

Интегрированная информационная система тематической обработки данных дистанционного зондирования Земли на основе автоматной модели

М.В. Энгель¹, В.В. Белов^{1,2*}

¹Институт оптики атмосферы им. В.Е. Зуева СО РАН
634055, г. Томск, пл. Академика Зуева, 1

²Национальный исследовательский Томский государственный университет
634034, г. Томск, пр. Ленина, 36

Поступила в редакцию 24.05.2016 г.

Работа посвящена развитию средств интеграции информационного и программного обеспечения для задач тематической обработки данных дистанционного зондирования Земли. Представлена автоматная модель организации обработки распределенных данных, обеспечивающая интеграцию разнородных распределенных данных и алгоритмов тематической обработки в рамках информационной системы. Используемый авторами подход показан на примере создания прототипа системы интеграции, включающей алгоритм расчета корректирующих атмосферных поправок на основе физической модели формирования оптических изображений земной поверхности.

Ключевые слова: спутниковые данные, распределенные информационные системы, интеграция ресурсов, атмосферная коррекция; satellite data, distributed information systems, integration of resources, atmospheric correction.

Введение

Современная тенденция развития данных дистанционного зондирования Земли (ДЗЗ) и технологий их использования определяется следующими факторами:

- 1) значительное увеличение номенклатуры и объема продуктов на основе данных ДЗЗ;
- 2) появление возможности перехода на веб-технологии поиска и заказа данных;
- 3) начало обмена данными с использованием телекоммуникационных сетей [1].

Основные принципы организации систем хранения и обработки данных ДЗЗ были сформулированы и реализованы в программе NASA Earth Science Enterprise (ESE), принятой с целью комплексного изучения планеты Земля. Ключевыми компонентами этой программы являются система, объединяющая множество спутниковых приборов, предназначенных для мониторинга различных параметров окружающей среды (Earth Observing System – EOS), и созданная для хранения и распространения стандартных продуктов обработки данных спутниковых измерений распределенная компьютерная сеть (Earth Observing System Data and Information System – EOSDIS) [2]. Методы и технологии, разработанные при создании EOSDIS, в той

или иной степени использовались в дальнейшем при построении других масштабных систем обработки данных ДЗЗ.

В России существуют несколько крупных научных центров, в которых разрабатываются технологии хранения и обработки данных ДЗЗ. Работы в этом направлении активно ведутся в Институте космических исследований (ИКИ) РАН, где в конце 90-х гг. прошлого столетия была разработана технология построения системы архивации и долгосрочного хранения спутниковых данных, позволяющая организовывать автоматически пополняющиеся архивы данных с возможностью сетевого доступа [3]. На базе данной технологии была создана система, предоставляющая удаленный доступ к распределенным электронным архивам, расположенным в различных центрах приема и обработки спутниковых данных.

В Институте вычислительных технологий (ИВТ) СО РАН разрабатываются технологии хранения и обработки сверхбольших объемов пространственной информации; организация потоковых вычислений на вычислительных устройствах кластерного типа. На базе ИВТ СО РАН функционирует Центр коллективного пользования данными дистанционного зондирования Земли (ЦКП ДЗЗ СО РАН), на базе которого создана распределенная система приема и обработки спутниковых данных, полученных приемными комплексами ИВТ СО РАН, ИОА СО РАН, ВЦ ДВО РАН и АлГУ [4].

* Марина Владимировна Энгель (angel@iao.ru); Владимир Васильевич Белов (belov@iao.ru).

На основе результатов многолетних измерений, полученных различными спутниковыми системами, создаются информационные системы, обеспечивающие не только хранение и навигацию, но и инструментарий для анализа и управления данными. Примером такого подхода является портал GIOVANNI [5, 6], предназначенный для пространственного, временного и статистического анализа информации, хранящейся в архивах сети EOSDIS. Активно развиваются информационные технологии, направленные на создание сервисов для тематической обработки спутниковой информации. Среди российских разработок такого рода можно назвать технологию GEOSMIS [7], предназначенную для создания интерфейсов работы с большими распределенными многомерными архивами спутниковых данных. На основе данной технологии созданы сервисы, предназначенные для решения различных научных и прикладных задач [8–11].

Анализ современных технологий разработки информационных систем для поддержки решения различных прикладных задач на основе данных ДЗЗ показывает тенденцию развития таких систем в направлении стандартизации и интеграции уже существующего информационного и алгоритмического обеспечения. В данной статье обсуждается подход к разработке информационной системы интеграции распределенных разнородных информационных и алгоритмических ресурсов, основанный на использовании автоматного метода, с целью создания инструментов для решения различных прикладных и научных задач. Реализация предложенного подхода показана на примере разработки прототипа системы, включающей алгоритмы тематической обработки данных ДЗЗ для решения задачи атмосферной коррекции.

Постановка задачи

Система предназначена для интеграции методов тематической обработки данных, базирующихся на следующих условиях:

- математическая модель обработки данных;
- разнородные распределенные данные, необходимые для работы алгоритмов;
- информация может иметь различные характеристики и находиться на распределенных ресурсах, количество которых может быть произвольным и меняться в процессе жизнедеятельности системы.

Для интеграции данных в системе сформулированы следующие требования:

- а) к данным возможен доступ по стандартным сетевым протоколам;
- б) данные записаны с использованием стандартного формата;
- в) предоставлен доступ к метаданным.

Функциональные требования к системе определены следующим образом.

1. Выбор алгоритма тематической обработки.
2. Задание параметров расчета.
3. Определение набора входных данных.
4. Определение источников входных данных.

5. Получение набора входных данных.
6. Оценка релевантности набора входных данных.
7. Проведение расчетов.
8. Предоставление результатов расчетов.

С учетом данных требований предложена информационная модель системы, описывающая работу системы от определения задания на проведение расчетов до получения результатов расчетов, выполненных в автоматическом режиме с использованием разнородных распределенных данных (рис. 1).

В целях описания информационной структуры системы разработана и используется инфологическая схема распределенных данных, которая включает параметры, определяющие идентификацию распределенных источников и доступ к данным: URL, тип данных, приоритет, сетевой протокол, параметры авторизации. Инфологическая схема обеспечивает конвертацию разнородных данных к внутреннему представлению математических моделей, реализующих алгоритмы тематической обработки (рис. 2).

Автоматная модель системы

К решению задачи интеграции существуют различные подходы. Например, в работах [8, 12, 13] для организации обработки распределенных данных используется парадигма событийно-ориентированного программирования или программная логика реализуется в виде набора сценариев и библиотек.

Нами предложена модель интеграции разнородных распределенных данных ДЗЗ и алгоритмов тематической обработки, основанная на использовании теории автоматов. Автоматный подход к построению информационных систем основан на понятии «состояние». При проектировании системы на основе автоматного подхода определяются все возможные состояния, в которых может находиться система, и переходы между ними в результате влияния некоторых событий. Из множества возможных состояний выделяются базовые управляющие состояния, которые определяют логику работы системы. На основе множества базовых состояний задаются входной и выходной алфавиты автоматной модели [14, 15].

Управляющие состояния, в которых может находиться система, определены следующим образом:

- формирование параметров расчета;
- определение входной информации для расчетов на основании заданных параметров;
- получение данных для расчетов из сетевых источников;
- оценка релевантности полученных данных;
- проведение расчетов;
- нормальное завершение выполнения задания;
- аварийное завершение выполнения задания;
- формирование сообщения о выполнении задания.

Управляющие состояния и переходы между ними показаны на UML-диаграмме состояний системы (рис. 3).

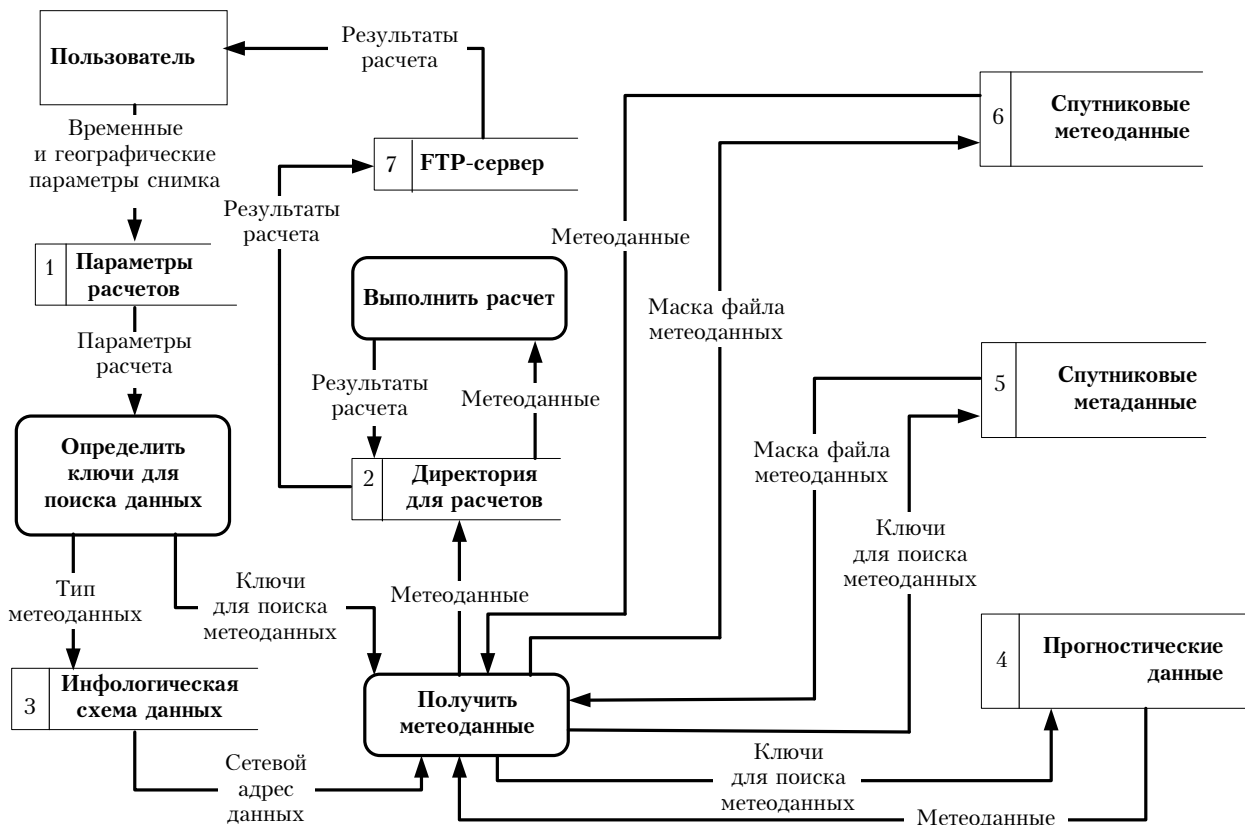


Рис. 1. Информационная модель системы интеграции

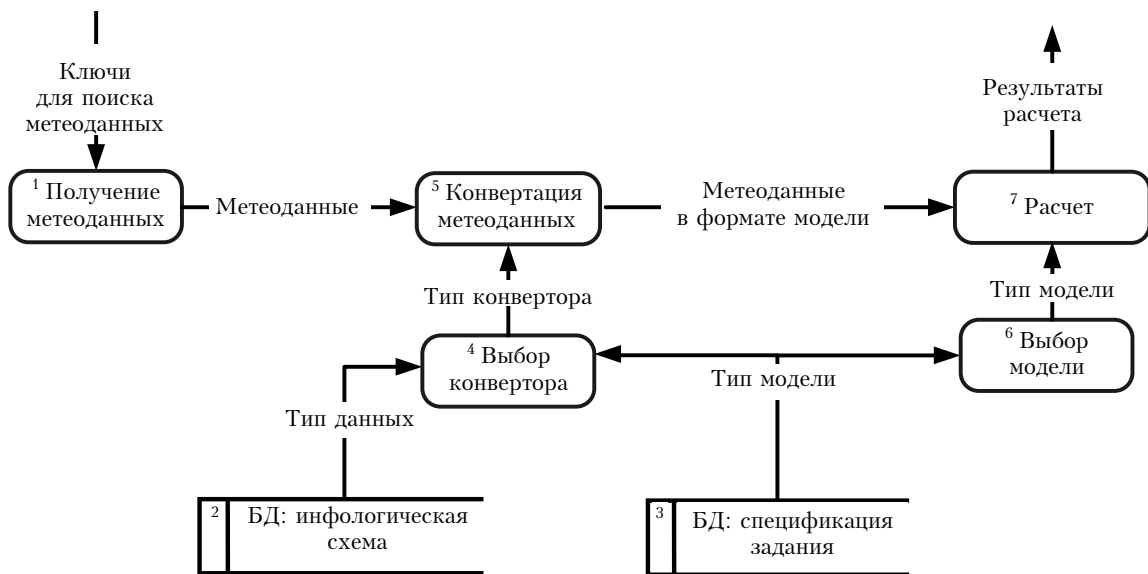


Рис. 2. Схема настройки данных к внутреннему формату модели

Каждое управляющее состояние определяется как конечный автомат. Механизм взаимодействия автоматов реализован в виде обмена сообщениями через общую память. Выбор именно этого способа взаимодействия обусловлен асинхронным характером процессов, выполняемых системой, и, как следствие, невозможностью синхронизировать работу автоматов между собой. Асинхронность порождает

в первую очередь возможной недетерминированностью входной информации, процедура получения которой зависит от различных внешних по отношению к системе факторов и может быть определена только в процессе выполнения. В целом управляющая часть системы представляет собой совокупность автоматов, взаимодействующих между собой через общую память.

Пример разработки на основе автоматного подхода

Реализация предложенной модели показана на примере создания информационной системы атмосферной коррекции (АК) изображений в ИК-каналах спутниковых зондировщиков [16, 17].

Процедура АК заключается в удалении атмосферных фонов рассеяния из измерений спутниковых приборов. Наиболее точная коррекция возможна на основе физического подхода, для реализации которого необходимо выполнение следующих условий:

- наличие математической модели атмосферы;
- информация о параметрах оптико-метеорологического состояния атмосферы в момент проведения ДЗЗ из космоса.

Эффективный подход к проблеме получения оптимального набора различных оптико-метеорологических параметров предполагает комплексное использование результатов спутникового зондирования атмосферы и данных моделирования. Выбор данных ДЗЗ в качестве возможных источников оптико-метеорологических атмосферных параметров, а также конкретной спутниковой системы, на данный момент оптимальной для реализации физического подхода к АК, был предложен и обоснован в работах [18, 19]. Здесь в качестве данных о состоянии атмосферы предлагается использовать тематические продукты, полученные на основе измерений спектрорадиометра MODIS [20], а в качестве альтернативного источника метеоинформации – прогностические модели NCEP (National Centers for Environmental Prediction) [21]. Ресурсы для распространения данных обоих типов представлены в Интернете. Одним из основных сетевых ресурсов для распространения спутниковых данных об атмосфере можно назвать LAADS Web (Level 1 and Atmosphere Archive and Distribution System), Goddard Space Flight Center, NASA [22]. Тематические атмосферные продукты MODIS представлены в архивах LAADS Web за весь период измерений.

Ресурс LAADS Web может быть интегрирован в рамках информационной системы АК в качестве источника информации о параметрах оптико-метеорологического состояния атмосферы, поскольку удовлетворяет сформулированным ранее требованиям, а именно:

- для записи и хранения данных используются стандартные документированные форматы (HDF-EOS, GRIB);
- данные в формате HDF-EOS записаны вместе с метаданными;
- к данным предоставлен доступ по протоколу ftp.

Распределенная инфраструктура ЦКП ДЗЗ СО РАН и электронные архивы спутниковых данных ИОА СО РАН также зарегистрированы в системе как сетевые информационные ресурсы.

Необходимым условием использования при расчетах информации, полученной из распределен-

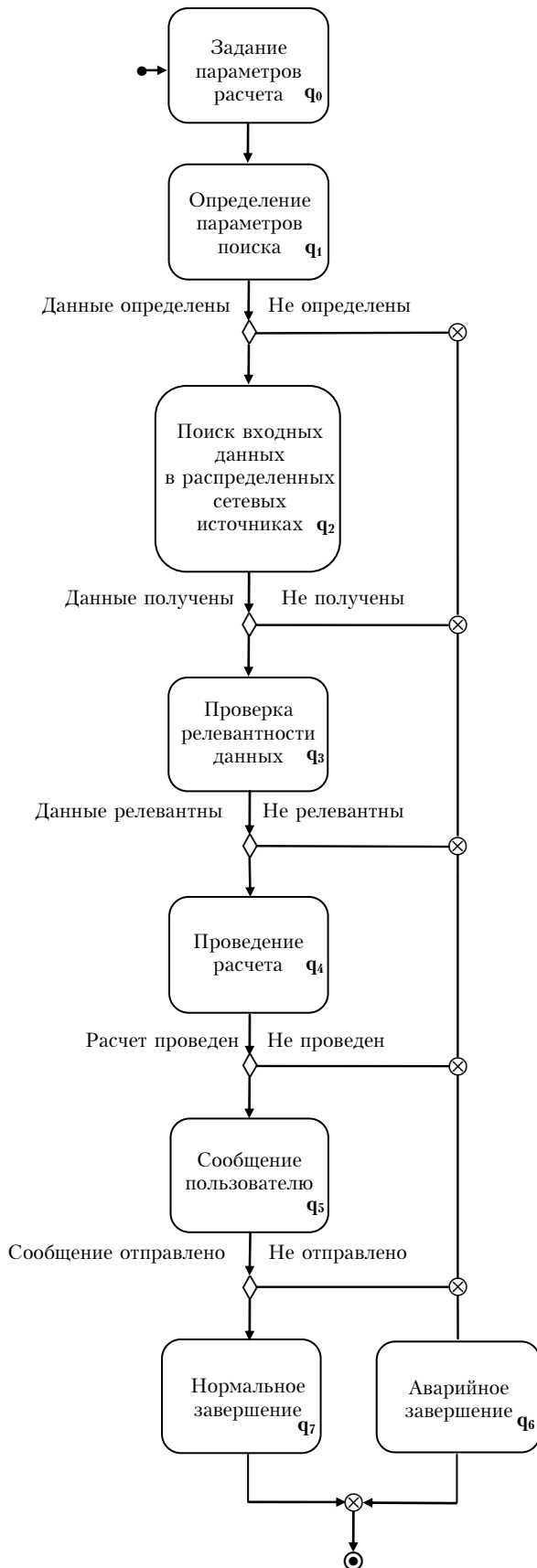


Рис. 3. Диаграмма состояний системы

ных сетевых источников, является оценка ее релевантности. Релевантность определяется по двум критериям:

1) критерий соответствия спутникового снимка заданным временным и географическим параметрам;

2) критерий качества спутниковых измерений.

Первый критерий базируется на использовании метаданных. Для оценки качества спутниковой метеоинформации MODIS предложен метод, основанный на сравнении температур земной поверхности, восстановленных по результатам одновременных спутниковых измерений в ИК-каналах [23].

При проведении расчетов в качестве модели переноса излучения используется программа MODTRAN. В результате работы информационной системы для каждого пикселя спутникового изображения производится вычисление корректирующих атмосферных поправок: значения функции пропускания, интенсивности собственного излучения атмосферы, интенсивности ослабленного атмосферой излучения поверхности, интенсивности отраженного от поверхности излучения атмосферы и Солнца, — восстановленных для каждой точки измерений спутникового прибора для заданных каналов сенсора.

На первом этапе разработки в системе реализован алгоритм расчета корректирующих атмосферных поправок для результатов измерений ИК-каналов спутниковой аппаратуры, использующий при расчетах распределенные разнородные данные. В настоящий момент ведутся работы по интеграции в систему алгоритма атмосферной коррекции в видимом диапазоне длин волн, сочетающего в себе учет основных факторов, влияющих на перенос излучения [24, 25]. Дальнейшее развитие системы предполагается в направлении интеграции новых алгоритмов тематической обработки данных ДЗЗ и типов данных, а также увеличения производительности системы за счет использования технологий параллельных вычислений.

Заключение

Рассмотрен подход к интеграции информационных и алгоритмических ресурсов, основанный на использовании автоматной модели и инфологической схемы данных, с целью создания программных инструментов тематической обработки данных дистанционного зондирования Земли. Использование автоматного подхода позволяет автоматизировать проведение процедур обработки данных. Вычислительная среда в автоматическом режиме планирует процесс выполнения задачи таким образом, чтобы процесс максимально соответствовал установленному критерию оптимальности. Автоматная модель системы обеспечивает расширяемость, наблюдаемость и сервис-ориентированный характер системы.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 15-07-06811 А.

1. *Луян 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.
2. *Ramapriyan H., Behnke J., Sofinowski E., Lowe D., Esfandiari M.* Evolution of the Earth Observing System (EOS) Data and Information System (EOSDIS) // Standard-Based Data and Information Systems for Earth Observation. Springer, 2010. P. 63–92.
3. *Луян Е.А., Мазуров А.А., Назиров Р.Р., Прошин А.А., Флитман Е.В.* Технологии построения автоматизированных информационных систем сбора, обработки, хранения и распространения спутниковых данных для решения научных и прикладных задач // Современные проблемы дистанционного зондирования Земли из космоса. 2004. Т. 1, вып. 1. С. 81–88.
4. *Шокин Ю.И., Антонов В.Н., Добрецов Н.Н., Кухтенко 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.
5. *The GES-DISC Interactive Online Visualization AND aNalysis Infrastructure (Giovanni).* URL: <http://giovanni.gsfc.nasa.gov/>
6. *Acker J.G., Leptoukh G.* Online Analysis Enhances Use of NASA Earth Science Data // EOS, Trans. American Geophysical Union. 2007. V. 88, iss. 2. P. 14–17.
7. *Толтин В.А., Балашов И.В., Ефремов В.Ю., Луян 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.
8. *Информационная система дистанционного мониторинга Федерального агентства лесного хозяйства.* URL: https://nffc.aviales.ru/main_pages/index.shtml
9. *Инструмент научного анализа данных спутниковых наблюдений биосферы BEGA-science.* URL: <http://sci-vega.ru>
10. *Спутниковый сервис мониторинга Мирового океана SEE THE SEA.* URL: <http://ocean.smislab.ru/static/index.sht>
11. *Информационный сервис «Дистанционный мониторинг активности вулканов Камчатки и Курил» VolSatView.* URL: <http://volcanoes.smislab.ru/static/index.sht>
12. *Ермаков Д.М., Емельянов К.С., Саворский В.П., Чернушиц А.П.* Реализация событийно управляемой архитектуры быстрого коллективного доступа к информационным ресурсам ДЗЗ на базе технологии Stream Handler // Современные проблемы дистанционного зондирования Земли из космоса. 2013. Т. 10, № 4. С. 118–126.
13. *Балашов И.В., Халикова О.А., Бурцев М.А., Луян Е.А., Матвеев А.М.* Организация автоматического получения наборов информационных продуктов из центров архивации и распространения спутниковых и метеоинформации // Современные проблемы дистанционного зондирования Земли из космоса. 2013. Т. 10, № 3. С. 9–20.
14. *Цетлин М.Л.* Исследования по теории автоматов и моделированию биологических систем. М.: Наука, 1969. 316 с.

15. Полицарпова Н.И., Шальто А.А. Автоматное программирование. 2-е изд. СПб.: Питер, 2011. 176 с.
16. Энгель М.В., Афонин С.В., Белов В.В. Интеграция междисциплинарных данных и результатов исследований, базирующихся на пространственных характеристиках и признаках. Инфраструктура пространственных данных, метаданные, веб- и геосервисы // Геоинформационные технологии и математические модели для мониторинга и управления экологическими и социально-экономическими системами / ред. кол. Ю.И. Шокин и др.; под ред. И.Н. Ротановой; Рос. акад. наук. Сиб. отд.-е. Институт водных и экологич. проблем. Барнаул: Пять плюс, 2011. С. 84–163.
17. Энгель М.В., Афонин С.В. Программное обеспечение информационно-вычислительной системы расчета данных для проведения атмосферной коррекции спутниковых изображений // Оптика атмосф. и океана. 2014. Т. 27, № 7. С. 628–633.
18. Афонин С.В., Соломатов Д.В. Методика учета оптико-метеорологического состояния атмосферы для решения задач атмосферной коррекции спутниковых ИК-измерений // Оптика атмосф. и океана. 2008. Т. 21, № 2. С. 147–153.
19. Афонин С.В. К вопросу о применимости восстановленных из космоса метеоданных MODIS для атмосферной коррекции спутниковых ИК-измерений // Оптика атмосф. и океана. 2010. Т. 23, № 8. С. 684–688; *Afonin S.V. Applicability of Space – Derived Meteorological Data to Atmospheric Correction of Satellite Infrared Measurements // Atmos. Ocean. Opt.* 2011. V. 24, N 1. P. 56–63.
20. *Level-2 MODIS Atmosphere Products.* URL: <http://modis-atmos.gsfc.nasa.gov/products.html>
21. *National Centers for Environmental Prediction.* URL: <http://www.ncep.noaa.gov/>
22. *LAADS Web.* URL: <https://ladsweb.nascom.nasa.gov/>
23. Энгель М.В., Афонин С.В., Белов В.В. Методика предварительной оценки точности метеоданных MODIS при атмосферной коррекции спутниковых ИК-измерений // Оптика атмосф. и океана. 2013. Т. 26, № 8. С. 692–694; *Engel' M.V., Afonin S.V., Belov V.V. Method for Preliminarily Estimating the Accuracy of MODIS Meteorological Data in the Atmospheric Correction of Satellite IR Measurements // Atmos. Ocean. Opt.* 2014. V. 27, N 1. P. 68–70.
24. Белов В.В., Тарасенков М.В. О точности и быстродействии RTM-алгоритмов атмосферной коррекции спутниковых изображений в видимом и УФ-диапазонах // Оптика атмосф. и океана. 2013. Т. 26, № 7. С. 564–571; *Belov V.V., Tarasenkov M.V. On the Accuracy and Operation Speed of RTM Algorithms for Atmospheric Correction of Satellite Images in the Visible and UV Ranges // Atmos. Ocean. Opt.* 2014. V. 27, N 1. P. 54–61.
25. Тарасенков М.В., Белов В.В. Комплекс программ восстановления отражательных свойств земной поверхности в видимом и УФ-диапазонах // Оптика атмосф. и океана. 2014. Т. 27, № 7. С. 622–627; *Tarasenkov M.V., Belov V.V. Software Package for Reconstructing Reflective Properties of the Earth' Surface in the Visible and UV ranges // Atmos. Ocean. Opt.* 2015. V. 28, N 1. P. 89–94.

M.V. Engel, V.V. Belov. Integrated information system based on the automated model for thematic processing of the Earth remote sensing data.

This paper deals with the development of tools for integrating data and software for thematic processing of Earth remote sensing data (ERS). An automaton model of organizing distributed data processing, ensuring the integration of heterogeneous distributed data and algorithms of thematic processing within the information system is presented. The approach is described by the example of the prototype of an integrated system including an atmospheric correction algorithm for calculating the corrections on the basis of a physical model of the formation of optical images of the Earth's surface.