rezul'tatov obrabotki sputnikovyykh dannykh tsentrov priema NITs "Planeta" (Distributed satellite data processing products archives operation system in the SRC Planeta centres), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2012, Vol. 9, No. 5, pp. 55-76.
10. Bourtsev M.A., Efremov V.Yu., Balashov I.V., Mazurov A.A., Proshin A.A., Loupian E.A., Milekhin O.E. Sistema dostupa k dannym Evropeiskogo, Sibirskogo i Dal'nevostochnogo tsentrov priema Rosgidrometa (Data handling system for the European, Siberian and Far Eastern Centres of Roshydromet), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 3, pp. 113-119.
11. Efremov V.Yu., Balashov I.V., Kotelnikov R.V., Loupian E.A., Mazurov A.A., Proshin A.A., Tolpin V.A., Uvarov I.A., Flitman E.V. Ob"edinennyi kartograficheskii interfeis dlya raboty s dannymi ISDM Rosleskhoz (Integrated mapping interface for operations with ISDM-Rosleskhoz data), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 3, pp. 129-139.
12. Efremov V.Yu., Girina O.A., Kramareva L.S., Loupian E.A., Manevich A.G., Mel'nikov D.V., Matveev A.M., Proshin A.A., Sorokin A.A., Flitman E.V. Sozdanie informatsionnogo servisa "Distantsionnyi monitoring aktivnosti vulkanov Kamchatki i Kuril" (Creating the information service "Remote Monitoring of Active Volcanoes of Kamchatka and the Kuril Islands"), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2012, Vol. 9, No. 5, pp. 155-170.
13. Kussul' N.N., Shelestov A.Yu. *Grid-sistemy dlya zadach issledovaniya Zemli* (Grid systems in the Earth observation area), Kyiv, Naukova dumka, 2008, 452 p.
14. Loupian E.A., Mazurov A.A., Nazirov R.R., Proshin A.A., Flitman E.V., Krasheninnikova Yu.S. Tekhnologii postroeniya informatsionnykh sistem distantsionnogo monitoringa (Technologies for Building Remote Monitoring Information Systems), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 1, pp. 26-43.
15. Loupian E.A., Matveev A.A., Uvarov I.A., Bocharova T.Yu., Lavrova O.Yu., Mityagina M.I. Sputnikovyi servis See the Sea – instrument dlya izucheniya razlichnykh yavlenii na poverkhnosti okeana (The Satellite Service See the Sea - a tool for the study of oceanic phenomena and processes), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2012, Vol. 9, No. 2, pp. 251-262.
16. Loupian E.A., Savin I.Yu., Bartalev S.A., Tolpin V.A., Balashov I.V., Plotnikov D.E. Sputnikovyi servis monitoringa sostoyaniya rastitel'nosti VEGA (Satellite service for vegetation monitoring VEGA), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 1, pp. 190-198.
17. Loupian E.A., Savorskiy V.P., Shokin Yu.I., Aleksanin A.I., Nazirov R.R., Nedoluzhko I.V., Panova O.Yu. *Sovremennye podkhody i tekhnologii organizatsii raboty s dannymi distantsionnogo zondirovaniya Zemli dlya*

- resheniya nauchnykh zadach (Up-to-date approaches and technology arrangement of Earth observation data applications aimed to solve scientific tasks), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2012, Vol. 9, No. 5, pp. 21-44.
18. Miklashevich S.E., Balashov I.V., Bourtsev M.A., Efremov V.Yu., Mazurov A.A., Matveev A.M., Proshin A.A., Radchenko M.V., Flitman E.V. Programmno-apparatnyi kompleks dlya sbora, obrabotki, arkhivatsii i rasprostraneniya sputnikovykh dannykh i produktov ikh tematicheskoi obrabotki (Complex system for the receiving, processing, archiving and distribution of satellite data and product of thematic processing), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2012, Vol. 9, No. 4, pp. 47-56.
  19. Tolpin V.A., Balashov I.V., Efremov V.Yu., Loupian E.A., Proshin A.A., Uvarov I.A., Flitman E.V. Sozdanie interfeisov dlya raboty s dannymi sovremennykh sistem distantsionnogo monitoringa (sistema GEOSMIS) (GEOSMIS: developing interfaces to operate data in modern remote monitoring systems), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 3, pp. 93-108.
  20. Shokin Yu.I., Antonov V.N., Dobretsov N.N., Kikhtenko V.A., Lagutin A.A., Smirnov V.V., Chubarov D.L., Chubarov L.B. Raspredeleonnaya sistema priema i obrabotki sputnikovykh dannykh Sibiri i Dal'nego Vostoka. Tekushchee sostoyanie i perspektivy razvitiya (Development and status of the distributed system for satellite data acquisition and processing for Siberia and the Far East), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2012, Vol. 9, No. 5, pp. 45-54.
  21. Acker J. G., Leptoukh G. Online analysis enhances use of NASA earth science data, *Eos, Transactions American Geophysical Union*, 2007, Vol. 88, No. 2, pp. 14-17.
  22. Bouman C. A., Shapiro M. A multiscale random field model for Bayesian image segmentation, *IEEE Trans. Image Process.*, 1994, Vol.3, No. 2, pp. 162-177.
  23. Moore R. T., Hansen M. C. Google Earth Engine: a new cloud-computing platform for global-scale Earth observation data and analysis, *AGU Fall Meeting, Abstracts*, 2011, Vol.1, pp. 2.
  24. Neteler, M.; Bowman, M.H.; Landa, M.; Metz, M. GRASS GIS: A multi-purpose open source GIS, *Environmental Modelling & Software*, 2012, Vol. 31, pp. 124-130.
  25. Warmerdam F. The geospatial data abstraction library, In: *Open Source Approaches in Spatial Data Handling*, Springer Berlin Heidelberg, 2008, pp. 87-104.