

ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНАЯ ТЕХНОЛОГИЯ ОБРАБОТКИ И ИНТЕРПРЕТАЦИИ СПУТНИКОВЫХ ДАННЫХ О ВЫБРОСАХ В АТМОСФЕРУ SO₂

Д.В. Зуев, В.Б. Кашкин, К.В. Симонов
Сибирский федеральный университет

e-mail: dmzvew@yandex.ru

Сибирский федеральный университет

e-mail: rtcvbk@ Rambler.ru

Институт вычислительного моделирования СО РАН

e-mail: simonovkv@icm.krasn.ru

Аннотация

Работа посвящена оценке экологической обстановки в районе расположения «ГМК «Норильский Никель» по выбросам в атмосферу диоксида серы на основе спутниковых методов мониторинга окружающей среды. Анализировались данные, получаемые от прибора Ozone Monitoring Instrument (ОМІ), установленного на искусственный спутник Земли АURA (США). Посредством ОМІ становится возможным выявлять концентрации диоксида серы на различных высотах за период действия программы ОМІ (2004–2011 гг.). Выполнен сравнительный анализ данных наблюдений о концентрации выбросов диоксида серы на различных высотах и выявлены особенности ее пространственного распределения за 2011 год для Норильского промышленного района. Построена карта концентрации диоксида серы для участка ограниченного координатами 66,8–69,7 градусов северной широты и 86,8–89,6 градусов восточной долготы. Показаны места с повышенной концентрацией и определены пики выбросов с привязкой по координатам.

Введение

Одной из наиболее актуальных проблем является проблема охраны природной среды. Научно-технический прогресс привел не только к появлению новых эффективных технологий, но и к увеличению техногенного давления на природу, а также к расходованию до опасной черты невозобновляемых природных ресурсов. Наиболее интенсивно загрязняются городские и пригородные территории, промышленные зоны. Особо остро проблема экологической безопасности стоит перед рядом российских регионов и, в частности, перед Норильским промышленным районом, который занимает первое место по степени загрязнения территории [1-2].

Особенностью исследуемого района является легкая экологическая уязвимость природной среды [2]. Экологическая обстановка в Норильском промышленном районе формируется, главным образом, в результате производственной деятельности предприятий Норильского горно-металлургического комбината, в настоящее время – Заполярного филиала ОАО «Горно-металлургическая компания «Норильский Никель», мощного многоотраслевого производственно-хозяйственного комплекса. Основными источниками поступления вредных (загрязняющих) веществ в природную среду являются металлургические переделы, а основным загрязняющим веществом – диоксид серы (SO₂). В

периоды безветрия (периоды, неблагоприятных метеоусловий – НМУ) в Норильске образуется высокий уровень загрязнения атмосферы диоксидом серы.

Практически ежедневно (порядка 350 дней в году), в течение 50-60 % от календарного времени, в различных районах города фиксируется загрязнение атмосферы диоксидом серы с превышением предельно допустимой нормы, в том числе порядка 190 дней – с пятикратным и более превышением, из них 90 дней с десятикратным и более превышением. Основной «вклад» в загрязнение атмосферы диоксидом серы в периоды НМУ вносит медный завод – 55 % от общей продолжительности загрязнения, а аглофабрика и никелевый завод – порядка 40 %. Уровни загрязнения также значительны при работе медного завода (максимальные концентрации SO₂ в периоды НМУ достигает 40-60 ПДК). Продолжительность отдельного случая загазованности различных районов города составляют в среднем 2-3 часа, а максимально – до суток. Работа промышленного комплекса связана со значительными техногенными воздействиями на естественную природную обстановку. Это приводит также к нарушению поверхностных покровов, водно-теплового режима сезонно-талого слоя и интенсивному развитию неблагоприятных криогенных физико-геологических процессов, протекающих в естественной обстановке обычно с гораздо меньшей скоростью.

Учитывая значимость указанной проблемы, требуется комплексный подход к проведению наблюдений, выдаваемых оценок и прогнозов в данной области. Известно [3-4], что диоксид серы поступает в воздух в результате сжигания топлива и плавки руд, содержащих серу. Основные источники загрязнения атмосферы SO₂: энергетические установки, предприятия цветной металлургии и сернокислотное производство. Менее значительны выбросы предприятий черной металлургии и машиностроения, угольной, нефтеперерабатывающей промышленности, производства суперфосфата, транспорта.

Как показывают данные космического мониторинга, выбросы SO₂ загрязняют воздух на значительное расстояние от источника (на тысячу и более километров). Диоксид серы считается одной из основных действующих составных частей «токсичных туманов» и одним из активных компонентов формирования смога. Сернистый ангидрид может вызывать общее отравление организма, проявляющееся в изменении состава крови, поражении органов дыхания, повышении восприимчивости к инфекционным заболеваниям, нарушении обмена веществ, повышении артериального давления у детей, ларингит, конъюнктивит, ринит, бронхопневмония, аллергические реакции, острые заболевания верхних дыхательных путей и системы кровообращения.

При кратковременном воздействии – раздражение слизистой оболочки глаз, слезотечение, затруднение дыхания, тошнота, рвота, головные боли. Повышение уровней общей заболеваемости смертности. Повышенная утомляемость, ослабление мышечной силы, снижение памяти. Замедление восприятия, ослабление функциональной способности сердца, изменение бактерицидности кожи. Диоксид серы может нарушать углеводный и белковый обмен, способствует образованию метгемоглобина, снижению иммуннозащитных свойств организма.

Взаимодействуя с атмосферной влагой, оксиды серы образуют кислотные осадки ($Ph \ll 4,5$), которые оказывают как прямое повреждающее действие на биоту, так и косвенное, закисляя почвы и водоемы. Установлено что при закислении почвы снижаются доступность для растений питательных элементов (Ca, Mg, Mn) и плодородие почвы.

Защеление уменьшает скорость разложения органических остатков, поскольку для жизнедеятельности большинства бактерий и грибов необходима нейтральная среда, снижается продуктивность азотфиксирующих бактерий (при $Ph < 5,0$ азотобактерий погибает), что приводит к ограничению поступления связанного азота в растения и торможению их роста.

Изменение структуры почвы (снижение грануляции, слияние частиц, уплотнение почвы и резкое уменьшение воздухопроницаемости) негативно сказывается на функционировании корневой системы растений, что наносит колоссальный вред лесным массивам. В кислой почве увеличивается подвижность ионов тяжелых металлов, которые накапливаются в растениях. Некоторые из них, например ионы железа и марганца блокируют поступление фосфора в растения.

1. Спутниковый мониторинг выбросов SO_2 в атмосферу

Дистанционные методы позволяют прогнозировать изменение газового состава атмосферы под воздействием выбросов (или постоянного поступления) загрязняющих веществ антропогенного и природного происхождения, а также рационально и оперативно реагировать на эти изменения. Глобальный мониторинг концентраций SO_2 осуществляется на основе измерений в ультрафиолетовом диапазоне (UV) приборами, установленными на искусственных спутниках Земли. Алгоритм обработки данных даёт информацию о повышенных концентрациях диоксида серы на четырёх высотах над уровнем моря: 0,9 км, 5 км, 10 км и 15 км.

Для мониторинга содержания и концентраций диоксида серы в атмосферном воздухе, на околоземной орбите находится два прибора – SCIAMACHY (на борту спутника ENVISAT) и OMI (на борту спутника AURA). Инструмент OMI является совместной разработкой Голландского аэрокосмического агентства, Финского метеорологического института и NASA [2]. Этот прибор в 2004 году сменил на околоземной орбите своего предшественника – TOMS (Total Ozone Mapping Spectrometer). OMI осуществляет ежедневный глобальный мониторинг.

Метод основан на измерении спектральных характеристик солнечного света рассеянного в атмосфере и отражённого от поверхности Земли. Сравнение спектральной интенсивности исходного и рассеянного в атмосфере «назад» излучения в ультрафиолетовом диапазоне даёт информацию о распределении и концентрации примесей диоксида серы, так как этот газ поглощает часть приходящего солнечного излучения [3-4]. В качестве единиц измерения концентрации SO_2 используются единицы Добсона (е.Д или DU). Одна е.Д. равна 0,01 мм толщины осажденного слоя диоксида серы при 0° С и атмосферном давлении 1013 ГПа. Типичное фоновое значение концентрации SO_2 в атмосфере составляет менее 1 е.Д. Погрешность оценки содержания диоксида серы прибором OMI равна ~ 1 е.Д.

Инструмент OMI является результатом дальнейшего развития разработанных в европейском космическом агентстве (ЕСА) инструментов GOME и SCIAMACHY, на которых была отработана технология измерений с высоким спектральным разрешением, ведущихся в широкой полосе спектра, включающей одновременно видимый, УФ- и ИК-диапазоны, что, например, позволяло получать данные о наличии сразу нескольких незначительных газовых примесей по данным одного измерения. Предшественником OMI является его американский аналог TOMS, разработанный NASA. Он использовал 8

спектральных диапазонов, позволяющих замерять количество газовых примесей в атмосферном столбе. ОМІ проводит ежедневные глобальные измерения в широкой полосе спектра (УФ- и видимый диапазоны) с очень высоким пространственным разрешением (13 км × 24 км).

Прибор ОМІ представляет собой направленный в надир спектрограф. Инструмент ОМІ способен осуществлять мониторинг состояния атмосферы в диапазоне длин волн 270-500 нм со спектральным разрешением 0,5 нм. Ширина снимаемой полосы поверхности, при угле обзора камеры в 114°, составляет порядка 2600 км, что позволяет вести измерения в глобальных масштабах. В нормальном режиме работы ОМІ размер пикселя (при съемке в направлении надира) составляет 13 км вдоль и 24 км поперек полосы съемки соответственно. В режиме увеличения пространственное разрешение можно увеличить, уменьшив размер пикселя до 13 км вдоль и 24 км поперек полосы съемки соответственно. Уменьшение размера пикселя позволяет вести наблюдения сквозь просветы в облачном покрове, что является важным при наблюдениях за состоянием тропосферы.

Для обработки данных, получаемых инструментом ОМІ, используется алгоритм, разработанный группой ученых из NASA (Kai Yang, Nickolay A. Krotkov, Arlin J. Krueger, Simon A. Carn, Pawan K. Bhartia, and Pieternel F. Levelt). Для расчета SO₂ используются диапазоны, выбранные для алгоритма измерения общего количества озона (ОМТОЗ). ОМТОЗ включает в себя 6 диапазонов, имевшихся ранее и у Earth-Probe TOMS, 4 диапазона выбранных для BRD-алгоритма (Band Residual Difference, метод разности диапазонов), и 2 дополнительных спектральных диапазона, в которых отсутствует поглощение, для повышения качества результатов ОМТОЗ (для калибровки). Этот набор каналов, включающих различные области спектра поглощения SO₂ и озона, содержит в себе обе (сильную и слабую) области поглощения. Данных, полученных в этих каналах достаточно для расчета количества SO₂.

Алгоритм обработки данных наблюдений. Результаты измерений отраженного ультрафиолетового солнечного излучения I_m (нормализованного потока падающего солнечного излучения) связаны с геофизическими характеристиками через модель переноса излучения, с помощью которой определяется поток излучения в верхних слоях атмосферы I . Спектр поглощения диоксида серы в УФ-диапазоне показан на рисунке 1. Эти характеристики определяются профилями содержания в атмосфере озона, (Ω), SO₂ (Ξ), отражательной способности поверхности (R), и других параметров. Сделав выборку из измеренных и смоделированных значений потока излучения в заданном канале, можно вывести их соотношение:

$$\begin{aligned} \log_{10} I_m &= \log_{10} I(\Omega, \Xi, R) + \varepsilon \\ N_m &= N(\Omega, \Xi, R) + \varepsilon_T, \end{aligned} \quad (1)$$

где $N = -100 \log_{10} I$, и $\varepsilon_T = -100 * \varepsilon$ обозначает абсолютную ошибку, слагаемую ошибками измерений и самой модели для канала, N – безразмерная величина.

В целом, поток излучения в верхних слоях атмосферы зависит от вертикальных профилей абсорбентов, которыми в нашем случае являются озон и SO₂. Подробная детализация профиля нецелесообразна ввиду значительных погрешностей измерений на длинах волн более 310 нм, ограниченности информации о вертикальном распределении компонентов профиля.

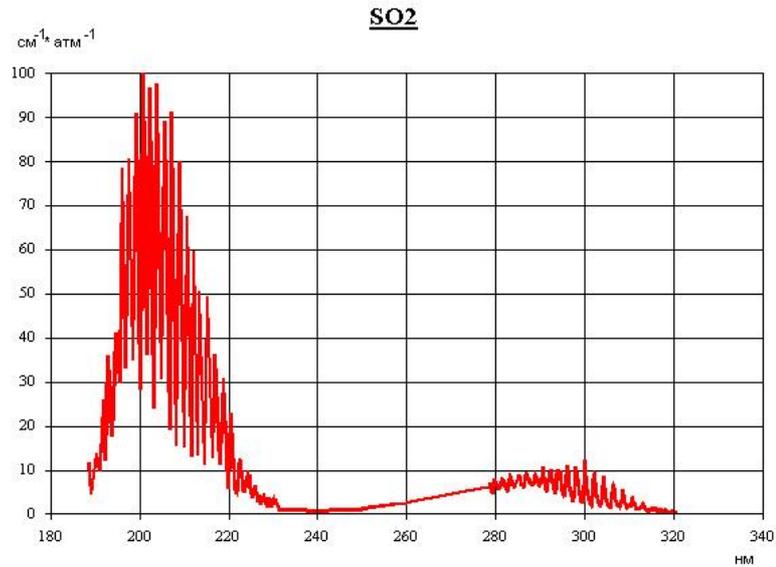


Рис. 1. Спектр поглощения SO₂ в ультрафиолетовом диапазоне.

Большинство существующих алгоритмов анализа первичных данных позволяют выразить поток излучения в верхних слоях атмосферы в зависимости от общего количества абсорбента, как в формуле (1), в которой эти профили стали частью параметров модели, и, соответственно не требуют измерения. Эти алгоритмы вносят ограничения в профили абсорбентов. В рассматриваемом алгоритме данные ограничения, взятый априори озон и заданный SO₂, служат для выяснения зависимости между профилями и размерами их столбов данных веществ. Эти и другие параметры модели, явно выраженные в (1), необходимы в качестве входных данных для измерителей потока излучения в верхних слоях атмосферы и рассмотрены далее. В формуле (1) сделано допущение, что скалярная величина R, которая может зависеть от длины волны, используется для описания отражения от различных поверхностей при возможности их ламбертовской аппроксимации.

С такими допущениями, вывод геофизических параметров: Ω , Ξ , R – может быть определен путем сравнения до тех пор, пока значения потока излучения в верхних слоях атмосферы, полученные из модели не начнут совпадать с данными измерений на заданном канале. С учетом наличия ошибок в измерениях и модели, алгоритм можно описать как минимизацию суммы квадратов остатков на выбранных каналах, остаток – разница между измеренным N и рассчитанным, исходя из модели.

Проблема минимизации в дальнейшем решается путем линеаризации (1). Для этого базисная точка, описываемая исходным вектором $\{\Omega_0, \Xi_0, R_0, c_1=0, c_2=0\}$, выбирается в качестве первоначального решения для уравнения. Уравнение (1) может быть записано следующим образом:

$$\begin{aligned}
 N_m - N_0 = & \Delta\Omega \left. \frac{\partial N}{\partial \Omega} \right|_{\Omega=\Omega_0} + \Delta\Xi \left. \frac{\partial N}{\partial \Xi} \right|_{\Xi=\Xi_0} \\
 & + \left(\Delta R + \sum_{i=1}^n c_i (\lambda - \lambda_0)^i \right) \left. \frac{\partial N}{\partial R} \right|_{R=R_0} + \varepsilon_T,
 \end{aligned} \tag{2}$$

где $N_0 = N(\Omega_0, \Xi_0, R_0)$, $\Delta\Omega = \Omega - \Omega_0$, $\Delta\Xi = \Xi - \Xi_0$, $\Delta R = R - R_0$.

Члены высшего порядка в (2) сведены в ε_T . R_0 здесь представлено как независимая от длины волны, и полином низшего порядка (квадратный в нашем случае, т.е. $n=2$ а c_1 и c_2 – коэффициенты) используется для выяснения количественной зависимости длины волны от отражательной способности поверхности. Минимизация теперь может быть получена путем метода наименьших квадратов остатков ($N_m - N_0$) для набора измерений в разных каналах весовыми функциями, модулирующими полином длины волны λ , с λ_0 являющейся исходной длиной волны, которую выбирают когда требуется вывести R_0 .

Алгоритм, описанный выше, может использоваться для всех гиперспектральных измерений в каком либо «окне», или только в нескольких отдельных диапазонах. Главным преимуществом большого количества измерений в малом поддиапазоне, является то, что таким образом можно снизить влияние систематических и случайных ошибок в измерениях, повышая при этом точность полученных результатов.

Например, спектральная чувствительность прибора и его спектральное разрешение при гиперспектральных измерениях могут быть улучшены путем эталонирования (сравнения с эталоном) и калибровки. В примерах для данного алгоритма показано, что расчеты с использованием гиперспектральных измерений GOME дают большую точность, чем использование отдельных диапазонов (как в TOMS) при тех же условиях.

Модель OMTO3, названная TOMRAD, предназначена для расчета значений потока излучения в верхних слоях атмосферы на длинах волн, соответствующих весовым функциям. Этот вектор функции переноса использует модель упругого молекулярного рассеяния, модель поглощения газовыми компонентами (включая все последовательные этапы рассеяния) и модель физико-спектральной коррекции. Модели позволяют достигнуть большей надёжности при обработке измерений отражённого УФ излучения для зенитных углов и при больших углах, до 88° .

Упрощающее допущение сделано относительно непрозрачности поверхности и Ламбертовского характера отражения (Lambert-equivalent reflectivity, LER). Кроме того, этот метод LER в сочетании с независимой аппроксимацией пикселей (MLER), учитывает эффекты тонких или разорванных облаков как это обычно и делается в расчете незначительных газовых примесей. В MLER-методе, при измерении потока излучения в верхних слоях атмосферы, значение пикселя, частично захватывающего облако, принимается как взвешенная сумма плотностей излучения, взятых у чистого и закрытого облаком субпикселей с фиксированной отражательной способностью (обычно 0.15 или 0.8).

Значения эффективного давления облака необходимые для алгоритма MLER берутся из данных спутниковой инфракрасной климатологической съемки. Аэрозоли не выражаются явно в TOMRAD, но их воздействие на поток излучения в верхних слоях атмосферы частично (за исключением возможного поглощения на аэрозолях) рассчитываются путем уравнивания и принятия эффективной отражательной способности как зависимости от длины волны с применением полинома второго порядка. В этой модели профиль озона принимается детерминированным, полученным на основе измерений, произведённых в определённый день года и на определённой территории.

Данные о климатологии температурного режима, полученные TOMS-V8 необходимы для расчета распределенных по времени и широтам вертикальных профилей температуры с тех пор, как сечения поглощения озона и SO_2 , использующиеся в модели стали зависимыми от температуры. Использование озона и temperature climatology уточняет расчеты потока

излучения в верхних слоях атмосферы за счет внесения в них вариаций зависящих от сезонных и широтных колебаний.

В УФ диапазоне вращательный эффект Рамана (RRS), который является нелинейным и включает в себя примерно только 4% от общего молекулярного рассеяния, изменяет длину волны рассеянного излучения, приводя к сглаживанию рельефных областей спектра, таких как солнечные линии Фраунгофера («кольца») и атмосферные полосы поглощения (эффект теллура). Этот RRS-эффект может оказать сильное влияние на расчет незначительных газовых примесей, если они не были должным образом подсчитаны.

В данной перспективной модели RRS-эффекты включены путем корректирования измеренных TOMRAD плотностей излучения с filling-in факторами, вычисленными с использованием скалярной LIDORT-RSS программы расчета переноса излучения. Факторы заполнения (filling-in факторы) определяются как отношение коэффициентов нелинейно (комбинационно) и линейно рассеянного потоков излучения, и рассчитываются в сходных с TOMRAD атмосферных, поверхностных и геометрических условиях, что гарантирует правильный учет зависимости от RRS-эффектов полосы поглощения, отражения поверхности, угла обзора и геометрии освещенности. Надо заметить, что, т.к. RRS слабо поляризуются, filling-in факторы могут быть точно подсчитаны без включения параметров поляризации излучения.

Обработка данных мониторинга. На рис. 2 представлен пример визуализации данных о концентрации диоксида серы по измерениям OMI с сайта NASA за 2 мая 2011 г. [5]. Можно видеть, что уровень концентрации в районе г. Норильска существенно повышен. При этом максимальный уровень зарегистрирован в точке с координатами 69,26° северной широты, 86,98° восточной долготы и составляет 5,32 *е.Д.*, что соответствует более чем пятикратному превышению фонового уровня. Двукратное превышение концентрации соответствует участку общей площадью около 7000 *км²*, ограниченной пределами по широте 68,5°-69,6° северной широты, по долготе 84°-88° восточной долготы.

В течение суток 02.05.2011 г. выбросы диоксида серы составили около 1000 *тонн*, площадь поражения при этом 16000 *км²*, что сравнимо с выбросами при извержении вулкана Безымянный 14 октября 2007 г., длившегося двое суток, когда выбросы SO₂ составили 3000 *тонн*. Анализируя данные о выбросах диоксида серы за июль 2011 г. (рис. 3), можно сделать вывод, что в течение месяца концентрация SO₂ не превышала фонового значения, кроме 27.07.2011 г., когда наблюдается двукратное превышение выбросов SO₂ на площади в 1000 *км²*, максимальное значение концентрации диоксида серы в 2,37 *е.Д.* регистрировалось в точке с координатами: 69,32° северной широты и 86,15° восточной долготы. В течение этого дня общее количество выбросов диоксида серы составило 260 *тонн*, площадь поражения при этом составила 6633 *км²*.

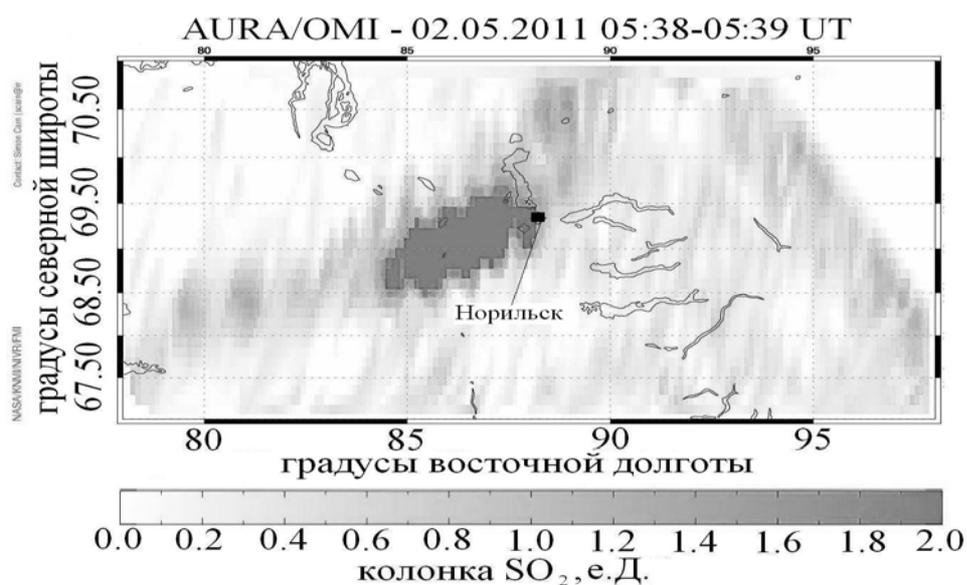


Рис. 2. Визуализация данных SO_2 за 02.05.2011, общая масса SO_2 1000 тонн, максимум SO_2 5,32 е.Д. на $69,26^\circ$ северной широты, $86,98^\circ$ восточной долготы.

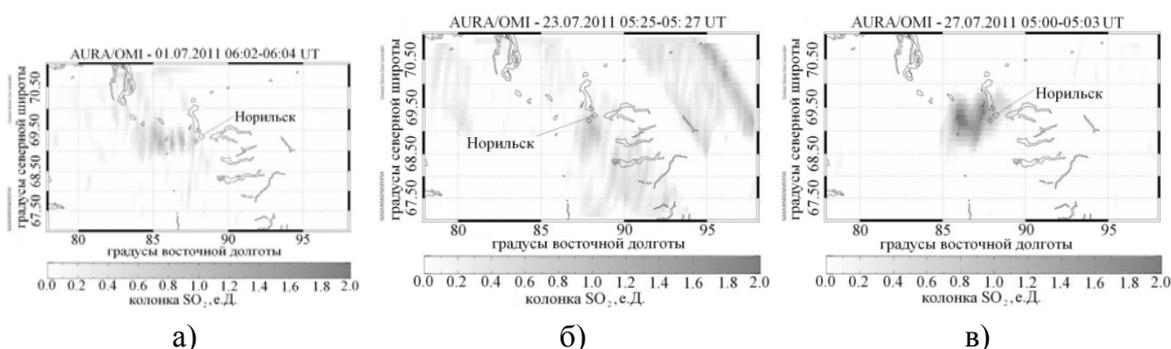


Рис. 3. Визуализация данных SO_2 за 01.07.2011, 23.07.2011 и 27.07.2011:
 а) – 01.07.2011, общая масса SO_2 15 тонн, максимум SO_2 1,39 е.Д. на $69,26^\circ$ северной широты, $86,27^\circ$ восточной долготы.
 б) – 23.07.2011, общая масса SO_2 5 тонн, максимум SO_2 1,50 е.Д. на $70,28^\circ$ северной широты, $96,14^\circ$ восточной долготы.
 в) – 27.07.2011, общая масса SO_2 261 тонн, максимум SO_2 2,37 е.Д. на $69,32^\circ$ северной широты, $86,15^\circ$ восточной долготы.

2. Методики анализа данных и результаты интерпретации

Алгоритмы реализации метода динамических ядер. Рассмотрим метод динамических ядер [12], который используется для обработки наблюдений космического мониторинга Земли [13-18]. Этот метод позволяет автоматически разделять точки из базы данных по классам, отличающимся значением концентрации и координатами, определяя тем самым участки с повышенным уровнем концентрации диоксида серы. Входными данными является набор n -мерных точек. Общее количество точек обозначим через m . Для каждой точки нужно определить её принадлежность одному из k классов, где $k \leq m$. Для этого выполняется следующая последовательность действий (алгоритм-1):

1. Выбирается k точек, называемых начальными ядрами: a^1, \dots, a^k , точка a^i будет ядром класса с номером i . Результатом всего процесса будет вектор \mathbf{C} размерности m , где компоненты вектора c_j принадлежит множеству $\{1, 2, \dots, k\}$. Два следующих пункта представляют собой итерационный процесс.

2. Для каждой точки b^j из входного списка вычисляется расстояние от этой точки до каждого из ядер. Полагается, что точка b^j относится к тому классу, расстояние до ядра которого наименьшее. Соответствующий номер класса заносится в c_j . В качестве расстояния используется евклидово расстояние:

$$D_{ij} = \sqrt{(b_1^j - a_1^i)^2 + (b_2^j - a_2^i)^2 + \dots + (b_n^j - a_n^i)^2} \quad (3)$$

Отметим, что для метода динамических ядер важно значение расстояний между разными точками. Если полученный на данном этапе вектор \mathbf{C} не отличается от вектора, полученного на предыдущем этапе, то процесс останавливается, вектор \mathbf{C} объявляется результатом работы алгоритма на ядрах a^1, \dots, a^k , в противном случае процесс переходит к пункту 3.

3. Теперь для каждого класса выбираем все точки, принадлежащие этому классу, и вычисляем среднюю для класса точку, которая станет новым ядром этого класса. Например, для s точек

$$p^1 = (p_1^1, p_2^1, \dots, p_n^1), p^2 = (p_1^2, p_2^2, \dots, p_n^2), \dots, p^s = (p_1^s, p_2^s, \dots, p_n^s) \quad (4)$$

из класса с номером i средней точкой будет точка

$$p_i^s = \left(\frac{\sum_{j=1}^s p_1^j}{s}, \frac{\sum_{j=1}^s p_2^j}{s}, \dots, \frac{\sum_{j=1}^s p_n^j}{s} \right) \quad (5)$$

С полученными ядрами процесс переходит к пункту 2.

Описанный итерационный процесс представляет собой один шаг алгоритма динамических ядер. Задача программной реализации заключается в том, чтобы перебрать все возможные значения различных начальных ядер без повторений.

Пусть $\mathbf{W} = (w_1, \dots, w_k)$ – вектор индексов начальных ядер. На первой итерации в качестве начальных ядер используются первые k точек, т.е. вектор имеет вид $\mathbf{W}=(1, 2, \dots, k)$. *Инкрементом* этого вектора назовем его преобразование по следующему алгоритму (алгоритм-2):

1. Начиная с $i=k$, выполняются $w_i = w_i + 1$. Если w_i стало равным $m - k + i$, то уменьшаем i на единицу и снова выполняем первый шаг, в противном случае переходим к следующему шагу.

2. Для полученного I выполняется следующая операция: $w_{i+1} = w_i + 1$; $i = i + 1$, пока i не станет равным k . Последним возможным вектором при переборе таким способом будет вектор $\mathbf{W}=(m-k+1, m-k+2, \dots, m)$.

При большом объеме данных чаще всего невозможно перебрать все такие векторы даже с использованием параллельной вычислительной машины, поэтому процесс прерывают после перебора заданного количества индексов начальных ядер или же по команде пользователя. Практика показывает, что для различных наборов начальных ядер получаются одинаковые (с учётом переименования) результаты. Поэтому разработанный программный комплекс выдаёт ответ в виде пары $\langle N, \mathbf{C} \rangle$, где \mathbf{C} – результирующий вектор, N – количество таких векторов. Кроме того, очень полезно сохранять индексы начальных ядер для каждого отдельного варианта разбиения множества. За окончательный результат можно принять

вектор, значение N для которого наиболее велико или другой вектор по выбору пользователя [12].

Теперь допустим, что идёт разбиение некоторых данных на два класса, при котором один шаг метода применяется сначала с исходными ядрами ($a^1=1$, $a^2=2$), а потом с ядрами ($a^1=2$, $a^2=1$), причем результаты этих разбиений окажутся одинаковыми с точностью до наименования классов. Практика показывает, что такие результаты часто получаются при любом k даже при использовании алгоритма-2, поэтому, полезно иметь алгоритм, приводящий номера классов в результирующих векторах к универсальному виду. Предлагается использовать алгоритм-3 [12]:

1. Введем вектор \mathbf{P} размерности k . Это будет вектор универсальных имён.

2. Пробегаем по вектору \mathbf{C} . На некотором шаге встретится число c_i , которое ранее не встречалось. Пусть это r -ое из встреченных нами различных чисел. Например, число c_1 встретится нам самым первым, т.е. r для него будет равно 1.

3. Присвоим компоненту P_{ci} вектора \mathbf{P} номер r , вектор \mathbf{c} нужно просматривать по порядку до тех пор, пока не будет заполнен весь вектор \mathbf{P} , т.е. пока мы не встретим все k различных чисел. Если какой-либо класс пуст, то вектор будет просмотрен до конца.

4. Получив вектор \mathbf{P} , пройдём по вектору \mathbf{C} от начала до конца, выполняя $c_i=P_{ci}$.

Представленный алгоритм метода динамических ядер в виде итерационной процедуры эффективной кластеризации данных наблюдений, в котором напрямую минимизируется суммарная мера близости, является успешной альтернативой методам обучения сетей Кохонена. Отметим, что в методе динамических ядер важным является вопрос о хорошем выборе начального приближения (первоначальных значений ядер).

Приведенные алгоритмы реализованы на кластерной вычислительной установке с распределенной памятью, при этом каждый узел имеет несколько вычислительных ядер, память для которых является общей. На большинстве современных кластеров установлена ОС *Linux* и коммуникационная библиотека *MPI*, которая будет использоваться для взаимодействия между узлами [19]. В каждом узле будет запущен один процесс, который внутри себя породит потоки, использующие один общий буфер входных данных. Для управления потоками внутри процесса используется интерфейс *pthread*, доступный практически на любой *UNIX*-подобной платформе.

Один шаг описанного выше алгоритма на различных начальных ядрах может выполняться за различное количество итераций. Более того, после выполнения шага результирующий вектор нужно сравнить с множеством уже полученных векторов, чтобы в случае повторения лишь увеличить количество векторов такого вида, а не сохранять дубли уже имеющихся векторов. Назовём этот процесс *сборкой*. Сборка может завершиться раньше или позже, в зависимости от того, будет ли найдено повторение и насколько быстро оно будет найдено. По этим двум причинам нецелесообразно заранее вычислять наборы индексов начальных ядер для раздачи их исполняющим потокам, т.к. часть потоков сделают свою работу раньше других и будут простаивать.

Запуск программы осуществляется командой `./cluster-аргументы`. С помощью команды `cat` создаем файл со столбцом кластеризации. Следующий шаг: совмещаем исходный файл с файлом кластеризации, используя полученный файл, строим искомый график, где представлено решение поставленной задачи.

Описание базы данных. Используемая в работе база данных космического мониторинга OMI (NASA) по выбросам в атмосферу диоксида серы в районе г. Норильска, имеет вид текстовой таблицы за период с 2004 по 2011 гг., всего 156167 строк [5]. Ниже приведен фрагмент таблицы (табл. 1).

Табл. 1 Фрагмент базы данных OMI.

Дата	с. (UTC)	широта	долгота	SO ₂ _0,9км	SO ₂ _5км	SO ₂ _10км	SO ₂ _15км
01102004	10468	68,98	88,93	4,281	0,813	0,18	0,073
01102004	10470	69,01	88,61	5,12	7,724	1,804	0,767
01102004	10472	69,05	88,28	2,674	6,496	1,377	0,57
01102004	10474	69,08	87,96	1,801	6,355	1,253	0,516
01102004	10476	69,11	87,63	1,007	5,276	1,133	0,462
01102004	10478	69,14	87,3	3,513	6,849	1,453	0,593
01102004	16269	69,23	89,28	4,425	1,732	0,646	0,41
01102004	16271	69,3	89,02	1,035	1,14	0,428	0,275

При выполнении расчетов использовались столбцы, где приведены широта, долгота, SO₂_0,9 км – концентрация SO₂ на высоте 0,9 км в е.Д., SO₂_5 км – концентрация SO₂ на высоте 5 км в е.Д., SO₂_10 км – концентрация SO₂ на высоте 10 км в е.Д., SO₂_15 км – концентрация SO₂ на высоте 15 км в е.Д.

Описание результатов. Выполнен предварительный анализ данных наблюдений OMI (рис. 4–8), за 2011 г., построены области-поверхности, которые характеризуют особенности данных космического мониторинга. На рис. 4 представлен результат обработки выборки данных наблюдений за весь 2011 г., высота столба SO₂ от 0,9 до 5 км. На рисунке видно, что пики выбросов в большей степени распределены с востока на запад. Максимальная концентрация выбросов SO₂ соответствует областям с координатами 69,16° северной широты и 87,2° восточной долготы, 69,2° северной широты и 88,6° восточной долготы, 69,29° северной широты и 88,9 восточной долготы. Среднее значение концентрации SO₂ в исследуемой области 69°- 70° северной широты и 87°- 90° восточной долготы соответствует уровню 7 е.Д. По мере удаления от г. Норильска постепенно снижение концентрации SO₂.



Рис. 4. Результат кластеризации выборки данных наблюдений за 2011 г., высота столба SO₂ от 0 до 0,9 км.

Сравнивая концентрацию SO_2 на высотах 0,9 и 5 км (рис. 4 и 5), можно видеть существенное уменьшение концентрации на высоте 5 км. Просматриваются пики концентраций, которые соответствуют областям с координатами $69,2^\circ$ северной широты и $89,1^\circ$ восточной долготы, 69° северной широты и $88,8^\circ$ восточной долготы. Заметно уменьшение концентрации с запада на восток. Среднее значение концентрации выбросов диоксида серы соответствует уровню в 4 е.Д.

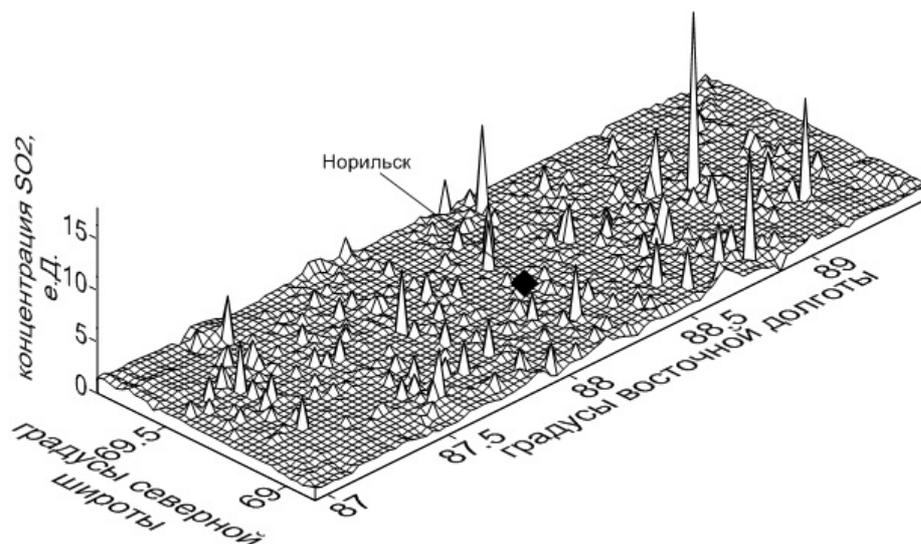


Рис. 5. Результат кластеризации выборки данных наблюдений за 2011 г., высота столба SO_2 от 0,9 до 5 км.

На рис. 6 представлен результат обработки данных SO_2 , на высотах 5-10 км. Среднее значение концентрации соответствует 4 е.Д. Пик выбросов в 15 е.Д. соответствует области с координатами $69,5^\circ$ северной широты и 89° восточной долготы.

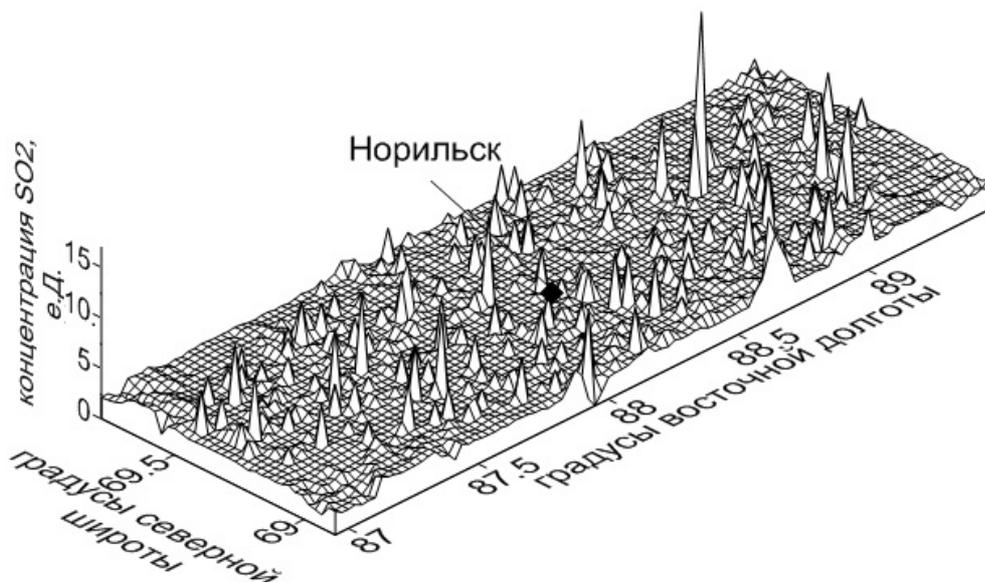


Рис. 6. Результат кластеризации выборки данных наблюдений за 2011 г, высота столба SO_2 от 5 до 10 км.

Анализируя выбросы диоксида серы на высотах 10–15 км (рис. 7), можно увидеть, что среднее значение концентрации соответствует 3 е.Д., что является минимальным значением среди всех ранее анализируемых высот.

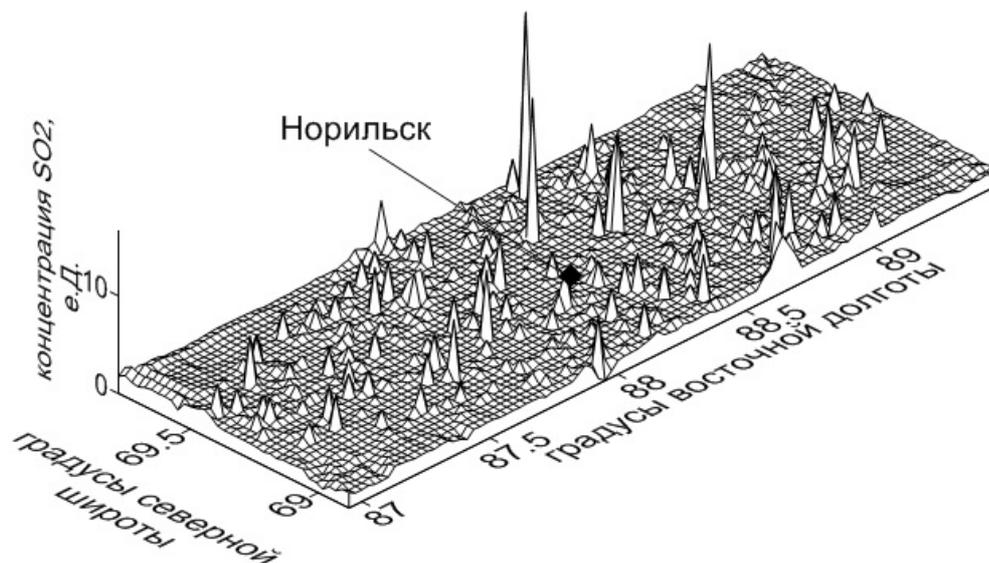


Рис. 7. Результат кластеризации выборки данных наблюдений за 2011 г., высота столба SO₂ от 10 до 15 км.

На рис. 8 представлен результат обработки данных выбросов диоксида серы за 2011 г. на высоте от 0 до 15 км. Видно, что среднее значение концентрации вблизи г. Норильск соответствует 20 е.Д. Пики выбросов достигают значения в 60 е.Д. По мере удаления от г. Норильска концентрация выбросов диоксида серы снижается. Явное снижение концентрации наблюдается в западном направлении.

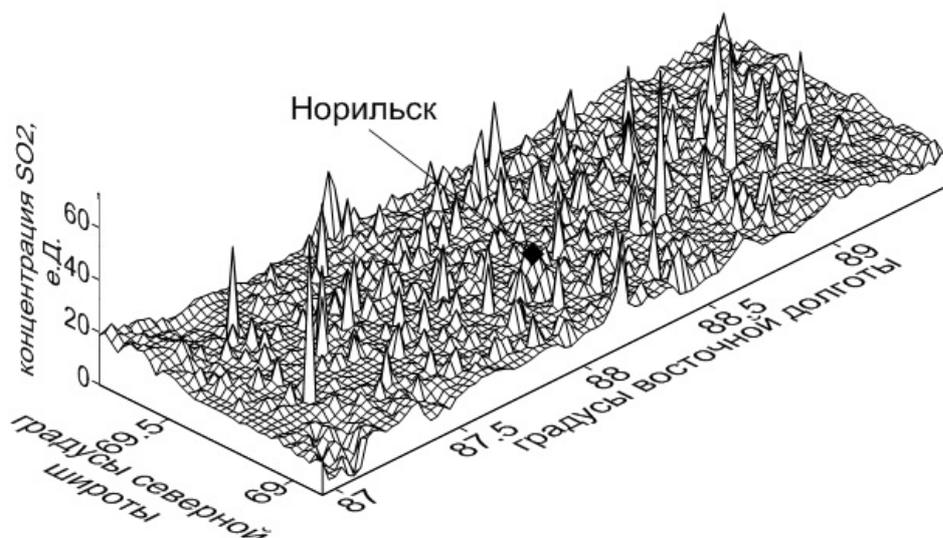


Рис. 8. Результат кластеризации выборки данных наблюдений за 2011 г., высота столба SO₂ от 0 до 15 км.

Сравнительный анализ результатов обработки данных для различных высот показывает, что в приземных слоях атмосферы среднее значение концентрации диоксида серы максимально, и убывает с высотой. На высотах 0–0,9 м среднее значение концентрации SO_2 в 2 раза больше, чем на высотах от 0,9 до 5 км. В некоторых точках для различных высот выделяются пики концентрации диоксида серы, что, по-видимому, вызвано особенностью рельефа местности, розы ветров и расположением дымоотводных труб. Например, в интервалах высот 0,9–5, 5–10, 10–15 км наблюдается пики концентрации SO_2 в точке с координатами $69,42^\circ$ северной широты и $88,90^\circ$ восточной долготы.

На рис. 9 представлен вариант визуализации данных наблюдений для одного дня (01.11.2011 г.), видно, что вблизи г. Норильск существенно превышен уровень концентрации диоксида серы.

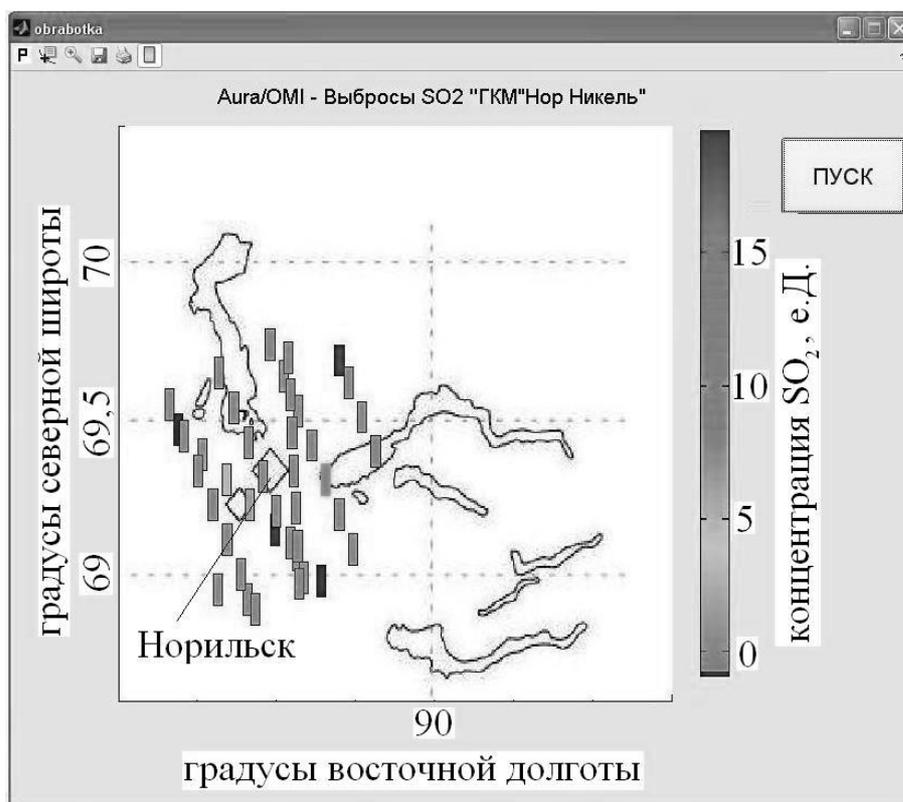


Рис. 9. Визуализация данных измерений OMI в среде MATLAB.

Тестовые расчеты показали, что разработанное алгоритмическое и программное обеспечение достаточно быстро работает на суперкомпьютере Сибирского федерального университета [20] для исследуемого набора данных космического мониторинга атмосферы в районе производства «ГМК «Норильский Никель». В составе вычислительного комплекса находятся 224 вычислительных узла IBM Blade HS21. Каждый узел включает в себя 16 Gb оперативной памяти и два четырехядерных процессора Xeon quad core E5345@2.33 GHz. Система хранения данных IBM DS3400 представляет собой единое дисковое пространство емкостью 20 Tb. Пропускная способность вычислительной сети Infiniband, которая объединяет узлы кластера, достигает 20 Gbit/s. Таким образом, разработанное алгоритмическое и программное обеспечение позволяет анализировать и визуализировать результаты обработки данных в автоматическом режиме за выбранный период космического мониторинга концентрации диоксида серы для разных высот: 0,9, 5, 10 и 15 км.

Заключение

Спутниковые методы исследований являются неотъемлемой частью системы мониторинга экологической обстановки крупных промышленных объектов. Значительно расширить эти наблюдения и предоставить новую информацию о выбросах промышленных объектов позволяют данные, получаемые из космоса прибором Ozone Monitoring Instrument. При помощи ОМІ становится возможным выявлять повышенные концентрации диоксида серы в различных фазах и на отдельных участках производства за короткий промежуток времени.

Анализ результатов обработки данных наблюдений для различных высот показывает, что в приземных слоях атмосферы среднее значение концентрации диоксида серы максимально, и убывает с высотой. На высотах 0-0,9 м среднее значение концентрации SO₂ в 2 раза больше, чем на высотах от 0,9 до 5 км. В некоторых точках для различных высот выделяются пики концентрации диоксида серы, что связано с особенностями рельефа местности, розы ветров и расположением дымоотводных труб. Например, в интервалах высот 0,9-5, 5-10, 10-15 км наблюдается пики концентрации SO₂ в точке с координатами 69,42° северной широты и 88,90° восточной долготы.

Наблюдения за концентрациями SO₂ являются важным направлением исследований для понимания экологической обстановки района «ГМК «Норильский Никель», обеспечения здоровья населения, безопасности полётов авиации. Анализ данных независимого спутникового мониторинга позволяет контролировать концентрацию выбросов SO₂ в районе комбината, а в случаях резкого увеличения объема выбросов диоксида серы может явиться сигналом к уменьшению объемов производства в период неблагоприятных метеорологических условий. Существует возможность за весь период действия программы ОМІ (2004–2011 гг.) оценить эффективность мероприятий комбината по улучшению экологической обстановки региона на основании объективного источника информации о выбросах диоксида серы AURA OMI.

Литература

1. <http://www.nornik.ru/> – Официальный сайт «ГМК «Норильский Никель».
2. Кармановская Н.В. Разработка организационно-методического обеспечения экологической устойчивости в НПП // Диссертация на соискание ученой степени кандидата технических наук. – Норильск: «Норильский индустриальный институт», 2006. – 133 с.
3. Протасов В.Ф. Экология, здоровье и охрана окружающей среды в России. – М.: «Финансы и статистика», 2001. – 672 с.
4. Скурлатов Ю.И. Введение в экологическую химию. – М.: «Высшая школа», 1994. – 400 с.
5. <http://SO2.umbc.edu/omi/> – Сайт NASA по глобальному мониторингу диоксида серы.
6. Krotkov, N.A., Carn, S.A., Krueger, P.K., Yang, K. Band residual difference algorithm for retrieval of SO₂ from the Aura Ozone Monitoring Instrument (OMI) // IEEE Trans. Geosci. Remote Sensing, AURA Special Issue, 44(5), 1259-1266, 2002. – 608 с.

7. Lee C., R.V. Martin, A. van Donkelaar, H. Lee, R. Dickerson, J.C. Hains, N.A. Krotkov, A. Richter, K. Vinnikov, and J.J. Schwab (2011). SO₂ emissions and lifetimes: Estimates from inverse modeling using in situ and global, space-based (SCIAMACHY and OMI) observations // *J. Geophys. Res.*, 116 (D06304).
8. Fioletov, V.E., C.A. McLinden, N.A. Krotkov, M.D. Moran, and K. Yang (2011). Estimation of SO₂ emissions using OMI retrievals// *Geophys. Res. Lett.*, 38 (L21811).
9. Krueger A., N.A. Krotkov, K. Yang, S. Carn, G. Vicente, and W. Schroeder (2009). Applications of Satellite-based Sulfur Dioxide Monitoring // *IEEE*, 2, N4 (December 2009), 293-298.
10. Kazadzisl S., A. Bais, A. Arola, N.A. Krotkov, N. Kouremen, and C. Meleti (2009). Ozone Monitoring Instrument spectral UV irradiance products: comparison with ground based measurements at an urban environment// *Atmos. Chem. Phys.*, 585-594.
11. Lee. C., V. Martin, A. van Donkelaar, G. O Byrne, N.A. Krotkov, A. Richter, L.G. Huey, and J.S. Holloway (2009). Retrieval of vertical columns of sulfur dioxide from SCIAMACHY and OMI: Air mass factor algorithm development, validation, and error analysis// *J. Geophys. Res.*, 114 (D22303).
12. Карчевский С.С., Жуков Л.А., Симонов К.В. Параллельная реализация модифицированного метода динамических ядер для кластерной вычислительной установки. // Материалы XII Всероссийской конференции «Проблемы информатизации региона» ПИР-2011. – Красноярск: СФУ, 2011. – С. 98-102.
13. Кашкин В.Б., Сухинин А.И. Дистанционное зондирование Земли из Космоса. – М: Логос. – 2001. – 264 с.
14. Зуев В.Е., Зуев В.В. Дистанционное оптическое зондирование атмосферы. – М: Гидрометеиздат, 1992. – 232 с.
15. Кароль И.Л., Розанов В.В., Тимофеев Ю.М. Газовые примеси в атмосфере. – Л.: Гидрометеиздат, 1983. – 192 с.
16. Кондратьев К.Я. Научный план системы EOS // Исследования Земли из космоса. – 2000. – № 3. – С. 82–91.
17. Ивлев Л. С., Андреев С. Д. Оптические свойства атмосферных аэрозолей. – Л.: Изд-во ЛГУ, 1986. – 359 с.
18. Креков Г.М., Рахимов Р.Ф. Оптические модели атмосферного аэрозоля. – Томск: Изд-во Томского филиала СО АН СССР, 1986. – 294 с.
19. Воеводин В.В., Воеводин Вл.В. Параллельные вычисления. – М: СПб.: БХВ-Петербург, 2002. – 608 с.
20. Кузьмин Д.А., Маколов С.В., Бугай А.П., Астриков Д.Ю. Комплекс высокопроизводительных вычислений СФУ // sfu-kras.ru: сайт Сибирского Федерального Университета. URL: <http://cluster.sfu-kras.ru/> (дата обращения: 03.11.2011).