

Построение архива спутниковых данных для анализа динамики растительности

М.А. Бурцев, А.А. Мазуров, И.А. Нейштадт., А.А. Прошин

*Институт космических исследований РАН
117997 Москва, ул. Профсоюзная, 84/32
E-mail: info@smis.iki.rssi.ru*

В настоящей статье приведено описание созданного в ИКИ РАН архива данных Terra-MODIS, используемого для решения задач анализа динамики растительного покрова. Представлено описание архитектуры архива, его программных компонент, а также программных и аппаратных особенностей реализации. Приведены перспективы развития архива.

В условиях глобальных изменений одной из наиболее актуальных является задача исследования динамики растительного покрова (см. например [1-7]). Незаменимую информацию для ее решения, безусловно, дают системы дистанционного зондирования Земли, позволяющие получать объективную устойчивую информацию наблюдения всех регионов Земли. К спутниковым данным и результатам из обработки данным, которые используются для изучения долговременной динамики растительности, предъявляются достаточно жёсткие требования. Эти данные должны быть стабильны, иметь хорошую географическую привязку и быть предварительно обработаны, то есть очищены от различных помех, шумов, искажений, влияния атмосферы и условий наблюдений и т.д.

Следует также учитывать, что при решении подобных задач приходится иметь дело с большими временными рядами данных, имеющими достаточно большой объем. Обеспечение полноценной и удобной работы с такими рядами является довольно сложной и трудоёмкой задачей. Для её решения необходима организация специализированных архивов данных. Такие архивы должны включать в себя не только систему непосредственно архивации, но и систему автоматизированного получения и усвоения данных, а также систему предоставления данных системам обработки. В данной работе рассматривается разработанная в ИКИ РАН система архивации данных TERRA/MODIS [10], предназначенная для решения вышеперечисленных задач. Данная система создавалась для обеспечения работ по проектам, связанным с анализом динамики растительности на территории северной Евразии [11-13].

Для обеспечения решения задач, возникающих в этих проектах, к создаваемой системе предъявлялись следующие основные требования:

- Высокий уровень автоматизация архивации и предоставления данных для обработки;
- Возможность простого масштабирования;
- Гибкая и удобная настройка цепочек обработки;
- Возможность хранения данных как в непосредственном доступе, так и в удалённых хранилищах и на автономных носителях;
- Обеспечение удалённого контроля и управления функционированием процессов получения данных, их архивации и обработки.

В качестве исходных данных в рассматриваемом архиве используются ежедневные данные TERRA/MODIS уровня обработки Level 2, получаемые из центра LP DAAC [14] (Land Processes Distributed Active Archive Center), входящего в состав USGS (U.S. Geological Survey). Для обеспечения получения оперативной информации в системе реализована процедура автоматической перекачки и усвоения данных из оперативного хранилища данных LP DAAC (<ftp://e0dps01u.ecs.nasa.gov/MOLT>). Для получения исторических данных используется портал EOS Data Gateway [15], они получают в основном на магнитных лентах и также в автоматизированном режиме усваиваются в архивах созданной системы.

По мере поступления в архив исходные данные автоматически проходят предварительную обработку. Алгоритм предварительной обработки спутниковых данных включает в себя несколько последовательных

шагов. На первом этапе из исходных продуктов спутниковых данных исключаются наблюдения с пространственным разрешением хуже заданного, а из нескольких измерений одной и той же точки сделанных с одного витка выбирается одно. Это позволяет уменьшить объем данных для последующей обработки приблизительно в три раза. Следующим этапом обработки является детектирование пикселей изображений облаков и снежного покрова. В основе алгоритма детектирования снежного и облачного покровов лежит использование Нормализованного Разностного Индекса Снега NDSI [16].

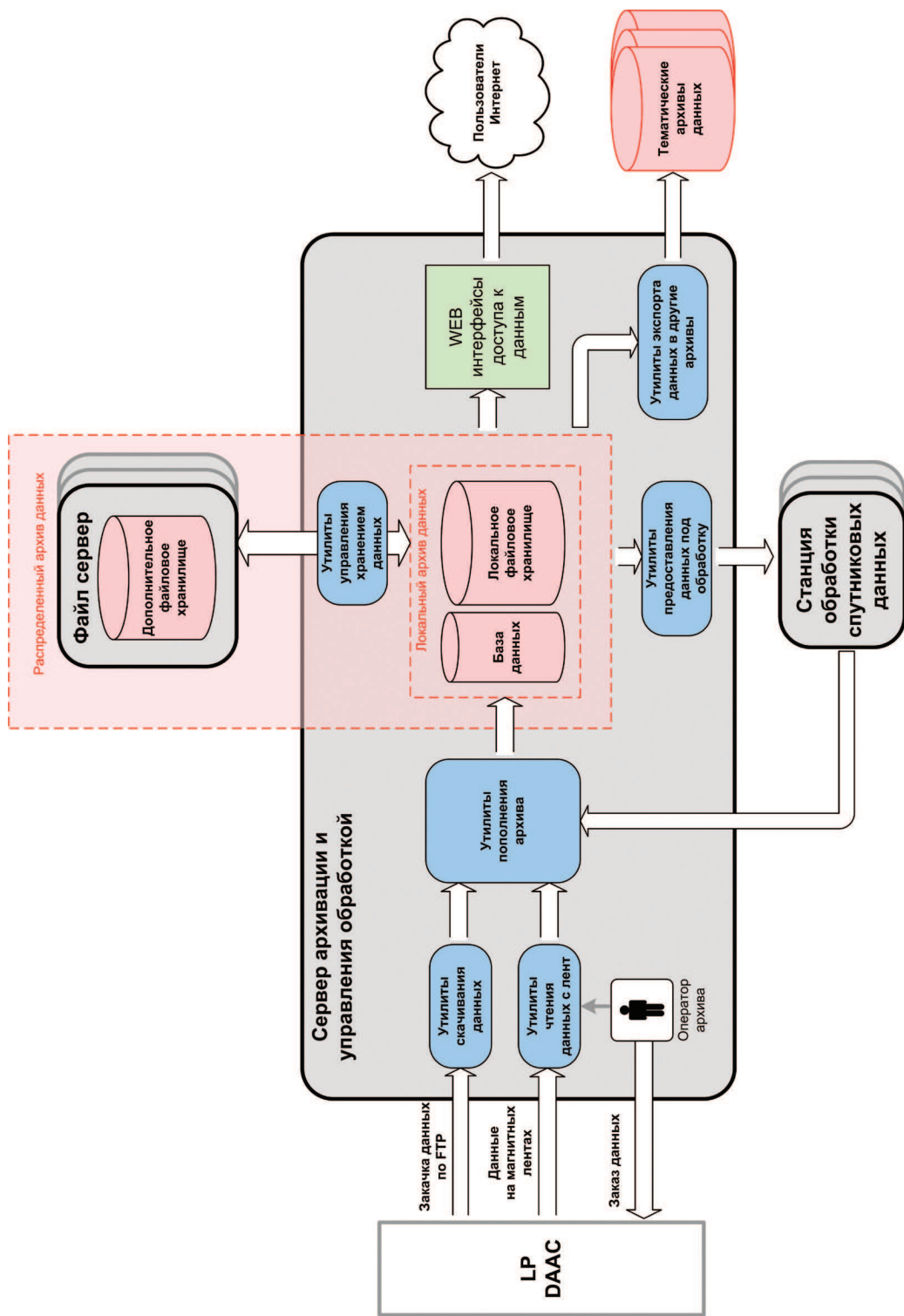


Рис. 1. Функциональная схема архива

Спутниковые данные, как правило, содержат достаточно большое количество пикселей, непригодных для анализа состояния растительности. К таким пикселям в первую очередь относятся пиксели, в которых поверхность закрыта облачностью или снежным покровом, а также пиксели, расположенные на краях скана и имеющие достаточно большие геометрические искажения. В то же время, многие задачи тематической обработки, в том числе задачи анализа динамики растительного покрова, не требуют использования данных ежедневных спутниковых наблюдений, что допускает использование композитных изображений, полученных с использованием данных за некоторый промежуток времени.

Программное обеспечение для ведения архив разработано на основе созданной в ИКИ РАН [17, 18] технологии построения автоматизированных систем хранения спутниковых данных. Функциональная схема архива представлена на рис. 1. Хранение непосредственно файлов в архиве организовано с использованием программного пакета FDB, позволяющего осуществлять различные операции с файлами данных в архиве в соответствии с изменениями БД. Поступающие данные перемещаются в выделенное локальное файловое хранилище на сервере архивации и управления обработкой, информация о них заносится в БД. Из локального хранилища данные могут быть представлены под обработку посредством специальных утилит, предоставлены пользователям через специализированные Web-интерфейсы, а также перемещены в дополнительное файловое хранилище. В качестве дополнительных файловых хранилищ используются специализированные файловые сервера, оборудованные комплектом сменных жёстких дисков. Состав комплектов дисков зависит от текущих задач обработки и предоставления данных. Данные, находящиеся в дополнительных хранилищах, также остаются доступными для пользователя. Таким образом, архитектура системы позволяет легко наращивать мощности по хранению данных в непосредственном доступе при возникновении такой необходимости. На данный момент используется три сервера хранения данных, суммарный объём данных, находящихся в непосредственном доступе составляет около 4,3 терабайта. Общий объём архива превышает 8 терабайт.

Для предоставления данных из архива под обработку была разработана специализированная система выборки данных. В силу специфики обрабатываемых данных, основным требованием к этой системе стала гибкость и лёгкость настроек. Система позволяет формировать несколько очередей обработки различных продуктов за разные интервалы времени. Конфигурируются длины и директории назначения очередей, диапазоны дат, за которые необходимо обработать данные, территориальное покрытие, типы и размеры окна для композитов, необходимые версии исходных данных. Также система позволяет проверять выкладываемые данные на целостность заархивированных файлов и на полноту выкладываемых комплектов. Каждая сформированная очередь может обрабатываться одним или несколькими компьютерами. Их число ограничено только пропускной способностью локальной сети и производительностью дисковой системы сервера управления. На настоящее время в системе используется около десяти параллельно функционирующих обработчиков.

Архитектура системы позволяет не только распараллеливание обработки, но и распараллеливание предоставления данных. В данный момент в системе функционирует один основной сервер архивации и управления и один дополнительный сервер управления, обеспечивающий исключительно выкладку данных под обработку, что позволяет снизить нагрузку на основной сервер и более равномерно загрузить обрабатывающие машины при больших потоках данных, поступающих в систему.

Для доступа к данным архива в системе разработан набор специализированных web-интерфейсов. При создании интерфейсов использовалась технология, разработанная в ИКИ РАН [19]. Функциональность основного интерфейса (http://193.232.9.113/modis_products/) включает в себя возможность фильтрации данных по типу, территориальному покрытию, версии продукта и времени создания продукта. При необходимости есть возможность послать оператору архива запрос на восстановление необходимых данных в режим непосредственного доступа. Также в системе реализован набор служебных интерфейсов, позволяющих получать разнообразную статистику о заполнении архива.

Для контроля функционирования серверов системы используется программный пакет PMS (Process Monitoring System) [20]. На основе использования специализированного WEB интерфейса пользователю

предоставляется информация о протоколах выполнения программ, их возвратных значениях, и, что наиболее важно, о возникновении ошибочных ситуаций, требующих немедленного вмешательства квалифицированного персонала. Для оперативного оповещения о таких сбойных ситуациях используется электронная почта.

Все сервера системы функционируют на ПО, разработанном в ИКИ РАН, и открытом ПО. Из последнего используются ОС FreeBSD версий 5.x и 6.0, интерпретатор Perl версии 5.8.6, web-сервер Apache версии 1.3, сервер баз данных MySQL версии 4.1, программа для доступа к файлам UNIX-сервера из локальной сети Windows Samba версии 2.2.

Обрабатывающие компьютеры функционируют под управлением ОС Windows 2000 или Windows XP. Стандартный набор устанавливаемого на них ПО включает программный пакет обработки спутниковых данных Sputnik, планировщик xv-cron и программу контроля состояния обработки remwatch [21], разработанные в ИКИ РАН, а также программу удалённого управления Windows-компьютером TightVNC.

В настоящее время в архиве имеются следующие продукты обработки спутниковых данных:

- Ежедневные продукты разрешения 250, 500 и 1000 м.;
- Ежедневные маски снега и облачности разрешения 250, 500 и 1000 м.;
- 8-дневные композиты разрешения 250, 500 и 1000 м.;
- 8-дневные композиты NDVI;
- 16-дневные композиты разрешения 250, 500 и 1000 м.;
- 16-дневные композиты NDVI;

Территориальное покрытие данного архива представлено на рис. 2. Вертикальной штриховкой показаны области, по которым имеется покрытие с 2002 года до настоящего времени. Горизонтальной штриховкой показаны области, работы по заполнению историческими данными в настоящее время ведется. В дальнейшем архив должен обеспечить покрытие всей территории России, а затем и всей территории северной Евразии с начала 2001 года.

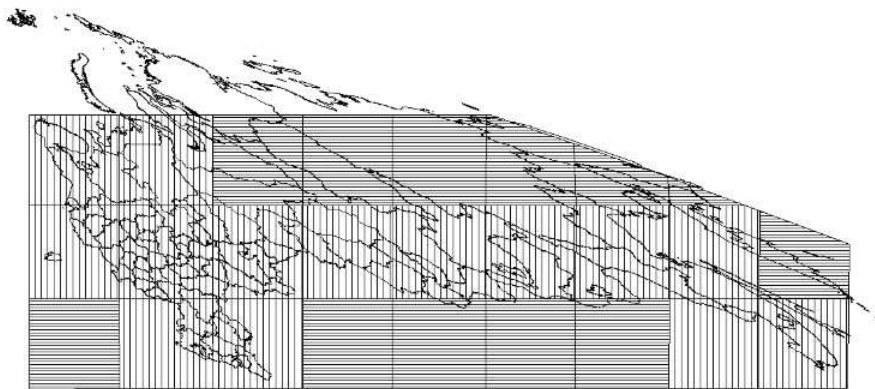


Рис. 2. Территориальное покрытие архива

Литература

1. Султангазин У.М., Муратова Н.Р., Терехов А.Г. Использование Космического мониторинга в Планировании и прогнозировании параметров зернового производства // Современные проблемы дистанционного зондирования Земли из космоса: Физические основы, методы и технологии мониторинга окружающей среды, потенциально опасных объектов и явлений. Сборник научных статей Москва Полиграф сервис, С. 291-297.
2. Justice C.O. et al. An overview of MODIS Land data processing and product status // Remote Sensing of Environment, 2002, №83, P.3-15.

3. *Santhosh K. Seelan et al.* Remote sensing applications for precision agriculture: A learning community approach // *Remote Sensing of Environment*, 2003, №88., P.157-169.
4. *Стивак Л.Ф., Архипкин О.П., Нургалиев С.Г., Шагарова Л.В.* Дистанционная оценка площадей зерновых в Казахстане по данным гиперспектрального радиометра MODIS // *Исследование Земли из Космоса*, 2003, №2. С.80-84.
5. *Wiegand C.L., Richardson A.J., Escobar D.E., Gerbermann A.H.* // *Vegetation Indices in Crop Assessments // Remote Sensing of Environment*, 1991, №35, P.105-119.
6. *Н.М. Вандышева, Г.И. Василенко, А.Ф. Гуров и др.* // Мониторинг сельскохозяйственных земель на базе разномасштабных спутниковых данных // *Исследование Земли из Космоса*, 2003, №3 72-84.
7. *И.Ю. Савин, Т. Негрэ* // О новом подходе к использованию NDVI для мониторинга посевов сельскохозяйственных культур // *Исследование Земли из Космоса*, 2003, №4 91-96.
8. *Нейштадт И.А., Барталев С.А., Ершов Д.М., Лупян Е.А., Савин И.Ю.*, Алгоритмы анализа данных спутниковых наблюдений TERRA-MODIS для мониторинга сельскохозяйственных земель // *Геоинформатика. Международная научно-техническая конференция, посвященная 225-летию МИИГАиК, Москва, 2004, С.205-209.*
9. *С.А. Барталев, Е.А. Лупян, И.А. Нейштадт, И.Ю. Савин* Дистанционная оценка параметров сельскохозяйственных земель по спутниковым данным спектро радиометра MODIS // *Настоящий сборник.*
10. MODIS Technical Specifications <http://modis.gsfc.nasa.gov/news>.
11. *Барталев С.А., Лупян Е.А., Мельник Н.Н., Темников В.А.* Состояние и перспективы развития российской системы спутникового мониторинга сельскохозяйственных земель. // *Настоящий сборник.*
12. *Егоров В.А., Барталев С.А.* Анализ временных серий спутниковых данных SPOT-Vegetation для детектирования поврежденной пожарами растительности Северной Евразии // *Современные проблемы дистанционного зондирования Земли из космоса.* – М.: ИКИ РАН, 2004. - С. 380-387.
13. *Уваров И.А., Барталев С.А., Егоров В.А., Лупян Е.А., Нейштадт И.А., Ховератович Т.С.,* Структура и функциональные возможности информационной системы TerraNorte для поддержки спутникового мониторинга бореальных экосистем // *Настоящий сборник.*
14. LP DAAC Home Page - <http://lpdaac.usgs.gov/>.
15. Earth Observing System Data Gateway, <http://edcimswww.cr.usgs.gov/pub/imswelcome/plain.html>.
16. *Hall D.K., Riggs G.A., Salomonson V.V.* Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data // *Remote Sensing of Environment*, 1995, №54, P.127-140.
17. *Е.А. Лупян, А.А. Мазуров, Р.Р. Назиров, А.А. Прошин, Е.В. Флитман* Универсальная технология построения систем хранения спутниковых данных // *Препринт ИКИ РАН. Пр-2024. М. 2000. 22 с.*
18. *Ефремов В.Ю., Лупян Е.А., Мазуров А.А., Прошин А.А., Флитман Е.В.* Технология построения автоматизированных систем хранения спутниковых данных // *Современные проблемы дистанционного зондирования Земли из космоса: Физические основы, методы и технологии мониторинга окружающей среды, потенциально опасных объектов и явлений. Сборник научных статей Москва Полиграф сервис, 2004 с 437-443.*
19. *Андреев М.В., Ефремов В.Ю., Лупян Е.А., Мазуров А.А., Наглин Ю.Ф., Прошин А.А., Флитман Е.В.* Построение интерфейсов для организации работы с архивами спутниковых данных удаленных пользователей // *Современные проблемы дистанционного зондирования Земли из космоса: Физические основы, методы и технологии мониторинга окружающей среды, потенциально опасных объектов и явлений. Сборник научных статей Москва Полиграф сервис, 2004 с 514-520.*
20. *Ефремов В.Ю., Лупян Е.А., Мазуров А.А., Наглин Ю.Ф., Прошин А.А., Флитман Е.В.* Управление и контроль работоспособности распределенных систем обработки спутниковых данных // *Современные проблемы дистанционного зондирования Земли из космоса: Физические основы, методы и технологии мониторинга окружающей среды, потенциально опасных объектов и явлений. Сборник научных статей Москва Полиграф сервис, 2004 с 467-475.*
21. *Егоров В.А., Ильин В.О., Лупян Е.А., Мазуров А.А., Флитман Е.В.* Возможности построения автоматизированных систем обработки спутниковых данных на основе программного комплекса XV SAT // *Современные проблемы дистанционного зондирования Земли из космоса: Физические основы, методы и технологии мониторинга окружающей среды, потенциально опасных объектов и явлений. Сборник научных статей Москва Полиграф сервис, 2004 с 431-436.*