

Система пакетной обработки океанологических спутниковых данных

С.В. Шеберстов

*Институт океанологии им. П.П.Ширшова РАН, Москва 117997, Россия
E-mail: sheberst@yandex.ru*

В работе представлен разработанный в ИО РАН программный комплекс (ПК) для обработки, визуализации и организации хранения океанологических спутниковых данных. В состав ПК входят программы для обслуживания базы данных, содержащей спутниковые файлы и результаты судовых измерений, и программы для обработки больших массивов данных спутниковых сканеров цвета океана и других океанологических данных. ПК содержит также головную программу, предназначенную для визуализации спутниковых и судовых данных, составления запросов к базе данных и формирования с помощью графического интерфейса управляющих файлов для программ пакетной обработки. Набор программ пакетной обработки включает модули для расчета новых продуктов и записи результата в форматах HDF и NetCDF, усреднение данных, экспорт карт в заданных проекциях, построение временных рядов и т. д. Особенности программ данной системы – динамическая загрузка библиотек, которые могут быть разработаны пользователями и внедрены без перекомпиляции программ, и наличие удобного и достаточно универсального языка управления заданиями. Приводятся примеры применения ПК для решения океанологических задач, связанных с разработкой биооптических алгоритмов и статистического анализа временных рядов спутниковых наблюдений.

Ключевые слова: программное обеспечение, океанологические данные, спутниковые данные, алгоритмы, интерпретатор

Введение

Программный комплекс, представленный в настоящей работе, предназначен для обработки, визуализации и организации хранения океанологических спутниковых данных, таких как спектры выходящего из воды излучения, биооптические характеристики, температура поверхности океана, скорость ветра, состояние ледового покрова и т. п.

В состав ПК входят программы для массовой обработки спутниковых данных, реализующие следующие функции:

- генерация производных продуктов с возможностью импорта разнородных данных в различных форматах (HDF, NetCDF, ASCII, N1, ...);
- создание файлов уровня обработки L3 путем отображения данных на различные пространственные и временные сетки;
- построение временных рядов различных геофизических характеристик усредненных по заданному региону; границей региона может служить произвольный замкнутый контур;
- статистический анализ указанных временных рядов, в том числе анализ межгодовых трендов и обнаружение взаимозависимостей временного хода различных параметров, обусловленных климатическими изменениями;
- сопоставление спутниковых и судовых данных;
- создание наборов географических карт в заданных проекциях с наложением векторных объектов (береговые линии, разрезы, полигоны, станции).

Формирование подлежащих обработке наборов спутниковых и судовых данных может осуществляться с помощью входящей в состав системы реляционной базы данных.

В состав ПК входят также головная программа (управляющий модуль), набор программ пакетной обработки, утилиты (вспомогательные программы) и конверторы – программы для преобразования форматов спутниковых файлов.

Возможности создания процедур потоковой обработки данных сегодня существуют во многих программных комплексах, ориентированных на работу с данными дистанционного зондирования; в первую очередь следует отметить программные комплексы SeaDAS (Baith et.al., 2001) и BEAM-VISAT (<http://www.brockmann-consult.de/cms/web/beam>), в которых задачи пакетной обработки решаются с использованием языков программирования, таких как Java, Python, bash, Perl и др. Особенности предлагаемой системы – простота использования, интуитивно ясный язык, отсутствие необходимости использования языков программирования, встроенная база данных и возможность автоматической генерации заготовки скрипта.

Первая версия данной системы программ была разработана в 2007 г. и представлена в работе (Sheberstov, Lukyanova, 2007), в которой основное внимание было уделено описанию головной программы и базы данных. В настоящей работе отражены изменения и дополнения новой версии и более подробно рассмотрены программы пакетной обработки и язык управления заданиями.

Головная программа

Головная программа системы разработана как приложение для ОС Windows с многооконным интерфейсом на основе технологии MDI (multi-document interface) с применением MFC (Microsoft Foundation Classes). Программа может работать с документами четырех типов – файлы судовых измерений, графические файлы, спутниковые файлы в формате HDF и текстовые документы с результатами запросов к базе данных. Документы всех четырех типов могут быть открыты одновременно.

Программа содержит графический интерфейс для отображения данных всех рассмотренных выше типов – спутниковые данные могут быть представлены в виде географических карт в одной из проекций по выбору пользователя, на тех же картах могут отображаться точечные и линейные объекты (положение станций, на которых производился сбор данных, траектории рейсов).

Одной из важных функций программы является формирование с помощью графического интерфейса управляющих файлов для программ пакетной обработки.

Более подробное описание головной программы и базы данных см. в (Sheberstov, Lukyanova, 2007).

База данных

База данных реализована в архитектуре клиент – сервер под управлением СУБД MySQL. Сервер под управлением ОС Linux, UNIX или Windows содержит часть архива

спутниковых файлов в формате HDF и реляционную базу данных, содержащую следующие таблицы:

- судовые данные в формате базы данных SeaBASS (Werdell , Bailey, 2002);
- метаданные спутниковых файлов – координаты угловых точек, время съемки, список продуктов, версия ПО и т. п.;
- метаданные графических файлов в формате GIF, JPG, BMP, или PNG;
- вспомогательные таблицы, используемые для реализации графического интерфейса формирования запроса.

Спутниковые и графические файлы не конвертируются к формату MySQL и хранятся на сервере отдельно, поскольку их размер может превышать максимально допустимые для MySQL значения.

Клиентская часть программного обеспечения состоит из следующих программ: программа формирования запроса, FTP-клиент, программа для загрузки необходимых спутниковых файлы с компакт-диска.

Составление запроса к базе данных на языке SQL может выполняться как в ручном режиме, так и с помощью графического интерфейса пользователя (GUI). Графический интерфейс позволяет формировать наборы судовых и спутниковых данных по координатам, по времени выполнения измерений, по набору продуктов спутниковых файлов. Кроме того существует возможность предварительного отбора файлов для процедуры сопоставления судовых и спутниковых данных программой пакетной обработки *seatrue*.

В результате выполнения запроса формируется список выбранных файлов, с помощью FTP-клиента эти файлы загружаются на компьютер пользователя.

Далее, посредством механизма сообщений (messages) Windows, этот список передается головной программе, которая в свою очередь формирует командный файл для программ пакетной обработки.

Программы пакетной обработки

Командные файлы всех перечисленных ниже программ имеют сходную структуру: каждый такой файл должен содержать по меньшей мере три раздела: *files* – список спутниковых файлов; *common* – имена других входных файлов, файлов, содержащих результаты обработки, координаты региона и т. д.; *products* – последовательность команд для вычисления новых продуктов. Синтаксис этих команд изложен в следующем разделе.

Набор программ пакетной обработки постоянно пополняется; ниже приведен список по состоянию на конец 2015 г.:

- *l2tol2* – предназначена для расчета новых продуктов. Для каждого файла из входного набора создается новый файл с новым набором продуктов. Каждому пикселю нового файла соответствует один пиксель входного файла с сохранением географической привязки;

- *l2tol3* – усреднение группы входных файлов по равномерной градусной сетке в заданной области;
- *asciiexport* – экспорт спутниковых данных в текстовом формате;
- *imagex* – создание наборов географических карт в заданных проекциях с наложением векторных объектов;
- *mctable* – построение временных рядов для заданной точки или региона;
- *seatruth* – сопоставление судовых и спутниковых данных;
- *mimport* – импорт данных с сохранением географической привязки;
- *fillgaps* – заполнение пробелов, пространственная интерполяция;

Для каждой программы пакетной обработки существуют три варианта – Windows приложение с графическим интерфейсом, консольное приложение Windows и консольное приложение для Linux.

Язык управления заданиями

Для управления различными модулями системы разработан единый язык управления заданиями, который позволяет любой программе пакетной обработки из состава ПК генерировать новые продукты как функции продуктов, содержащихся в исходных файлах. Если алгоритм расчета нового продукта может быть представлен одной формулой, для интерпретации этой формулы используется встроенный синтаксический анализатор. В противном случае алгоритм расчета может быть загружен из динамически загружаемой библиотеки (DLL в операционной системе MS Windows, shared objects в UNIX – Linux). Включение нового алгоритма в систему не требует перекомпиляции программ.

Список разработанных пользователем алгоритмов и библиотек, в которые включены эти алгоритмы, содержится в текстовом файле *globalproductlist.dat*, размещаемом в той же директории, что и выполняемые модули программного комплекса. Каждая строка файла содержит описание одной функции: имя функции, имя библиотеки DLL (dll=), список аргументов (arguments=...).

Основная информация о выполняемых действиях содержится в секции *products* командного файла. Каждая строка этой секции содержит оператор присвоения вида <имя переменной>= <формула> и набор пар вида <ключевое слово> = <значение>, разделенных знаком “:”. Здесь <формула> – математическое выражение, содержащее переменные и константы. Переменные – это имена продуктов спутникового файла, имена столбцов таблицы судовых данных и величины, вычисленные в предыдущих строках секции *products* командного файла.

Последовательность действий при интерпретации командного файла представлена на *рис. 1*. В первую очередь система читает командный файл, затем, если необходимо, файл *globalproductlist.dat*. Программа *seatrue* должна, кроме указанных двух файлов, прочитать таблицу судовых данных, содержащую координаты станций и время измерений для каждой

станции. После этого синтаксический анализатор извлекает имена переменных и функций из каждой строки секции *products* командного файла, создает таблицу переменных и отдельно таблицу функций. Далее для каждой функции выполняется поиск соответствующей библиотеки DLL, загрузка библиотеки в оперативную память, поиск функции в библиотеке и загрузка указателя на функцию в таблицу функций.

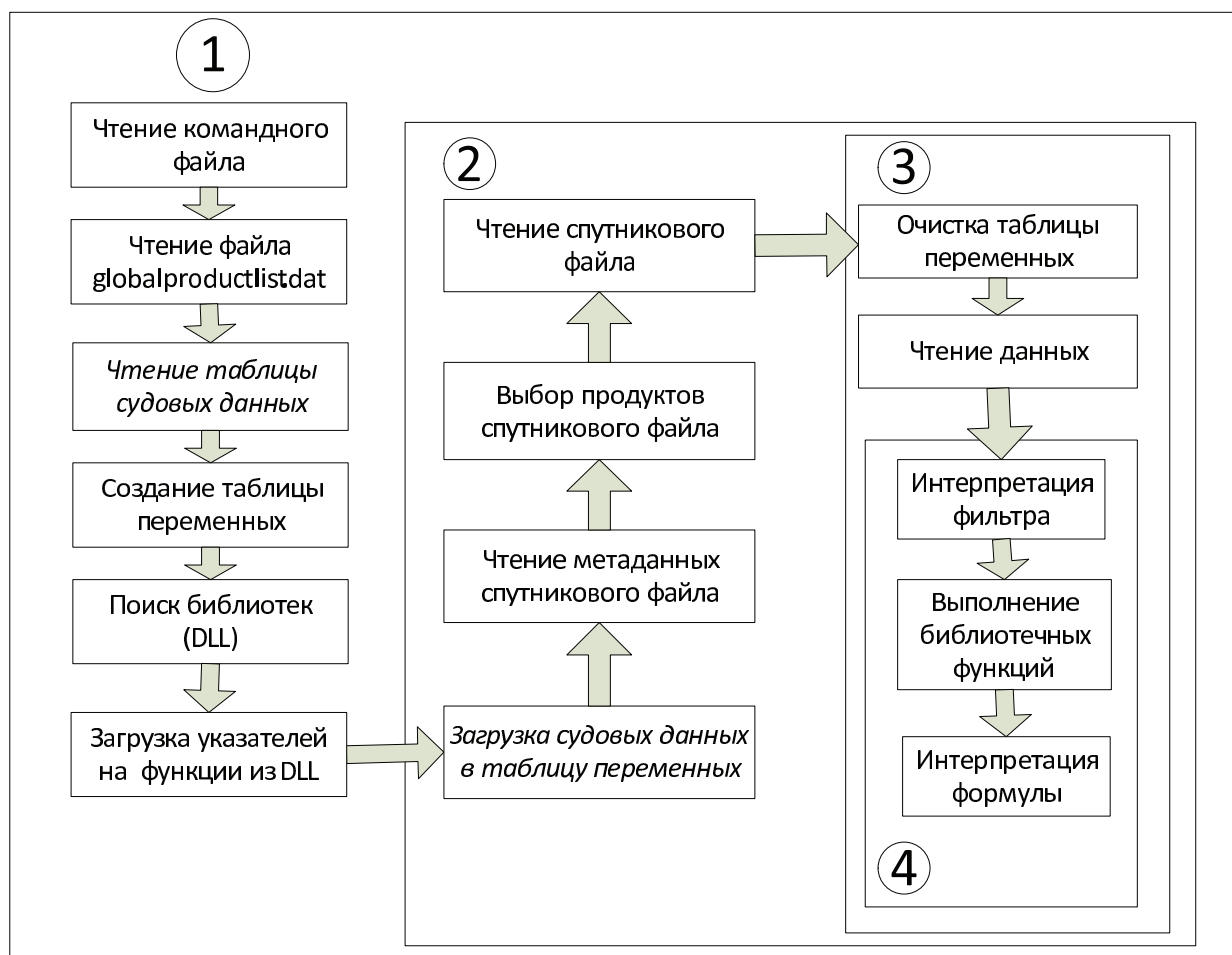


Рис. 1. Последовательность операций при выполнении командного файла.
 1 – операции, выполняемые однократно для всего набора данных, 2 – для каждого файла,
 3 – для каждого пикселя, 4 – для каждого продукта. Курсивом выделены операции,
 выполняемые только при сопоставлении спутниковых и судовых данных

Таблица переменных – это набор троек (имя, тип, значение). Целью дальнейших операций, показанных на рис. 1, является присвоение корректных значений для всех переменных для каждого пикселя входного спутникового файла и каждой строки секции *products* командного файла, после чего возможно применение стандартного интерпретатора формул.

Использование полученного результата зависит от программы – он может быть записан в текстовый файл, включен в файл формата HDF или NetCDF, использован для усреднения, статистического анализа или построения цветного изображения в графическом формате. Таким образом, результаты обработки могут быть представлены в виде файлов формата HDF или NetCDF, изображений форматов bmp, gif, png, jpeg, а также в виде таблиц, столбцы которых могут содержать координаты и время, значения параметров спутникового изобра-

жения в пикселе или пространственное и/или временное среднее по области, ограниченной произвольным (заданным пользователем) контуром. Многочисленные примеры результатов обработки можно найти на сайте Лаборатории оптики океана ИО РАН <http://optics.ocean.ru>

Использование системы

В настоящее время система включает следующие библиотеки, разработанные в Лаборатории оптики океана ИО РАН.

1. Набор биооптических алгоритмов для данных сканеров цвета океана SeaWiFS, MODIS, MERIS, VIIRS. Содержит региональные алгоритмы для расчета показателей обратного рассеяния взвесью (Копелевич и др., 2010), показателя поглощения желтым веществом (Вазюля и др., 2014), концентрации кокколитофорид в Черном море (Kopelevich et al., 2014).

2. Программы для оценки проникновения солнечной радиации в толщу воды и расчета фотосинтетически активной радиации (PAR) (Вазюля, Копелевич, Шеберстов, 2009).

3. Алгоритмы атмосферной коррекции для сканеров цвета SeaWiFS и MODIS (Копелевич и др., 2009), (Лихачева, Копелевич, Шеберстов, 2009).

С помощью данного комплекса программ построен Атлас «Биооптические характеристики морей России по данным спутниковых сканеров цвета» (Kopelevich et al., 2013), выполнена валидация некоторых биооптических алгоритмов для морей России и Атлантического океана (Копелевич и др., 2009), проанализированы сезонные и межгодовые изменения биооптических характеристик и температуры поверхности океана в Черном море (Burenkov, Kopelevich, Sheberstov, 2011) и в Атлантическом океане (Копелевич, Шеберстов, 2010), (Шеберстов, Копелевич, Лукьянова, 2011).

Заключение

Опыт использования системы показал, что она является весьма простой и эффективной при решении стандартных задач спутниковой океанологии. Благодаря наличию базы данных, содержащей как спутниковые, так и судовые данные, встроенному в головную программу генератору скриптов и прозрачному синтаксису этих скриптов, время, необходимое для конфигурации системы для решения конкретной задачи, существенно меньше, чем в других системах обработки спутниковых данных (при условии, что для каждого продукта известна формула или скомпилирована DLL).

Для простых алгоритмов использование встроенного интерпретатора формул лишь несущественно увеличивает время выполнения характерных задач. При реализации сложных алгоритмов использование библиотек DLL, компилируемых отдельно от перечисленных выше программ пакетной обработки, позволяет применять средства повышения производительности вычислительных систем за счет использования параллельной обработки данных.

Работа выполнена за счет гранта Российского научного фонда (проект №14-17-00800), предоставленного через Институт океанологии им. П.П. Ширшова РАН .

Литература

1. *Вазюля С.В., Копелевич О.В., Шеберстов С.В.* Оценка баланса солнечного излучения в мелководном море по спутниковым данным о цвете вод. // Современные проблемы дистанционного зондирования Земли из космоса. 2009, Т. 6. № 1. С. 339–348.
2. *Вазюля С.В., Копелевич О.В., Шеберстов С.В., Артемьев В.А.* Оценка по спутниковым данным показателей поглощения окрашенного органического вещества и диффузного излучения в водах Белого и Карского морей // Современные проблемы дистанционного зондирования Земли из космоса. 2014. Т. 11. № 4. С. 31–41.
3. *Копелевич О.В., Буренков В.И., Шеберстов С.В., Прохоренко О.В.* Разработка региональных алгоритмов атмосферной коррекции данных спутниковых сканеров цвета. // Современные проблемы дистанционного зондирования Земли из космоса. 2009. Т. 6. № 1. С. 400–408.
4. *Копелевич О.В., Буренков В.И., Вазюля С.В., Шеберстов С.В.* Региональные алгоритмы расчета по спутниковым данным биооптических характеристик поверхностного слоя моря // Физические, геологические, и биологические исследования океанов и морей. М.: Научный мир, 2010. С. 115–135.
5. *Копелевич О.В., Шеберстов С.В.* Оценка по спутниковым данным межгодовых и сезонных изменений температуры поверхности Атлантического океана и концентрации хлорофилла в поверхностном слое в период 1998–2008 // Современные проблемы дистанционного зондирования Земли из космоса. 2010. Т. 7. № 1. С. 238–247.
6. *Лихачева М.В., Копелевич О.В., Шеберстов С.В.* Коррекция данных спутникового сканера MODIS на солнечный блик без использования дополнительных данных о скорости ветра. // Современные проблемы дистанционного зондирования Земли из космоса. 2009. Вып. 6. Т. 1. С. 421–428.
7. *Шеберстов С.В., Копелевич О.В., Лукьянова Е.А.* Анализ межгодовых трендов температуры поверхности океана и концентрации хлорофилла в Атлантическом океане по спутниковым данным // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 4. С. 274–282.
8. *Baith K., Lindsay R., and Gary Fu G., McClain Ch. R.,* Data Analysis System Developed for Ocean Color Satellite Sensors, *Eos*, 2001. Vol. 82. No. 18. 202 p.
9. *Burenkov V.I., Kopelevich O.V., Sheberstov S.V.* Seasonal and interannual variability of the biooptical characteristics of the Black Sea from satellite data // Proc. VI Int. Conf. “Current problems in Optics of Natural Waters (ONW 2011)”, St.Petersburg, September 6-9, 2011. St.Petersburg: Nauka, 2011. P. 140-142.
10. *Kopelevich O.V., Sheberstov S.V., Burenkov V.I., Vazyulya S.V., Sahling I.V.,* Bio-optical characteristics of the Russian seas from satellite ocean color data of 1998–2012 // Proceedings VII International Conference: Current Problems in Optics of Natural Waters. St.Petersburg: Nauka. 2013. P. 168–171.
11. *Kopelevich O., Burenkov V., Sheberstov S., Vazyulya S., Kravchishina M., Pautova L., Silkin V., Artemiev V., Grigoriev A.* Satellite monitoring of coccolithophore blooms in the Black Sea from ocean color data // Remote Sensing of Environment. 2014. Vol. 146. P. 113–123.
12. *Sheberstov S.V., Lukyanova E.A.,* A System for Acquisition, Processing and Storage of Satellite and Field Biooptical Data // Proceedings of IV International Conference: Current Problems in Optics of Natural Waters. Nizhny Novgorod, September 11–15. 2007. P. 179–183.
13. *Werdell P.J., Bailey S.W.* The SeaWiFS Bio-Optical Archive and Storage System (SeaBASS) : Current Architecture and Implementation // NASA/TM 2002–211617, Goddard Space Flight Center, Greenbelt, Maryland, 2002.

System for batch processing of oceanographic satellite data

S.V. Sheberstov

P.P. Shirshov Institute of Oceanology RAS, Moscow 117997, Russia

E-mail: sheberst@yandex.ru

The paper presents a software package for processing, visualization and storage organization of oceanographic satellite data developed at IO RAS. The software includes programs for maintenance of a database of satellite files and ship data and programs for batch processing of large sets of data from ocean color scanners and other oceanographic data. The software package also contains the main program intended for the visualization of data, querying the database, and creating scripts for batch processing using GUI. The set of programs for batch processing includes modules for the calculation of new products and writing the result in HDF and/or NetCDF format, averaging data, export maps in a given projection, building time series, etc. Essential features of this software are dynamic loading of libraries that can be developed and implemented by users without having to recompile programs, and the availability of convenient and versatile job control language. Examples of application the software to oceanographic problems related to the development of bio-optical algorithms and statistical analysis of time series of satellite observations are presented.

Keywords: software, oceanographic data, satellite data, algorithms, interpreter

References

1. Vazyulya S.V., Kopelevich O.V., Sheberstov S.V., Otsenka balansa solnechnogo izlucheniya v melkovodnom more po sputnikovym dannym o tsvete vod (Assessment of the budget of solar radiation in shallow waters from satellite data), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2009, Vol. 6, No. 1, pp. 339–348.
2. Vazyulya S.V., Kopelevich O.V., Sheberstov S.V., Artem'ev V.A., Otsenka po sputnikovym dannym pokazatelei pogloshcheniya okrashennogo organicheskogo veshchestva i diffuznogo izlucheniya v vodakh Belogo i Karskogo morei (Satellite estimation of the coefficients of CDOM absorption and diffuse attenuation in the White and Kara Seas), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2014, Vol. 11, No. 4, pp. 31–41.
3. Kopelevich O.V., Burenkov V.I., Sheberstov S.V., Prokhorenko O.V., Razrabotka regional'nykh algoritmov atmosferno korrektsii dannykh sputnikovyykh skanerov tsveta (Development of regional algorithms for the atmospheric correction satellite ocean color data), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2009, Vol. 6, No. 1, pp. 400–408.
4. Kopelevich O.V., Burenkov V.I., Vazyulya S.V., Sheberstov S.V., Regional'nye algoritmy rascheta po sputnikovym dannym bioopticheskikh kharakteristik poverkhnostnogo sloya morya (Regional algorithms for calculation of bio-optical characteristics of the surface layer of the sea using satellite data), *Fizicheskie, geologicheskie, i biologicheskie issledovaniya okeanov i morei*. Moscow: Nauchnyi mir, 2010, pp. 115–135.
5. Kopelevich O.V., Sheberstov S.V., Otsenka po sputnikovym dannym mezhgodovykh i sezonnykh izmenenii temperatury poverkhnosti Atlanticheskogo okeana i kontsentratsii khlorofilla v poverkhnostnom sloe v period 1998–2008 (Assessment of inter-annual and seasonal changes in sea surface temperature and chlorophyll concentration in the Atlantic Ocean in 1998–2008 from satellite data), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2010, Vol. 7, No. 1, pp. 238–247.
6. Likhacheva M.V., Kopelevich O.V., Sheberstov S.V., Korrektsiya dannykh sputnikovogo skanera MODIS na solnechnyi blik bez ispol'zovaniya dopolnitel'nykh dannykh o skorosti vetra (Sun glint correction of data from satellite scanner MODIS with no ancillary information on wind speed), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2009, Vol. 6, No. 1, pp. 421–428.
7. Sheberstov S.V., Kopelevich O.V., Lukyanova E.A., Analiz mezhgodovykh trendov temperatury poverkhnosti okeana i kontsentratsii khlorofilla v Atlanticheskoy okeane po sputnikovym dannym (Analysis of inter-annual trends of sea surface temperature and chlorophyll concentration in the Atlantic Ocean from satellite data), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 4, pp. 274–282.
8. Baith K., Lindsay R., and Gary Fu G., McClain Ch. R., Data Analysis System Developed for Ocean Color Satellite Sensors, *Eos*, 2001, Vol. 82, No. 18, 202 p.
9. Burenkov V.I., Kopelevich O.V., Sheberstov S.V., Seasonal and interannual variability of the biooptical characteristics of the Black Sea from satellite data, *Proceedings of VI International Conference: Current Problems in Optics of Natural Waters*, St.-Petersburg: Nauka, 2011, pp. 140–142.
10. Kopelevich O.V., Sheberstov S.V., Burenkov V.I., Vazyulya S.V., Sahling I.V., Bio-optical characteristics of the Russian seas from satellite ocean color data of 1998–2012, *Proceedings of VII International Conference: Current Problems in Optics of Natural Waters*, Saint Petersburg: Nauka, 2013, pp. 168–171.
11. Kopelevich O., Burenkov V., Sheberstov S., Vazyulya S., Kravchishina M., Pautova L., Silkin V., Artemiev V., Grigoriev A., Satellite monitoring of coccolithophore blooms in the Black Sea from ocean color data, *Remote Sensing of Environment*, 2014, Vol. 146, pp. 113–123.
12. Sheberstov S.V., Lukyanova E.A., A System for Acquisition, Processing and Storage of Satellite and Field Bio-optical Data, *Proceedings of IV International Conference: Current Problems in Optics of Natural Waters*. Nizhny Novgorod, September 11–15, 2007, pp. 179–183.
13. Werdell P.J., Bailey S.W., The SeaWiFS Bio-Optical Archive and Storage System (SeaBASS): Current Architecture and Implementation, NASA/TM 2002-211617, *Goddard Space Flight Center*, Greenbelt, Maryland, 2002.