

Интеграция параметров спектральных линий молекулы CO₂, содержащихся в банках данных CDSД, в Виртуальный центр атомных и молекулярных данных (VAMDC)

Р.В. Кочанов, В.И. Перевалов, С.А. Ташкун*

*Институт оптики атмосферы им. В.Е. Зуева СО РАН
634021, г. Томск, пл. Академика Зуева, 1*

Поступила в редакцию 22.11.2013 г.

Представлены информация об интеграции спектроскопических банков данных CDSД-296, CDSД-1000 и CDSД-4000, созданных в Институте оптики атмосферы СО РАН, в Виртуальный центр атомных и молекулярных данных (VAMDC) с целью расширения круга их пользователей и краткий обзор VAMDC, затрагивающий назначение системы, ее инфраструктуру и формат данных. Приведена реляционная структура CDSД, адаптированная для нужд VAMDC. Обсуждены решения ряда задач по оптимизации доступа к большим объемам спектроскопических данных.

Ключевые слова: молекулярная спектроскопия, базы данных, банки данных, углекислый газ, CO₂, CDSД, VAMDC, HITRAN; molecular spectroscopy, databases, databanks, carbon dioxide, CO₂, CDSД, VAMDC, HITRAN.

Введение

Знание параметров спектральных линий углекислого газа в инфракрасной области (ИК) спектра необходимо для широкого круга атмосферных и астрофизических приложений. В Институте оптики атмосферы на протяжении более 20 лет ведутся работы по созданию теоретических моделей, глобально описывающих ИК-спектры молекулы CO₂ и ее изотопических модификаций [1]. На их основе были созданы банки параметров спектральных линий для атмосферных CDSД-296 [2] и высокотемпературных CDSД-1000, CDSД-4000 [3, 4] приложений. Все они были востребованы научным сообществом [5, 6]. В частности, они вошли полностью или частично в широко используемые базы данных HITRAN-2008 [7], GEISA [8], HITEMP [9], информационно-вычислительную систему SPECTRA [10], а также в ряд специализированных информационных астрофизических систем, таких как, например, EXOMOL [11] и SPECTRA-FACTORY [12]. Банки данных CDSД доступны на сайте Института по адресу ftp.iao.ru/pub.

Основным препятствием, ограничивающим число возможных пользователей CDSД и затрудняющим их использование, являются форматы представления данных, а также способы хранения самих банков на

носителях информации. Текущие версии банка CDSД следуют стандарту представления информации о параметрах спектральной линии, принятому в текущей версии банка HITRAN [7]. С другой стороны, пользователи, чье программное обеспечение ориентировано на работу, например, с банком GEISA, для доступа к CDSД будут вынуждены адаптировать его под формат CDSД, что создаст дополнительные трудности. Способы хранения банков определяются их размерами. Банки CDSД-296 и CDSД-1000 хранятся на сайте Института в виде одного заархивированного текстового файла. Банк CDSД-4000, имеющий размер более 60 Гбайт, хранится в виде почти 3000 заархивированных текстовых файлов, каждый из которых имеет размер нескольких десятков Мбайт и перекрывает узкий диапазон частот (от 1 до 40 см⁻¹). Такая форма представления для CDSД-4000 обусловлена обеспечением возможности его скачивания по сравнительно медленным каналам связи. Очевидно, что пользователь банков должен быть, по возможности, избавлен от необходимости вникать в детали, не связанные напрямую с его сферой деятельности.

Для этого Европейским сообществом был инициирован проект VAMDC (Virtual Atomic and Molecular Data Centre) [13]. Основной его целью является обеспечение унифицированного доступа к широкому набору информационных систем и банков данных атомов и молекул, ориентированных на атмосферные, астрофизические и высокотемпературные приложения. По замыслу авторов проекта пользователь составляет запрос к VAMDC в терминах, близких к той

* Роман Викторович Кочанов (roman2400@rambler.ru); Валерий Иннокентьевич Перевалов (vip@lts.iao.ru); Сергей Анатольевич Ташкун (tashkun@rambler.ru).

предметной области, в которой он работает. При этом все аспекты, связанные с информационной структурой и форматами представления данных в конкретных системах или банках данных, к которым осуществляется обращение, формируются и обрабатываются самим центром. В частности, пользователь избавлен от необходимости хранить копию банка у себя на компьютере. В настоящее время при помощи VAMDC можно выполнять запросы к нескольким десяткам информационных систем и банков, расположенных по всему миру. Институт представлен в центре банками данных CDSD и информационными системами SPECTRA [10], S&MPO [14] и W@DIS [15]. Более подробная информация о текущем состоянии проекта может быть найдена на сайте VAMDC [16].

В настоящей статье представлен опыт по интегрированию банков данных CDSD в VAMDC. Обсуждены некоторые аспекты структуры и функционирования VAMDC с точки зрения информатики. Приведены некоторые приемы по оптимизации доступа к большим объемам данных в рамках VAMDC. Информация, изложенная в статье, может быть полезна для читателей, желающих ознакомиться с основными характеристиками VAMDC, а также методами интеграции своих банков данных в систему.

Описание проекта VAMDC

VAMDC представляет собой распределенную систему, предоставляющую централизованный доступ к атомно-молекулярным данным посредством веб-портала, расположенного по адресу <http://portal.vamdc.eu>. В настоящее время в распоряжении пользователя находится более 20 источников данных о молекулярных и атомных спектрах, которые могут быть использованы для атмосферных и метеорологических приложений, а также для задач астрофизики, астрохимии и планетологии. Среди них отметим, в частности, CDSD [2, 3], HITRAN [7], GEISA [8], W@DIS [15], S&MPO [17], ETHER [18], MeCaSDa/ECaSDa [19].

На рис. 1 схематически показана инфраструктура VAMDC. Пользователь обращается к узлам через специальный веб-портал. После обработки пользовательского запроса, составленного с помощью специального конструктора, портал пересылает запрос всем узлам, находящимся в реестре (перечень узлов приведен в нижней части диаграммы), и предоставляет пользователю унифицированный доступ к полученным данным. Средствами специальной расширяемой библиотеки программного обеспечения, установленной на сайте портала, пользователь может преобразовывать «сырой» xml-файл данных в различные форматы: csv, plain text, bibtext, html и т.д. Также предусмотрена возможность прямого обращения к узлам системы, минуя портал. В этом случае достаточно ввести в адресную строку веб-браузера html-адрес, представляющий запрос к конкретному узлу и написанный на специальном языке VSS2 (VAMDC SQL Subset) [20]. Однако в этом случае преобразование формата выходного XSAMS файла необходимо про-

изводить на компьютере конечного пользователя. Хотя данный подход является более гибким, он предполагает, что пользователю знакомы языки SQL, VSS (язык запроса), XSLT (язык преобразования). Для широкого пользователя предусмотрен визуальный конструктор запросов и графический веб-интерфейс, позволяющий на интуитивном уровне составлять сложные запросы и преобразовывать данные в нужный формат.

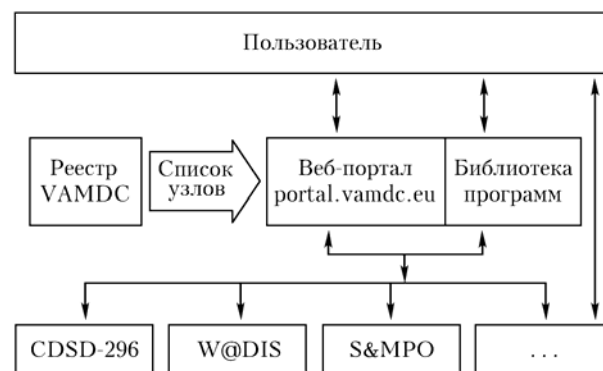


Рис. 1. Базовая инфраструктура VAMDC

Для описания разнообразных атомно-молекулярных данных в VAMDC используется XML-схема для атомной и молекулярной спектроскопии (XSAMS) [21], которая представляет собой иерархическое описание спектроскопической предметной области на языке XML. В настоящий момент в системе VAMDC спецификация XSAMS применяется как промежуточный формат данных, использующийся при их передаче от узла VAMDC к клиенту. Схема данных XSAMS позволяет получить информацию от узлов системы в едином формате, предназначенном для последующей программной обработки для встраивания в конечный пользовательский проект.

Пример запроса VAMDC

В качестве примера можно рассмотреть запрос на выдачу линий основного изотопа CO_2 с волновым числом в диапазоне от 2000 до 2500 cm^{-1} и отсечкой линий по коэффициенту Эйнштейна A от 10^{-5} $1/\text{с}$, причем энергии верхних колебательно-вращательных уровней должны быть больше 1000 cm^{-1} . На рис. 2 показан процесс создания этого запроса (portal.vamdc.eu).

По желанию пользователя запрос можно уточнять, добавляя различные дополнительные формы по атомам, молекулам, радиационным процессам, столкновениям, параметрам сред и т.д. Узлы, которые располагают необходимой информацией, соответствующей запросу, выделены зеленым цветом в списке, расположенном на экране монитора справа. Когда пользователь вводит все интересующие его ограничения, он может приступить к опросу узлов, нажав на кнопку «Find data». Результаты запроса можно будет либо сохранить на диск, либо просмотреть в браузере в виде таблицы.

Find data Save query

Molecule 1
Clear Remove <

Chemical name		<input type="text" value="Carbon dioxide"/>
Stoichiometric formula	Общее описание молекулы	<input type="text" value="CO<sub>2</sub>"/>
Structural formula		<input type="text" value="CO<sub>2</sub>"/>
Spin isomer		<input type="text"/>
Standard InChiKey	<input type="text" value="CURLTUGMZLYLDI-UHFFFAOYSA-N"/>	

Search by stoichiometric formula if no isotopologue is selected.

Isotopologue

<input checked="" type="checkbox"/> Carbon dioxide CO ₂	
<input type="checkbox"/> Carbon Dioxide ¹³ C ¹⁶ O ₂	Изотопологи
<input type="checkbox"/> Carbon Dioxide ¹⁶ O ¹² C ¹⁸ O	

Radiative
Clear Remove <

Wavenumber	<input type="text" value="2000"/>	to	<input type="text" value="2500"/>	<input type="text" value="cm<sup>-1</sup>"/>
Equivalent Wavelength	Wavelength from 40000.0 to 50000.0 A			
Upper state energy	<input type="text" value="1000"/>	to	<input type="text"/>	<input type="text" value="cm<sup>-1</sup>"/>
Equivalent to	Upper state energy from 1000.0 to 1/cm			
Lower state energy	<input type="text"/>	to	<input type="text"/>	<input type="text" value="1/cm"/>
Equivalent to	1/cm			
Probability, A	<input type="text" value="1e-5"/>	to	<input type="text"/>	<input type="text" value="1/s"/>

Рис. 2. Пример составления запроса к узлам VAMDC на сайте portal.vamdc.eu

Банки данных CDSD

Углекислый газ является важной составляющей в атмосфере Земли. Для успешного мониторинга содержания CO₂ в атмосфере необходимо располагать качественными теоретическими предсказаниями его спектров. Кроме того, моделирование спектров углекислого газа имеет широкое прикладное значение, начиная с атмосферных приложений и заканчивая высокотемпературными астрофизическими. Ранее в лаборатории теоретической спектроскопии ИОА в сотрудничестве с другими лабораториями было реализовано моделирование спектров высокого разрешения для изотопических модификаций, наиболее распространенных в земной атмосфере. Спектры были про-

моделированы для трех разных температур (296, 1000 и 4000 K) и соответственно разделены на три банка данных: CDSD-296 [2], CDSD-1000 [3] и CDSD-4000 [4].

В табл. 1 приведена краткая сводка по линиям для каждой изотопической модификации. Здесь N означает количество линий; ν – волновое число (см⁻¹); S – интенсивность линии (см⁻¹/(мол · см⁻²)) при $T = 296$ K, в случае CDSD-1000 при $T = 1000$ K).

Банки CDSD включают параметры линий, предусмотренные форматом HITRAN-2008 [7]. Основные параметры (волновые числа перехода, интенсивности, коэффициенты сдвига и уширения, идентификация и т.д.) интегрированы в VAMDC, как это показано в табл. 2.

Таблица 1

Сводка по КВ-переходам банков данных

Банк данных	iso	N	v_{\min}	v_{\max}	S_{\min}	S_{\max}
CDS-296	$^{12}\text{C}^{16}\text{O}_2$	170486	345,9	12784,1	1E-30	3,520E-18
	$^{13}\text{C}^{16}\text{O}_2$	68510	433,2	12462,0	1E-30	3,740E-20
	$^{16}\text{O}^{12}\text{C}^{18}\text{O}$	109852	5,9	11422,6	1E-30	6,870E-21
	$^{16}\text{O}^{12}\text{C}^{17}\text{O}$	18969	10,6	8270,1	1E-30	1,260E-21
	$^{16}\text{O}^{13}\text{C}^{18}\text{O}$	39007	449,7	6744,2	1E-30	7,810E-23
	$^{16}\text{O}^{13}\text{C}^{17}\text{O}$	2741	580,9	6768,6	1E-30	1,400E-23
	$^{12}\text{C}^{18}\text{O}_2$	10045	484,3	8162,7	1E-30	1,330E-23
CDS-1000	$^{12}\text{C}^{16}\text{O}_2$	2508800	257,9	9648,0	1E-27	6,240E-19
	$^{13}\text{C}^{16}\text{O}_2$	520260	418,2	8104,7	1E-27	6,400E-21
	$^{16}\text{O}^{12}\text{C}^{18}\text{O}$	565090	397,5	6937,4	1E-27	1,200E-21
	$^{16}\text{O}^{12}\text{C}^{17}\text{O}$	224966	428,4	2691,4	1E-27	2,210E-22
	$^{16}\text{O}^{13}\text{C}^{18}\text{O}$	80942	515,7	4958,0	1E-27	1,310E-23
	$^{16}\text{O}^{13}\text{C}^{17}\text{O}$	33653	546,6	4880,1	1E-27	2,380E-24
	$^{12}\text{C}^{18}\text{O}_2$	16842	520,1	3689,6	1E-27	2,310E-24
CDS-4000	$^{12}\text{C}^{16}\text{O}_2$	573881316	226,4	8309,8	7E-109	3,520E-18
	$^{13}\text{C}^{16}\text{O}_2$	30805168	405,5	6796,9	3E-105	3,750E-20
	$^{16}\text{O}^{12}\text{C}^{18}\text{O}$	18282396	393,7	5139,6	2E-95	6,880E-21
	$^{16}\text{O}^{12}\text{C}^{17}\text{O}$	5355574	456,5	4975,6	6E-85	1,260E-21

Таблица 2

Параметры линий CO_2 для банков CDS, интегрированные в VAMDC. Привязка параметров осуществлялась с помощью: *dict* – глобального словаря VAMDC (в скобках указано имя параметра в словаре); *case-by-case* – имен, предусмотренных в части схемы XSAMS Case By Case; *raw-xml* – чистого xml

Имя параметра	Значение	Интеграция в VAMDC (XSAMS)
ν	Волновое число	<i>dict</i> (RadTransWavenumber)
λ	Длина волны	<i>dict</i> (RadTransWavelength)
S	Интенсивность	<i>dict</i> (RadTransProbabilityLineStrength)
A	Коэффициент Эйнштейна	<i>dict</i> (RadTransProbabilityA)
γ_{air}	Коэффициент уширения воздухом (γ_{air})	<i>raw-xml</i>
γ_{self}	Коэффициент самоуширения (γ_{self})	<i>raw-xml</i>
n_{air}	Коэффициент температурной зависимости для γ_{air}	<i>raw-xml</i>
n_{self}	Коэффициент температурной зависимости для γ_{self}	<i>raw-xml</i>
δ_{air}	Коэффициент сдвига давлением воздуха	<i>dict</i> (RadTransShiftingParamValue)
E', E''	Энергия верхнего (') и нижнего (") состояний	<i>dict</i> (MoleculeStateEnergy)
ν'_1, ν''_1 ν'_2, ν''_2 l'_2, l''_2 ν'_3, ν''_3 r', r''	Спектроскопическая идентификация верхнего (') и нижнего (") колебательных состояний: квантовые числа для нормальных мод и резонансных полиад Ферми	<i>dict</i> (MoleculeQNv1) <i>dict</i> (MoleculeQNv2) <i>dict</i> (MoleculeQNl2) <i>dict</i> (MoleculeQNv3) <i>case-by-case</i>
P', P'' c', c'' n', n''	Глобальная идентификация верхнего (') и нижнего (") колебательных состояний: номер полиады (p), симметрия Ванга (c) и номер в блоке матрицы гамильтониана (n)	<i>raw-xml</i> <i>case-by-case</i> <i>raw-xml</i>
J', J''	Вращательные квантовые числа верхнего (') и нижнего (") состояний	<i>dict</i> (MoleculeQNJ)

Выбор реляционной структуры данных для CDS и ее оптимизация

В общей сложности банки CDS-296, CDS-1000 и CDS-4000 включают в себя 419610, 3950553 и 628324454 линий соответственно. Банк CDS-4000

в бинарном виде занимает дисковый объем порядка 60 Гбайт. Задачи обработки данных таких размеров требуют достаточно тщательного проектирования и организации схемы данных.

В качестве хранилища данных для всех трех банков была выбрана реляционная система управления базами данных MySQL v5.5 [22]. В качестве

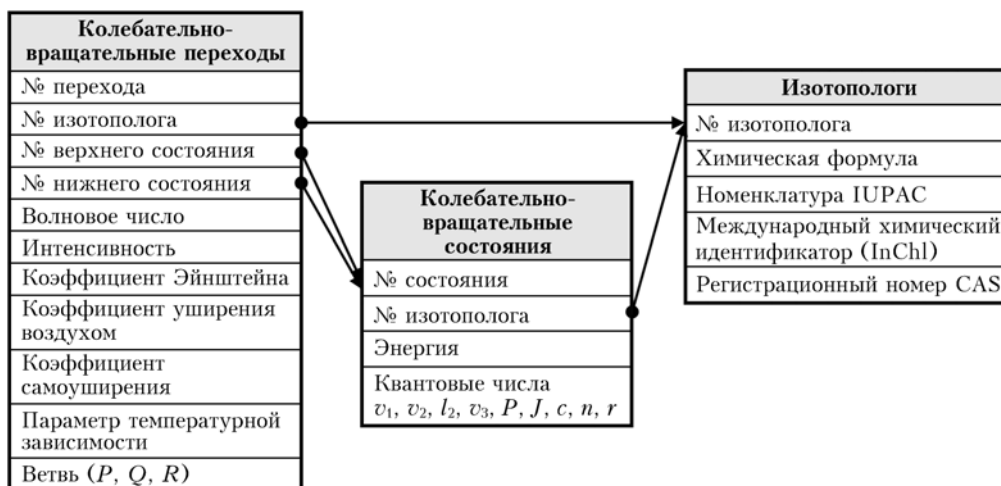


Рис. 3. Организация данных для банков CDS-296, CDS-1000 и CDS-4000 в рамках VAMDC

движка таблиц использован MyISAM. Этот движок позволяет оптимизировать базу данных по скорости выполнения запросов на чтение.

Для того чтобы оптимизировать используемое пространство, к банкам данных была применена *нормализация*. Это означает, что из таблиц, содержащих переходы с нижнего на верхнее колебательно-вращательное состояние, были выделены в явном виде таблицы состояний, затем полные идентификации состояний в таблицах переходов были заменены на целочисленные ссылки, тем самым была уменьшена избыточность данных, которая имеет место для баз в формате HITRAN.

Система MySQL 5.5 поддерживает разбиение таблицы данных на части (*partitioning*), что позволяет выполнять базовые запросы к таблицам больших размеров (например, таблица переходов из CDS-4000) за приемлемое время. В качестве основного решения проблемы больших объемов данных применено разбиение таблицы переходов для CDS-4000 по значению длины волны λ на 50 частей. Такое разбиение позволяет отсечь большую часть нерелевантных данных в случае, если в запросе присутствует интервальное ограничение на длину волны. В результате запрос на чтение с ограничением на λ для CDS-4000 выполняется за 1–5 с для достаточно плотной концентрации подходящих под запрос линий в указанном диапазоне длин волн. Дальнейшая проработка схемы разбиения находится в процессе.

Также для оптимизации времени скорости выполнения запросов на чтение было использовано индексирование по наиболее употребляемым параметрам спектральных линий, таким как волновые числа

переходов, длины волн, интенсивности, квантовые числа и номера изотопических модификаций.

Таким образом, все банки данных CDS, представленные в VAMDC, имеют одинаковую структуру, которая показана на рис. 3. База данных для каждой температуры разделена на таблицы переходов, уровней и изотопологов. Стрелками обозначены ссылки от одной таблицы к другой. К примеру, в таблице переходов каждый переход содержит ссылки на два состояния (верхнее и нижнее) и на соответствующий изотополог.

Интеграция банков данных CDS в VAMDC

Для функционирования в рамках VAMDC банков данных CDS-296, CDS-1000 и CDS-4000 было установлено специальное программное обеспечение (Т. Маркварт, [23]). Данное программное обеспечение написано на языке Python и использует возможности веб-сервера Apache для поддержки протокола передачи данных VAMDC. Модуль на Python позволяет серверу запускать ядро MySQL для выполнения запросов к базам данных.

В настоящее время реализовано по 3 узла на каждую из активных версий XSAMS (версии 0.3 и 1.0). Ссылки на узлы CDS-296, CDS-1000 и CDS-4000 содержатся в табл. 3.

Для проверки работоспособности узла, а также для информации о возвращаемых параметрах линий в рамках схемы XSAMS пользователь может зайти по ссылке *адрес_узла/тап/capabilities*, где адрес узла дан в табл. 3.

Таблица 3

Ссылки на узлы CDS для двух версий схемы данных XSAMS

Узел CDS	Версия XSAMS	
	v0.3	v1.0
CDS-296	http://lts.iao.ru/node/cdsd-296	http://lts.iao.ru/node/cdsd-296-xsams1
CDS-1000	http://lts.iao.ru/node/cdsd-1000	http://lts.iao.ru/node/cdsd-1000-xsams1
CDS-4000	http://lts.iao.ru/node/cdsd-4000	http://lts.iao.ru/node/cdsd-4000-xsams1

Для обращения к узлам напрямую, минуя портал VAMDC, пользователю необходимо составить http-запрос вида *адрес_узла/tap/sync?LANG=VSS2&REQUEST=doQuery&FORMAT=XSAMS&QUERY=* запрос, причем запрос формируется на языке VSS2 [20]. К примеру, введенный в адресную строку веб-браузера запрос *http://lts.iao.ru/node/cdsd-4000/tap/sync?LANG=VSS2&REQUEST=doQuery&FORMAT=XSAMS&QUERY=select * where RadTrans Wavelength>20000.0* возвратит все линии из банка CDS-4000, длина волны которых больше 20000 Å, в формате XSAMS. Для уточнения формата, актуального на момент прочтения данной статьи, авторы рекомендуют обратиться к сайту portal.vamdc.eu, на котором можно использовать конструктор создания запросов, выдающий явный вид запроса на языке VSS2.

Заключение

Проведенная работа по встраиванию банков данных CDS в систему VAMDC позволила упростить запросы к этим банкам и тем самым расширить круг их пользователей. Выполнена первичная оптимизация структуры соответствующих баз данных с целью ускорения запросов на чтение и уменьшения занимаемого пространства. Даны ссылки на интернет-ресурсы, предоставляющие доступ к узлам CDS как в рамках системы VAMDC, так и напрямую через интерфейс адресной строки веб-браузера.

1. *Перевалов В.И., Ташкун С.А., Тютерева Вл.Г., Люлин О.М.* Глобальное моделирование спектров высоко разрешения молекул атмосферных газов // *Оптика атмосф. и океана.* 2007. Т. 20, № 5. С. 398–407.
2. *Perevalov V.I., Tashkun S.A.* CDS-296 (Carbon Dioxide Spectroscopic Databank): updated and enlarged version for atmospheric applications // *Proc. the 10th HITRAN database conference.* June 2008, Cambridge, MA, USA. URL: <ftp://ftp.iao.ru/pub/CDS-2008/>
3. *Tashkun S.A., Perevalov V.I., Teffo J.L., Vykov A.D., Lavrentieva N.N.* CDS-1000, the high-temperature carbon dioxide spectroscopic databank // *J. Quant. Spectrosc. and Radiat. Transfer.* 2003. V. 82, N 1–4. P. 165–196.
4. *Tashkun S.A., Perevalov V.I.* CDS-4000: High-resolution, high-temperature carbon dioxide spectroscopic databank // *J. Quant. Spectrosc. and Radiat. Transfer.* 2011. V. 112, N 9. P. 1403–1410.
5. *Ташкун С.А., Перевалов В.И.* Радиационные свойства CO₂: спектроскопические банки данных для атмосферных и высокотемпературных приложений // *Оптика атмосф. и океана.* 2011. Т. 24, № 12. С. 1108–1112.
6. *Ченцов А.В., Воронина Ю.В., Чеснокова Т.Ю.* Моделирование атмосферного пропускания с различными контурами линий поглощения CO₂ // *Оптика атмосф. и океана.* 2013. Т. 26, № 9. С. 711–715.

7. *Rothman L.S., Gordon I.E., Barbe A., Benner D.C., Bernath P.F., Birk M., Boudon V., Brown L.R., Campargue A., Champion J.P., Chance K., Coudert L.H., Dana V., Devi V.M., Fally S., Flaud J.-M., Gamache R.R., Goldman A., Jacquemart D., Lacombe N., Lafferty W.J., Mandin J.-Y., Massie S.T., Mikhailenko S., Moazzen-Ahmadi N., Naumenko O., Nikitin A., Orphal J., Predoi-Cross A., Perevalov V., Perrin A., Rinsland C.P., Rotger M., Šimečková M., Smith M.A.H., Tashkun S., Tennyson J., Toth R.A., Vandaele A.-C., Vander Auwera J.* The HITRAN 2008 molecular spectroscopic database // *J. Quant. Spectrosc. and Radiat. Transfer.* 2009. V. 110, N 9–10. P. 533–572.
8. *Jacquinet-Husson N., Scott N.A., Chédin A., et al.* The GEISA spectroscopic database: Current and future archive for Earth and planetary atmosphere studies // *J. Quant. Spectrosc. and Radiat. Transfer.* 2008. V. 109, N 6. P. 1043–1059.
9. *Rothman L.S., Gordon I.E., Barber R.J., Dothe H., Gamache R.R., Goldman A., Perevalov V.I., Tashkun S.A., Tenmyson J.* HITEMP, the high-temperature molecular spectroscopic database // *J. Quant. Spectrosc. and Radiat. Transfer.* 2010. V. 111, N 15. P. 2139–2150.
10. *Михайленко С.Н., Бабилов Ю.Л., Головки В.Ф.* Информационно-вычислительная система «Спектроскопия атмосферных газов». Структура и основные функции // *Оптика атмосф. океана.* 2005. Т. 18, № 9. С. 765–776.
11. *Tennyson J., Yurchenko S.N.* ExoMol: molecular line lists for exoplanet and other atmospheres // *Mon. Not. Roy. Astron. Soc.* 2012. V. 425, N 1. P. 21–33.
12. *Cami J., Van Malderen R., Markwick A.J.* SPECTRAFACTORY.NET: A database of molecular model spectra // *Astrophys. J. Suppl. Series.* 2010. V. 187, N 2. P. 409–415.
13. *Dubernet M.L. et al.* Virtual atomic and molecular data centre // *J. Quant. Spectrosc. and Radiat. Transfer.* 2010. V. 111, N 15. P. 2151–2159.
14. *Михайленко С.Н., Бабилов Ю.Л., Тютерева Вл.Г., Барбе А.* Банк данных по спектроскопии озона доступный в интернете (S&MPO) // *Вычислительные технологии.* 2002. Т. 7. С. 64–70 [*Mikhailenko S.N., Babikov Yu.L., Tyutereva V.I.G., Barbe A.* The databank of ozone spectroscopy on WEB (S&MPO)] // *J. Comput. Technol.* 2002. V. 7, N 7. P. 64–70.
15. URL: <http://wadis.saga.iao.ru/saga2/>
16. URL: <http://www.vamdc.eu/>
17. *Babikov Y.L., Tyutereva V.I.G., Barbe A., Mikhailenko S.N., Starikova E.N.* O₃Calc (GSMA-S&MPO): Ozone database adaptation for VAMDC // *VAMDC Annual Meeting, Vienna, Austria, February 21–24. 2012.*
18. URL: <http://www.pole-ether.fr/>
19. *Ba Y.A. et al.* MeCaSDa and ECaSDa: Methane and ethane calculated spectroscopic data bases for the virtual atomic and molecular data centre // *J. Quant. Spectrosc. and Radiat. Transfer.* 2013. In press.
20. URL: <https://vamdc-standards.readthedocs.org/en/latest/queryLanguage/vss2.html>
21. URL: <http://www.vamdc.eu/documents/standards/dataModel/vamdcxsams/>
22. URL: <http://dev.mysql.com/>
23. URL: <http://www.vamdc.eu/software/>

R. V. Kochanov, V. I. Perevalov, S. A. Tashkun. Integration of CO₂ spectral line parameters from the CDS-296, CDS-1000 and CDS-4000 into the Virtual Atomic and Molecular Data Centre (VAMDC).

The spectroscopic databanks CDS-296, CDS-1000 and CDS-4000 are presented in the context of integration into the Virtual Atomic and Molecular Data Centre (VAMDC) in order to extend the number of their users. A brief review of the VAMDC system is given concerning the purpose of this project, its infrastructure, and the data representation format. The relational structure of CDS databases, adapted to VAMDC needs, is presented. Some of the technical problems connected with the huge volumes of stored information are considered.