

УДК 504.064.37

Modification of the Empirical Orthogonal Functions Method for Solving the Inverse Task of Retrieving of the CO₂ Total Content from Satellite Data

Mikhail Yu. Kataev*,
Andrey K. Lukyanov and Artur A. Bekerov
*Tomsk State University of Control Systems and Radioelectronics
40 Lenin Str., Tomsk, 634050, Russia*

Received 12.05.2017, received in revised form 17.07.2017, accepted 04.08.2017

This article discusses the use of a modified method of empirical orthogonal functions for solving the inverse problem of recovering the total content of carbon dioxide in the column of atmosphere above a particular territory, based on ground measurement points (TCCON station). Describes the method and results of processing real data satellite GOSAT instrument. Estimated error of the method in comparison with the ground truth data points and product GOSAT L2.

Keywords: inverse problem, empirical orthogonal functions, satellite measurements, atmospheric content of carbon dioxide.

Citation: Kataev M.Yu., Lukyanov A.K., Bekerov A.A. Modification of the empirical orthogonal functions method for solving the inverse task of retrieving of the CO₂ total content from satellite data, J. Sib. Fed. Univ. Eng. technol., 2018, 11(1), 77-85. DOI: 10.17516/1999-494X-0011.

Модификация метода эмпирических ортогональных функций решения обратной задачи восстановления общего содержания углекислого газа по спутниковым данным

М.Ю. Катаев, А.К. Лукьянов, А.А. Бекеров
*Томский государственный университет
систем управления и радиоэлектроники
Россия, 634050, Томск, ул. Ленина, 40*

В статье рассматривается применение модифицированного метода эмпирических ортогональных функций для решения обратной задачи восстановления общего содержания

© Siberian Federal University. All rights reserved

* Corresponding author E-mail address: k_m_y@mail.ru, hyena116@mail.ru

углекислого газа в столбе атмосферы над определенной территорией с опорой на подспутниковые точки измерений (станции TCCON). Приводится описание метода и результаты обработки реальных данных спутникового прибора GOSAT. Оценивается погрешность метода в сравнении с данными подспутниковых точек и продукта GOSAT уровня L2.

Ключевые слова: обратная задача, эмпирические ортогональные функции, спутниковые измерения, содержание углекислого газа.

Введение

В проблеме потепления климата важен вопрос о влиянии атмосферного углекислого газа. Одни ученые выясняют историческую взаимосвязь изменений климата и колебаний концентрации CO₂ в атмосфере по изучению льда. Другие – добиваются количественного объяснения изменений климата от климатообразующих факторов и их вариаций, оценки основных факторов влияния – естественных и антропогенных. Важным является направление моделирования климата, циркуляции атмосферы, переноса вещества. Однако во всех этих направлениях и многих других возникает центральная задача, связанная с верификацией полученных результатов с фактическими значениями содержания углекислого газа в атмосфере Земли [1-3].

Получение такого глобального набора данных с определенной частотой обновления и соответствующей точностью требует применения мониторинговых методов зондирования атмосферы и поверхности Земли из космоса. Многие космические агентства мира создают спутниковые приборы и разрабатывают математические алгоритмы обработки спутниковых данных.

Один из первых приборов, который относится к классу мониторинговых, прибор MODIS (36-канальный спектрорадиометр) [<http://modis.gsfc.nasa.gov>], установленный на спутниках Terra и Aqua, открыл принципиально новый этап глобальных мониторинговых наблюдений «системы Атмосфера+Земля» из космоса. Разработанные тематические продукты на основе данных измерений этим прибором позволяют получать данные об аэрозоле, радиационном балансе, облачности, содержании водяного пара, отражательной способности, температуре подстилающей поверхности и др. в глобальном масштабе на регулярной основе (несколько раз в день).

Однако уникальные пространственно-временные и радиометрические характеристики прибора MODIS не дают возможность получать информацию о газовом составе атмосферы (кроме водяного пара), в частности о содержании углекислого газа и метана в атмосфере. Среди существующих мониторинговых спутниковых приборов, которые позволяют получать информацию о газовом составе атмосферы, можно выделить IASI [<http://smc.cnes.fr/IASI>] и SCIAMACHY [<http://www.sciamachy.org/>]. Хотя эти приборы имеют возможность по своим спектральным характеристикам получать информацию об газовом составе атмосферы, они обладают достаточно большим пятном обзора (например, SCIAMACHY – 60 км поперек трассы и 30 км вдоль трассы в надире). Он существенно сужает рамки применения полученных экспериментальных результатов в научных направлениях, ограничиваясь в большей степени климатическими задачами.

В настоящее время эффективно работает на орбите прибор Фурье-спектрометр среднего разрешения GOSAT [<http://global.jaxa.jp/projects/sat/gosat>], который позволяет получать опера-

тивные данные о содержании в атмосфере климатически важных газов: CO₂ и CH₄. В данной статье рассматривается методика обработки данных GOSAT с целью получения информации о распределении общего содержания CO₂ над территорией. В основе методики лежит метод эмпирических ортогональных функций. Данная статья является продолжением ранее выполненных нами работ по разработке методики для одной пространственно локализованной географической точки [4, 5].

Выбор типа решения обратной задачи

Формально можно выделить две задачи в теории обработки данных оптики атмосферы, это прямая и обратная задачи [6, 7]. Прямая задача подразумевает построение математической модели измерений, в основу которой положены физические представления о процессах взаимодействия излучения с атмосферой и поверхностью Земли. Понятно, что при математической формализации выбранных физических процессов (положенных в основу модели) неизбежно возникают физические приближения, а при численной реализации – математические приближения. Эти приближения в приводят к тому, что полученная физическая модель представляет собой определенное упрощение реальности. Этот факт при решении обратной задачи может вызвать неоднозначность в оценках параметров модели (некорректность). Степень неоднозначности проверяется в сравнении с реальным экспериментом при различных условиях реализации процессов. Также отметим, что число параметров выбранной модели всегда ограничено.

Таким образом, решение прямой задачи – это алгоритм, позволяющий связать измеренные значения с искомыми параметрами физической модели, которую математически можно записать в виде:

$$y_m(x) = F(x, p), \quad (1)$$

где $y_m(x)$ – модельный вектор измеряемых величин на приборной сетке x ; $F(x, p)$ – функция от искомых параметров и приборной сетки x .

В нашем случае величины $y(x)$ содержат измерения отраженного от поверхности солнечного излучения в ближней инфракрасной области спектра.

Обратная задача определяет алгоритм, по которому по данным измерений $y(x)$ и физической модели $F(x, p)$ можно восстановить значения неизвестных параметров p . Отметим, что не все параметры p являются восстанавливаемыми (p_v), определенная их часть должна быть заранее известна (p_a). Формально, опираясь на выражение (1), можно записать выражение для решения обратной задачи в виде

$$p_v = F^{-1}(x, y, p_a), \quad (2)$$

здесь p_v – набор неизвестных восстанавливаемых параметров; $y(x)$ – вектор измеренных значений; p_a – набор известных параметров и $F^{-1}()$ – алгоритм решения обратной задачи, в основе которой лежит функция $F()$.

Заметим, что в общем случае система (2) может и не иметь решений, что может быть связано с тем, что измеряемый вектор имеет погрешности измерений ($y(x) = y_o(x) + \varepsilon(x)$, где $y_o(x)$ – незашумленный (истинный) вектор измерений и $\varepsilon(x)$ – погрешность измерений, а также тем, что функция $F(x, p)$ является лишь приближением к реальности.

Типично все алгоритмы решения обратных задач основаны на сравнении результатов измерения с моделированием функции (1). Типичная формулировка звучит так: следует найти такой набор параметров прямой задачи p , при котором разница измеренных и рассчитанных значений параметров будет минимальна:

$$\hat{p} = \min_p \Delta(y(x), ym(x, p)), \quad (3)$$

где \hat{p} – вектор восстановленных значений параметров (оценка) и $\Delta(y(x), ym(x, p))$ – функция сравнения измеренного и модельного вектора измерений. Для метода наименьших квадратов и ряда других методов Δ представляет собой невязку разности этих двух векторов.

Типичный вариант решения уравнения (3), с учетом априорной информации об изменениях параметров p , представлен следующим выражением:

$$\hat{p} = p_o + (J^T S_\varepsilon^{-1} J + S_p^{-1}) J^T S_\varepsilon^{-1} (y - y_m(p_o)), \quad (4)$$

здесь \hat{p} – вектор неизвестных параметров; p_o – вектор начального приближения параметров; y – вектор измерений; y_m – модельный вектор измерений, рассчитанный для начального значения параметров;

$J = \left. \frac{\partial y_m}{\partial p} \right|_{p_o}$ – матрица Якоби, рассчитанная для приближения p_o ; S_ε – ковариационная матрица шума измерений; S_p – ковариационная матрица априорной информации об искомым параметрах.

Анализ типично применяемого подхода решения обратной задачи (4) приводит к выводу [4, 5], что существует зависимость оценки \hat{p} : 1) от типа выбранной физической модели и приближений; 2) от объема и качества априорной информации; 3) от размерности вектора неизвестных параметров \hat{p} и измеренных значений; 4) от качества определения характеристик погрешности измерений.

Указанные выше и некоторые другие особенности решения, с одной стороны, позволяют получать соответствующего качества решения (оценки), однако эти оценки весьма затратны по времени при больших размерностях вектора измерений; существенно зависят от величины погрешности, могут при решении получать физически неверные результаты (обнуления, отрицательные значения и др.); требуют правильного и тщательного выбора физической модели и приближений.

Учитывая эти моменты, мы для решения обратной задачи применяем метод эмпирических ортогональных функций [8-11], который относится к классу непараметрических алгоритмов. Последнее свойство указывает на факт того, что не существует явной модели, связывающей измерения и искомую оценку неизвестных параметров.

Описание методики ЭОФ и ее модификация

Основная методика получения эмпирических ортогональных функций (ЭОФ) является классической процедурой и состоит из нескольких вычислительных этапов:

Этап 1. Расчет ковариационной матрицы $K=RR^T$, $R=(Y-\langle Y \rangle)$.

Этап 2. Расчет собственных векторов E и собственных значений Λ матрицы K из решения характеристического уравнения $(K-\Lambda I)=0$.

Этап 3. Вычисление эмпирических ортогональных функций $F=E^TR$.

Этап 4. Вычисление коэффициентов линейной регрессии $P=\langle P \rangle + cF$.

На первом этапе формируется матрица вычислений $Y=Y(i,j)$, где $i=1,N$ – число длин волн измеренных значений и $j=1,M$ – число измерений. Ковариационная матрица – квадратная с размером $[N \times N]$.

Модификация метода ЭОФ заключается в модификации матрицы измерений и включении в нее априорной информации. Априорной информацией является зенитный угол Солнца, рельеф, оптические толщии аэрозоля и водяного пара. Для каждого измерения значения априорной информации – различны, что не влияет на вычислительные свойства матрицы в целом, но повышает качество решения.

Результаты применения методики

Для решения климатических задач в 2009 г. японским космическим агентством был запущен спутник IBUKI с Фурье-спектрометром среднего разрешения GOSAT на борту. Каждый сигнал GOSAT представляет собой набор спектральных точек, получаемых Фурье-спектрометром в четырех каналах (около 20 тысяч точек в ближнем ИК (с центрами 0,7, 1,6, 2,1 мкм) и ИК (12 мкм) участках спектра). Три первых канала принимают отраженное солнечное излучение от поверхности, а четвертый (ИК-канал) получает уходящее тепловое излучение системы атмосферы и Земли. Спутник пролетает каждые три дня над одной и той же точкой на поверхности. Нами выбирались данные второго канала (около 8000 точек), из которых для обработки взяты данные (1800 точек), в которых расположены полосы поглощения углекислого газа и метана. Далее все выбранные спектральные точки сигналов за весь временной промежуток времени проходили оценку на вариабельность. Для дальнейшей работы оставляли наиболее изменчивые по величине спектральные точки сигнала, которых оказалось всего 280.

Нами проделана работа по развитию предлагаемой методики эмпирических ортогональных функций при работе не только для одной пространственной точки, но и для большой территории. Для этих целей нами была выбрана территория США, на которой существует сеть станций TCCON, а именно Lamont, Four Coners и Park Falls [<http://www.tcccon.caltech.edu>] (рис.



Рис. 1. Положение станций TCCON и промежуточных точек оценки содержания CO_2

Fig. 1. Geographical location of TCCON stations and intermediate points of total CO_2 content estimation

1). Координаты станций: Lamont (36.604 с.ш., 97.486 з.д., Park Falls (45.945 с.ш., 90.273 з.д.) и Four Corners (36.797 с.ш., 108.48 з.д.). Данные измерений общего содержания CO_2 на этих трех станциях являются базовыми для обучения метода ЭОФ и дальнейшего восстановления общего содержания CO_2 на всей территории США.

Для обучения нами были получены данные измерений GOSAT; например, для станции Lamont было получено 5785 сигналов за все время измерений (2009-2014 гг.), в каждом из которых было 280 спектральных точек.

Был выполнен процесс обучения для всех трех станций одновременно за все годы получения сигналов. Для дальнейшего сравнения качества работы нашего алгоритма для всей территории США приведем результаты сравнения для одной реперной точки – станции Lamont и произвольной другой точки территории США только за 2012 г. После обучения и решения обратной задачи стандартным подходом ЭОФ для тестовой выборки результаты восстановленного общего содержания CO_2 сравнивались со значениями для станции Lamont, которые показаны на рис. 2 и модифицированным методом на рис. 4. Такая же процедура выполнялась для метода ЭОФ для произвольных точек территории США, что показано на рис. 4-8. Модифицированным считается метод, когда матрица измерений дополнялась априорными значениями (зенитный угол Солнца, при котором получен сигнал, общее содержание аэрозольной составляющей атмосферы и водяного пара на основе данных GOSAT уровня L2).

Сравнение результатов работы стандартного и модифицированного алгоритмов показывает, что точность восстановления общего содержания CO_2 предлагаемым модифицирован-

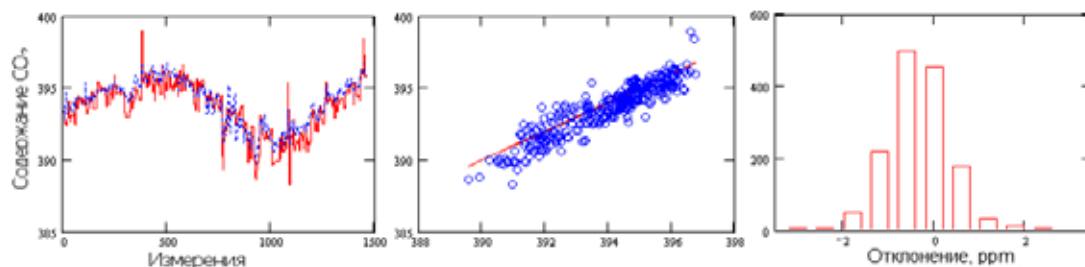


Рис. 2. Рассчитанное общее содержание CO_2 для станции Lamont стандартным методом

Fig. 2. Calculated total CO_2 content for Lamont station by standard method

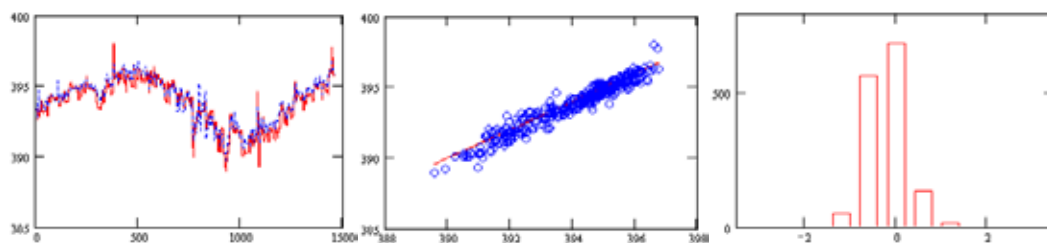


Рис. 3. Рассчитанное общее содержание CO_2 для точки с для станции Lamont модифицированным методом

Fig. 3. Calculated total CO_2 content for point of the Lamont station by a modified method

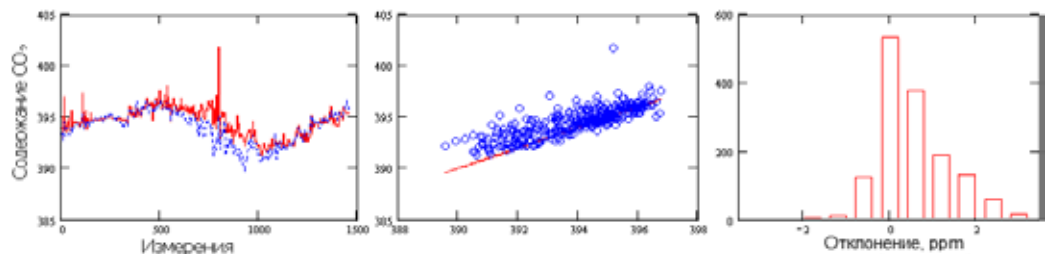


Рис. 4. Рассчитанное общее содержание CO_2 для точки с координатами (41.38 с.ш., 96.84 з.д.) стандартным методом

Fig. 4. Calculated total CO_2 content for a point with coordinates (41.38 N, 96.84 W) by the standard method

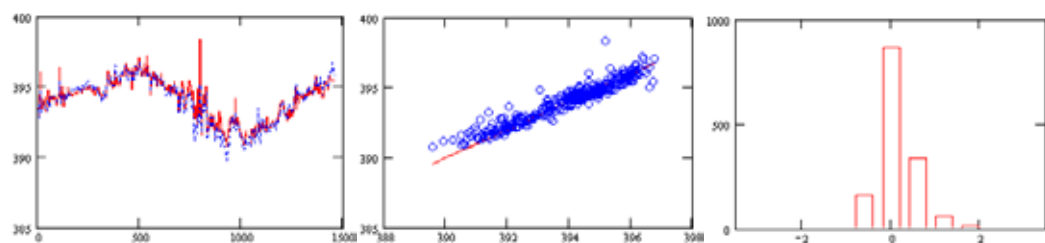


Рис. 5. Рассчитанное общее содержание CO_2 для точки с координатами (41.38 с.ш., 96.84 з.д.) модифицированным методом

Fig. 5. Calculated total CO_2 content for a point with coordinates (41.38 N, 96.84 W) by a modified method

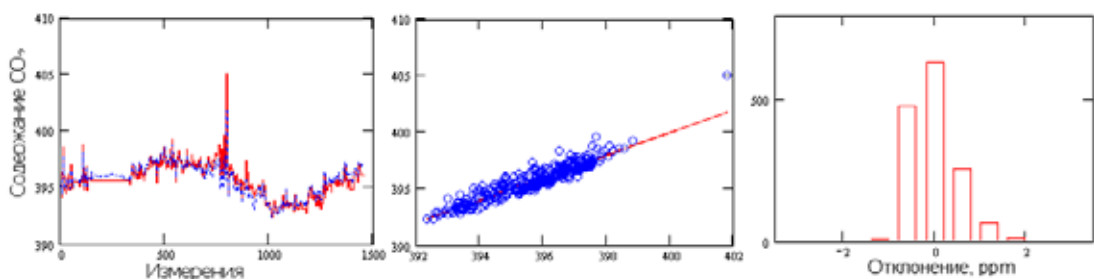


Рис. 6. Рассчитанное общее содержание CO_2 для точки с координатами (38.24 с.ш., 105.48 з.д., рядом со станцией Four Corners) модифицированным методом

Fig. 6. Calculated total CO_2 content for a point with coordinates (38.24 N, 105.48 W, near the Four Corners station) by a modified method

ным методом выше. Предлагаемый вариант обучения на трех точках, находящихся в разных частях исследуемой территории, позволяет с приемлемой точностью восстанавливать общее содержание CO_2 не только в выбранных точках, но и по всей территории. Графики гистограмм отклонений (крайний правый рисунок) показывают, что погрешность не превосходит 1 % (все результаты менее 2 ppm). Однако следует отметить, что при восстановлении возникает некоторое незначительное смещение (средний график рисунков), которое еще предстоит оценить и определить причину его появления в следующих работах. Можно лишь предположить, что на

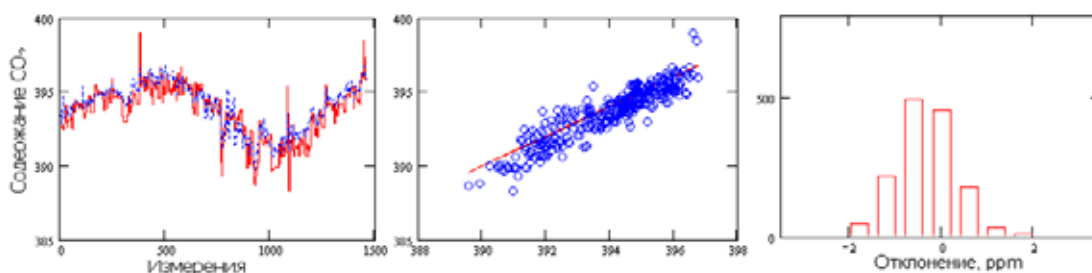


Рис. 7. Рассчитанное общее содержание CO_2 для точки с координатами (42.37 с.ш., 98.87 з.д.) модифицированным методом

Fig. 7. Calculated total CO_2 content for a point with coordinates (42.37 N, 98.87 W) by a modified method

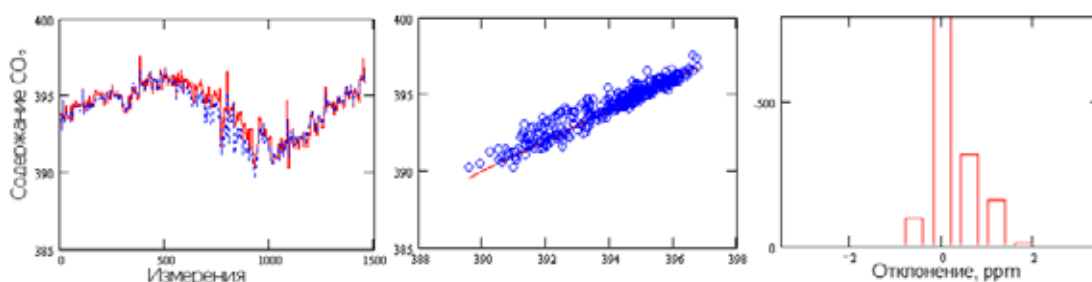


Рис. 8. Рассчитанное общее содержание CO_2 для точки с координатами (46.5 с.ш., 92.26 з.д., рядом со станцией Park Falls) модифицированным методом

Fig. 8. Calculated total CO_2 content for a point with coordinates (46.5 N, 92.26 W, next to the Park Falls station) by a modified method

восстановлении сказывается положение реперной точки, которое определяет тип воздушных масс, а значит, и структуру пространственно-временного поведения общего содержания CO_2 .

Заключение

В статье рассмотрен метод эмпирических ортогональных функций и его модификация при решении обратной задачи восстановления общего содержания углекислого газа по спутниковым данным измерений отраженного солнечного излучения Фурье-спектрометром среднего разрешения GOSAT. Для определения устойчивости и работоспособности предлагаемого метода была выбрана территория США, на которой присутствуют подспутниковые (реперные) станции измерений общего содержания CO_2 похожим на спутниковый Фурье-спектрометром. Результаты обработки реальных спутниковых данных показали, что метод применим не только для одной точки, но и для территории с несколькими территориально разнесенными реперными точками. Одним из возможных приложений результатов данной работы является их экологическая направленность [12].

Список литературы

- [1] Джирард Дж.Е. *Основы химии окружающей среды*. М.: ФИЗМАТЛИТ, 2008, 640 с. [Girard J.E. *Fundamentals of Environmental Chemistry*. М., FIZMATLIT, 2008, 640 p. (in Russian)]

- [2] Семенов С.М. *Парниковые газы и современный климат Земли*. М.: Изд. центр «Метеорология и гидрология», 2004, 175 с. [Semenov S.M. *Greenhouse gases and the modern climate of the Earth*. М., Izd. Center «Meteorology and Hydrology», 2004, 175 p. (in Russian)]
- [3] Кондратьев К.Я. *Глобальный климат*. СПб., Наука, 1992, 359 с. [Kondratiev K.Ya. *Global climate*. SPb., Science, 1992, 359 p. (in Russian)]
- [4] Катаев М.Ю., Катаев С.Г., Максюттов Ш., Андреев А.Г., Базельюк С.А., Лукьянов А.К. Математические алгоритмы обработки и анализа данных спутникового Фурье-спектрометра в ближней ИК-области спектра. *Известия вузов. Физика*, 2012, 3, 84–90 [Kataev, M.Yu., Kataev SG, Maksyutov Sh., Andreev A.G., Bazelyuk S.A., Lukyanov A.K. Mathematical algorithms for data processing and analysis of the satellite Fourier spectrometer in the near-infrared region of the spectrum. *Izvestiya Vuzov. Physics*, 2012, 3, 84-90 (in Russian)]
- [5] Катаев М.Ю., Лукьянов А.К. Восстановления общего содержания углекислого газа методом эмпирических ортогональных функций из спутниковых данных. *Доклады ТУСУР*, 2 (32), 2014, 230-237 [Kataev M.Yu., Lukyanov A.K. Retrieving of the total carbon dioxide content by the method of empirical orthogonal functions from satellite data. *Reports of TUSUR*, 2 (32), 2014, 230-237 (in Russian)]
- [6] Современные проблемы атмосферной оптики: [В 9 т.: Т. 7: Обратные задачи оптики атмосферы, В.Е. Зуев, И.Э. Наац, под общ. ред. В.Е. Зуева]. 1990, 286 p. [Modern problems of atmospheric optics: [In 9: Vol. 7: Inverse problems of atmospheric optics, V.E. Zuev, I.E. Naats Under the general. Ed. V.E. Zueva] 1990, 286 p. (in Russian)]
- [7] Ягола А.Г., Степанова И.Э., Титаренко В.Н. *Обратные задачи и методы их решения*. Приложения к геофизике М., БИНОМ. Лаб. знаний, 2014, 216 с. [Yagola A.G., Stepanova I.E., Titarenko V.N. *Inverse problems and methods for their solution*. Applications to geophysics, М., BINOM. Lab. knowledge, 2014, 216 p. (in Russian)]
- [8] Ахмед Н., Рао К.Р. *Ортогональные преобразования при обработке цифровых сигналов*. М., Связь, 1980, 248 с. [Ahmed N., Rao K.R. *Orthogonal transformations when processing digital signals*. М., Communication, 1980, 248 p. (in Russian)]
- [9] Большаков А.А., Каримов Р.Н. *Методы обработки многомерных данных и временных рядов*. М., Горячая линия – Телеком, 2014, 522 с. [Bolshakov A.A., Karimov R.N. *Methods for processing multidimensional data and time series*. М., Hot line – Telecom, 2014, 522 p. (in Russian)]
- [10] Иберла К. *Факторный анализ*. М.: Статистика, 1980, 398 с. [Iberla K. *Factor analysis*. М., Statistics, 1980, 398 p. (in Russian)]
- [11] Кендалл М., Стьюарт А. *Статистические выводы и связи*. М., Наука, 1973, 899 с. [Kendall M., Stewart A. *Statistical conclusions and connections*. М., Nauka, 1973, 899 p. (in Russian)]
- [12] Катаев М.Ю., Бекеров А.А. Обнаружение экологических изменений природной среды по данным спутниковых измерений. *Оптика атмосферы и океана*, 27 (7), 2014, 652-656. [Kataev M.Yu., Bekеров A.A. Detection of environmental changes in the natural environment from satellite measurements. *Optics of the atmosphere and ocean*, 27 (7), 2014, 652-656 (in Russian)]