

Развитие информационных систем в ИОА СО РАН

А.З. Фазлиев*

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

Поступила в редакцию 08.07.2009 г.

Дан краткий обзор развития информационных систем по молекулярной спектроскопии и атмосферной радиации, созданных в ИОА СО РАН за последние 25 лет. Информационные системы по молекулярной спектроскопии описаны с разных точек зрения: функциональной, структурной и технологической.

Ключевые слова: информационная система, молекулярная спектроскопия; information system, molecular spectroscopy.

Введение

Значимость информационных систем (ИС) в научной деятельности определена их функциями посредников между исследователями. В них собираются, хранятся и публикуются данные, извлекаются информация и знание. Автоматизированные информационные системы дают пользователям возможность проводить вычисления и в последние годы делать логические выводы в рамках онтологий. Заметим, что ниже термин «информационные системы» [1] используется в двух контекстах. В широком смысле он трактуется как результат представления знаний в предметной области. Исследовательская и технологическая деятельность в этом направлении в Институте оптики атмосферы (ИОА) СО РАН частично описана в работах [2–6]. При интерпретации в узком смысле имеется в виду создание собственно информационных систем [6–16] и их компонентов: баз данных и знаний [17–23], интерфейсов пользователя и приложений в предметных областях [24–28].

Исследования и разработки ИС и их компонентов, выполненные в ИОА СО РАН, опубликованы более чем в сотне статей и нескольких монографиях. Их существенная часть ориентирована на традиционные для ИОА СО РАН предметные области: дистанционное зондирование и распространение излучения в атмосфере, атмосферный аэрозоль, атмосферная химия и т.д. Работы в этом направлении интенсивно проводились в 80–90-е гг. и в настоящее время практически не ведутся. Дать анализ всех публикаций в исторической перспективе в рамках небольшой статьи не представляется возможным. По этой причине статья ограничена описанием двух предметных областей, в которых исследование предметных областей, их создание и модификации продолжаются в настоящее время: молекулярная спектроскопия [2–14, 17–20, 25, 26] и атмосферная радиация [15, 16, 21–25, 27, 28].

Как атмосферная спектроскопия, так и радиация имеют свои особенности. В молекулярной спектроскопии число исследуемых молекул составляет чуть более сотни, при этом исследователи используют объемы данных, превышающие 50 Гбайт, а сложные вычисления с большим временем счета в ИС не включаются. В атмосферной радиации количество спектральных данных невелико (0,2 Гбайт), но используются данные по атмосферному аэрозолю, облакам и т.д. Число исследуемых объектов возрастает почти вдвое. Естественным следствием этих особенностей являются более сложная структура информационных объектов в радиации и, как следствие, интерфейсы конечного пользователя. Для обеих этих предметных областей сложной, и в настоящее время окончательно не формализованной, является проблема достоверности данных. Они связаны между собой спектральными данными, требуемыми для расчетов в атмосферной радиации.

Статья состоит из двух частей. В первой части описана эволюция созданных в ИОА СО РАН информационных систем по молекулярной спектроскопии. Показан поэтапный переход, выполненный в 80–90-е гг., от ИС, работающих на клиентском месте, к ИС, работающим в глобальной сети. Важным моментом этого развития явился также переход от ИС с однослойной архитектурой к информационным системам с трехслойной архитектурой.

Во второй части дан краткий обзор информационных систем и их компонентов, созданных в ИОА СО РАН в области атмосферной радиации.

1. Информационные системы по молекулярной спектроскопии

Работы по созданию ИС по молекулярной спектроскопии были начаты в ИОА в начале 80-х гг. Технические средства для их реализации, которыми обладали разработчики, а именно алгоритмические языки и файловые системы для организации данных, позволяли им решать задачи соз-

* Александр Зарипович Фазлиев (faz@iao.ru).

дания ИС, которые с сегодняшних позиций кажутся неуклюжими. Тем не менее основные идеи по представлению молекулярной области в виде набора информационных объектов до сих пор оказываются не реализованными (рис. 1).

Основные задачи по созданию ИС, поставленные в то время, были разделены на три уровня:

1) формирование, хранение и организация выборки рекомендованных и экспериментальных данных, систематизация параметров спектральных линий, классификация источников данных;

2) интеграция программ решения обратных спектроскопических задач, статистического анализа и синтеза информации, а также оценка достоверности данных в ИС;

3) реализация математических моделей.

Предложенное трехуровневое разделение смешивает между собой задачи разных предметных областей: спектроскопии, математической обработки результатов исследований и информатики. Из рис. 1 в явном виде следует разделение интересов обработчиков экспериментальных данных (уровень I и II) и теоретиков, решающих прямые задачи спектроскопии (уровень III). Отмеченные на рис. 1 штрихпунктиром части (кроме блока графического представления информации) так и не были реализованы. Основными данными в ИС первого поколения были параметры спектральных линий, размещаемые в файловой системе. С помощью командной строки был организован диалог пользователя с ИС. Приложения ИС содержались в системе ввода данных эксперимента. Создание первого поколения информационных систем была закончено в 1990 г. [2, 4], и во многом определило направление

развития ее потомков. Отдельные компоненты этой генерации наследовались всеми последующими ИС.

Следующее поколение ИС было связано с появлением персональных компьютеров, позволивших использовать графический интерфейс. Единственным его представителем была ИС AirSentry [9]. По своему назначению она была ориентирована на решение конкретных задач, в первую очередь на вычисление спектральных функций для чистых газов и их смесей при задании разных аппаратных функций. Ее функциональность была существенно меньше, чем у спроектированных информационных систем первого поколения, а техническое решение акцентировано на графическом представлении результатов вычислений.

Третье поколение ИС связано с освоением технологий представления информации в глобальной сети. Первыми ИС по молекулярной спектроскопии с доступом по сети Интернет были системы «Атмосферная спектроскопия» (<http://spectra.iao.ru>) [10] и «Спектроскопия и молекулярные свойства озона» (<http://smpro.iao.ru>) [8]. ИС «Spectra» основана на модели молекулярной спектроскопии, использованной в [9]. Вторая ИС «S&MPO» дополнила эту же модель архивом спектроскопических параметров для молекулы озона, уровнями энергии озона, поверхностью потенциальной энергии молекулы, поверхностью дипольного момента, параметрами колебательно-вращательных гамильтонианов, волновыми функциями колебательных состояний и др. [8]. Для этой молекулы был сделан шаг вперед в реализации уровня III (см. рис. 1). К сожалению, для других молекул больше ничего подобного не создавалось. Схема этого поколения ИС показана на рис. 2.

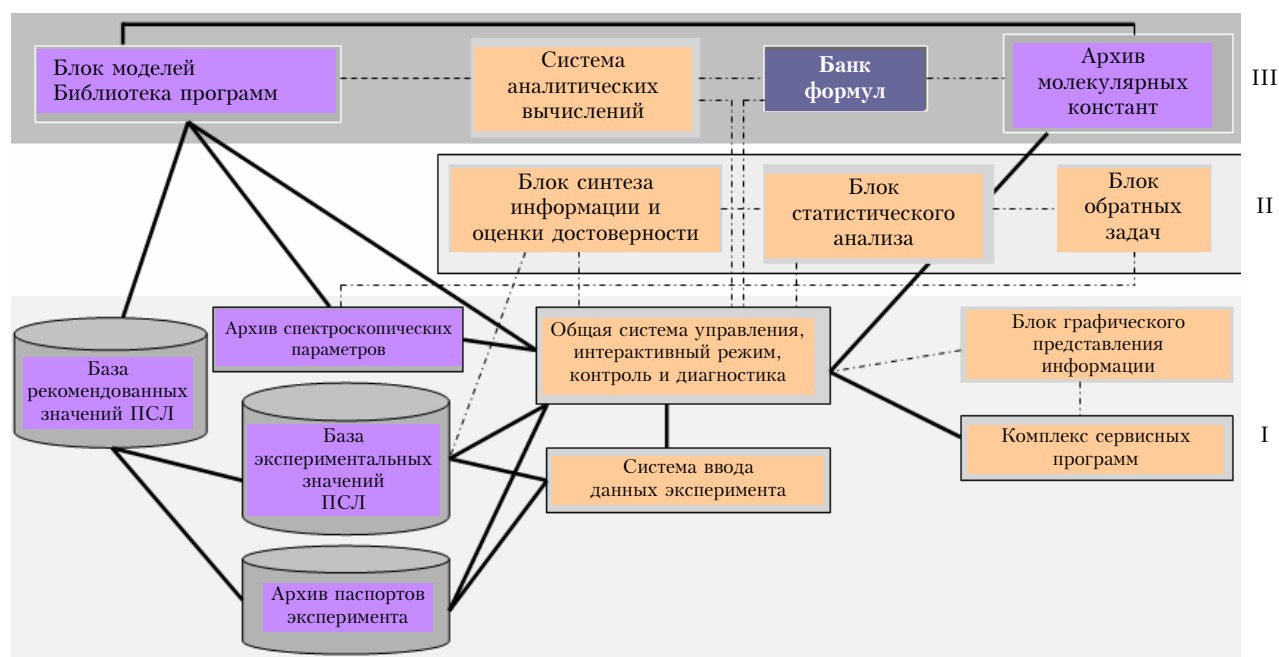


Рис. 1. Схема информационной системы по молекулярной спектроскопии [4]

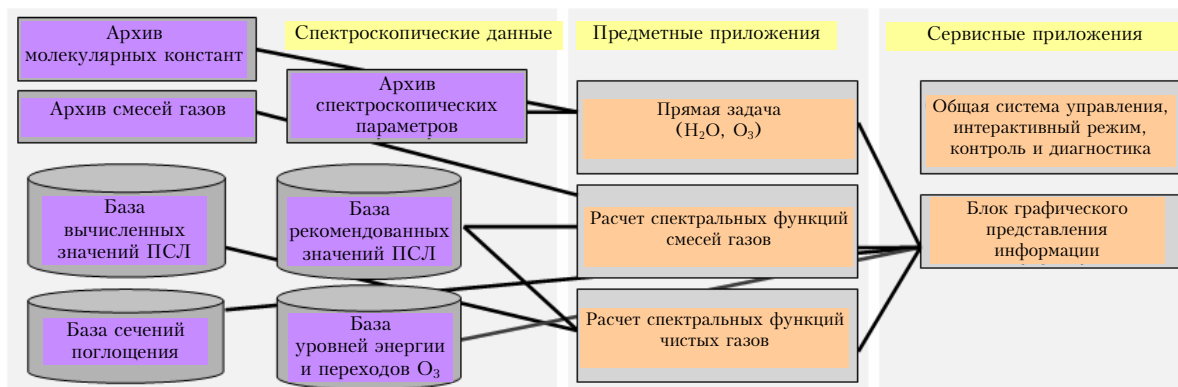


Рис. 2. Схема информационной системы по молекулярной спектроскопии [8, 10]. Предметная область представлена спектроскопическими данными и предметными приложениями. Все компоненты ИС были реализованы

В основу систематизации положен иной принцип по сравнению с использованным на рис. 1. Такого рода классификация соответствует принятой при общем описании автоматизированных ИС [1].

Существенной чертой этих систем было создание их зеркал во Франции. На наш взгляд, это была первая в России успешная попытка создания зеркал информационных ресурсов для совместной работы, но эти системы не были распределенными.

Архитектура описанных выше ИС содержала только слой данных и вычислений. Для пользователя это создавало определенные неудобства. В таких системах можно манипулировать данными, проводить вычисления, но вопрос достоверности полученных результатов остается неопределенным. Информационные системы с такой архитектурой до сих пор являются широко распространенными во многих предметных областях.

В конце 90-х гг. появились работы (например, [29]), предлагающие новую архитектуру ИС, ориентированную на работу с вычислительным и семантическим гридом. Эта архитектура включала в себя дополнительные два слоя в ИС: информационный слой и слой знаний. Информационный слой характеризовал свойства данных предметной области, а слой знаний был ориентирован на представление отдельных данных и их свойств в виде прикладных онтологий. Технологическая поддержка, необходимая для их организации, осуществлялась как W3C [принятие необходимых рекомендаций], в первую очередь, языков разметки (XML, RDF, OWL, ...), так и рядом крупных фирм, создавших свободное программное обеспечение для работы с размеченными ресурсами.

В работах [12–14] по созданию ИС *W@DIS* были реализованы все три слоя современной архитектуры ИС. Модель молекулярной спектроскопии в виде цепей прямых и обратных задач была использована в качестве структуры, группирующей параметры спектральных линий, лайнлисты, уровни энергий и спектральные функции. Математические модели, используемые в задачах, предоставили ограничения на значения физических сущностей, входящих в решения спектроскопических задач. Сильным ограничением, не связанным с мате-

матическими моделями физических объектов, стало ограничение, связанное с публикацией решений задач. Оно состояло в том, что неопубликованные данные, т.е. не существующие в сети Интернет, не рассматривались как достоверные в рамках информационного пространства. Введенные ограничения позволили определить понятие достоверности решений в явном виде. Тем самым стала возможной реализация блока синтеза информации и проверки достоверности (см. рис. 1).

ИС *W@DIS* содержала систему ввода решений прямых и обратных задач, обобщив функциональность системы ввода экспериментальных данных ИС первого поколения. Эта система, дополненная приложениями, обеспечивающими ей возможность работы со знаниями (рис. 3), стала в настоящее время базовой для развития распределенной информационной системы с узлами в Томске, Нижнем Новгороде и Санкт-Петербурге. Процесс создания распределенной ИС начался в середине 2000-х гг. и в настоящее время не закончен.

Информационные системы всех поколений содержали оригинальные расчетные и экспериментальные данные о параметрах спектральных линий, полученные в ИОА СО РАН. Среди публикаций можно выделить работы, содержащие данные о молекулах воды [18], оксидов азота и углерода [17, 20], сероводорода [19] и озона [8].

2. Информационные системы по атмосферной радиации и дистанционному зондированию атмосферы

Исследование физических процессов, знание которых требовалось для дистанционного зондирования атмосферы, долгое время являлось основным направлением в развитии ИОА СО РАН. Программное обеспечение, разрабатываемое для решения отдельных задач, с которыми сталкивались при дистанционном зондировании атмосферы, в ИОА СО РАН начали объединять в комплексы и пакеты программ с начала 80-х гг., и эта работа продолжалась в течение 20 лет [24, 27, 28].

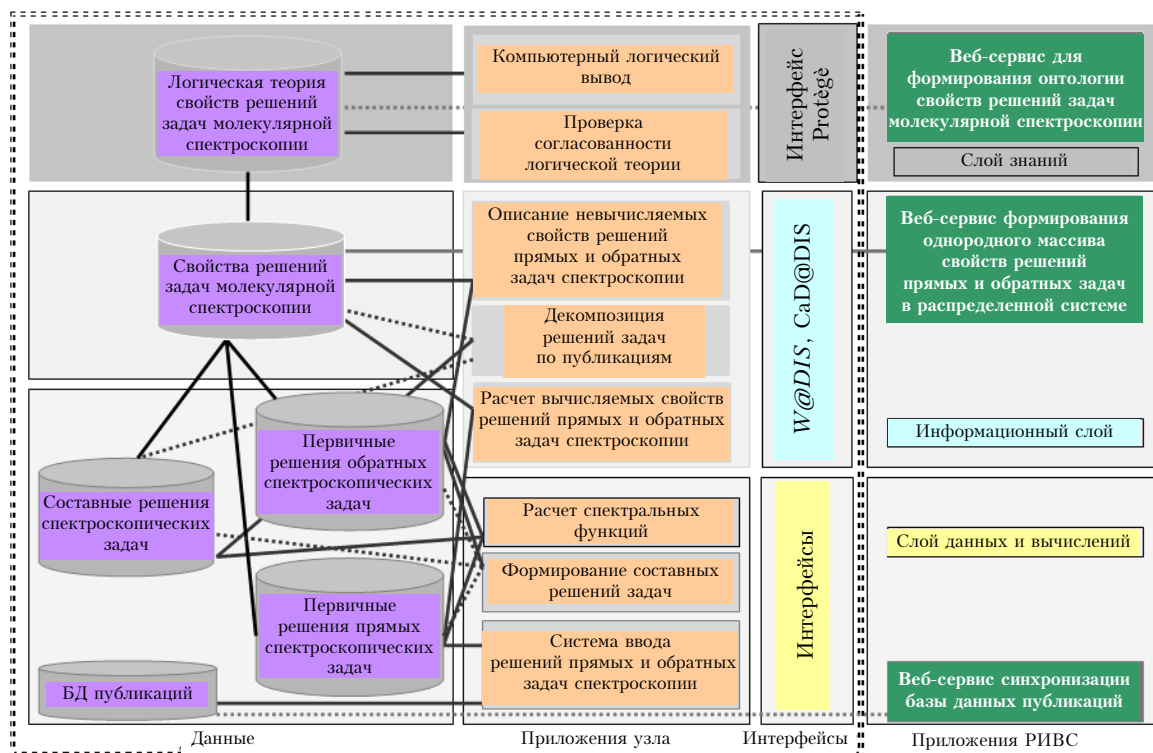


Рис. 3. Архитектура узла РИВС W@DIS. Пунктиром показаны не реализованные на момент публикации связи

В этот период времени возник пакет программ LARA, позволяющий проводить расчеты коэффициентов молекулярного и аэрозольного ослабления, функции пропускания для наклонных и горизонтальных атмосферных трасс.

В расчетах коэффициента поглощения использовался метод полинейного счета, а аэрозольное ослабление рассчитывалось по моделям атмосферного аэрозоля [30]. Учет континуального поглощения проводился с учетом теории крыльев линий, разработанной С.Д. Твороговым и др. [31]. Пакет программ AIRA [24] был ориентирован на задачи атмосферной оптики, пассивное зондирование газового состава атмосферы, оптической климатологии. В нем использовались упрощенный расчет коэффициентов поглощения (модельное представление полос поглощения) и региональные модели атмосферы [22]. Созданный пакет программ PARSEG был ориентирован на моделирование и обработку данных дистанционного зондирования газового состава атмосферы [23].

С начала 2000-х гг. в Интернет-доступной ИС по атмосферной радиации появились радиационные модели из других организаций [32, 33]. В последние годы при создании ИС по атмосферной радиации так же, как и в спектроскопии, стала преобладающей тенденция интеграции информационно-вычислительных ресурсов. Реализацией этой тенденции является создание распределенной информационной системы по атмосферной радиации [15, 16], имеющей узлы в Томске, Волгограде и Екатеринбурге. В основу этой системы положено промежуточное программное обеспечение, созданное в ИОА СО РАН при выполнении проекта ИНТАС «ATMOS:

a scientific WWW portal for the atmospheric environment».

Заключение

Информационные системы в области оптики атмосферы были одним из направлений в прикладной информатике, которое академик В.Е. Зуев выделил и поддерживал с начала 80-х гг. XX в. Эта поддержка позволила занять ИОА СО РАН лидирующее место в разработке и создании ИС, ориентированных на обеспечение решения информационных задач оптики атмосферы. Современное развитие информационных систем в ИОА СО РАН обеспечено поддержкой специалистов в области спектроскопии, радиации и атмосферного аэрозоля.

1. Guarino N. Formal Ontology and Information Systems // Proc. FOIS'98. Trento, Italy, 6–8 June 1998. Amsterdam: IOS Press, 1998. P. 3–15.
2. Войцеховская О.К., Розина А.В., Трифонова Н.Н. Информационная система по спектроскопии высокого разрешения. Новосибирск: Наука, 1988. 150 с.
3. Войцеховская О.К., Зуев В.Е., Тюттерев Вл.Г. Информационные системы молекулярной спектроскопии атмосферы // Оптика атмосф. 1988. Т. 1. № 3. С. 3–15.
4. Тюттерев Вл.Г. Глобальные вариационные и эффективные методы расчетов положений и интенсивностей спектральных линий трехатомных молекул: некоторые тенденции и особенности нового поколения спектроскопических информационных систем // Оптика атмосф. и океана. 2003. Т. 16. № 3. С. 245–255.
5. Быков А.Д., Науменко О.В., Родимова О.Б., Синеца Л.Н., Творогов С.Д., Тонков М.В., Фазлиев А.З., Филиппов Н.Н. Информационные аспекты молекулярной спектроскопии. Томск: Изд-во ИОА СО РАН, 2008. 356 с.

6. Gordov E.P., Lykosov V.N., Fazliev A.Z. Web portal on environmental sciences «ATMOS» // Adv. Geosci. 2006. V. 8. P. 33–38.
7. Войцеховская О.К., Кузнецов С.В., Сапожников С.В., Трифонова Н.Н., Черкасов М.Р. Информационная система по молекулярному поглощению излучения CO₂-лазера // Оптика атмосфер. 1991. Т. 4. № 9. С. 938–953.
8. Михайленко С.Н., Бабиков Ю.Л., Тютерев Вл.Г., Варбе А. Банк данных по спектроскопии озона доступный в Интернете (S&MPO) // Вычислительные технологии. 2002. Т. 7. Спец. выпуск. С. 64–70.
9. Golovko V.F., Nikitin A.V., Chursin A.A., Tyuterev V.I.G. Information system AIRSENTRY for modeling atmospheric IR-spectra and radiation transmission in the atmosphere // Proc. 2nd Int. Workshop AD-BIS'95. V. 2. Moscow, 1995. P. 12–14.
10. Михайленко С.Н., Ташкун С.А., Бабиков Ю.Л., Головкин В.Ф., Тютерев Вл.Г. Прямая спектроскопическая задача в рамках ИВС «Спектроскопия атмосферных газов» // Оптика атмосфер. и океана. 2004. Т. 17. № 11. С. 927–938.
11. Михайленко С.Н., Бабиков Ю.Л., Головкин В.Ф. Информационно-вычислительная система «Спектроскопия атмосферных газов». Структура и основные функции // Оптика атмосфер. и океана. 2005. Т. 18. № 9. С. 765–776.
12. Быков А.Д., Воронин Б.А., Козодоев А.В., Лаврентьев Н.А., Родимова О.Б., Фазлиев А.З. Информационная система для решения задач молекулярной спектроскопии. 1. Структура информационных ресурсов // Оптика атмосфер. и океана. 2004. Т. 17. № 11. С. 921–926.
13. Козодоев А.В., Привезенцев А.И., Фазлиев А.З. Информационная система для решения задач молекулярной спектроскопии. 3. Уровни энергии молекул // Оптика атмосфер. и океана. 2007. Т. 20. № 9. С. 805–809.
14. Лаврентьев Н.А., Привезенцев А.И., Фазлиев А.З. Информационная система для решения задач молекулярной спектроскопии. 4. Переходы в молекулах симметрии C_{2v} и C_s // Оптика атмосфер. и океана. 2008. Т. 21. № 11. С. 957–963.
15. Фирсов К.М., Фазлиев А.З., Сакерин С.М., Журавлева Т.Б., Фомин Б.А., Захаров В.И. Информационно-вычислительная система «Атмосферная радиация». Современное состояние, перспективы развития // Тр. 9-й Всерос. науч. конф. «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» – RCDL'2007. Переславль-Залесский, 2007. Ч. 1. С. 62–66.
16. Фирсов К.М., Фазлиев А.З., Чеснокова Т.Ю., Козодоева Е.М. Распределенная информационно-вычислительная система «Атмосферная радиация» // Тр. 11-й Всерос. науч. конф. «Электронные библиотеки: перспективные методы и технологии». Петрозаводск, 2009.
17. Войцеховская О.К., Кузнецов С.В., Остапенко С.П., Черкасов М.Р. База данных по параметрам спектральных линий оксидов азота и углерода в диапазоне температур 300–3000 К // Оптика атмосфер. и океана. 1992. Т. 5. № 9. С. 962–969.
18. Быков А.Д., Пономарев Ю.Н., Сеница Л.Н. База данных по коэффициентам сдвига и уширения линий поглощения H₂O в ИК- и видимой областях спектра // Оптика атмосфер. и океана. 1992. Т. 5. № 9. С. 931–938.
19. Науменко О.В., Половцева Е.Р. База данных по поглощению сероводорода в области 4400–11400 см⁻¹ // Оптика атмосфер. и океана. 2003. Т. 16. № 11. С. 985–991.
20. Tashkun S.A., Perevalov V.I., Teffo J.-L., Bykov A.D., Lavrentieva N.N. CDS-1000, the high-temperature carbon dioxide spectroscopic databank // J. Quant. Spectrosc. and Radiat. Transfer. 2003. V. 82. Issue 1. P. 165–196.
21. Журавлева Т.Б., Анисимова Е.М. База данных «Численная радиационная модель однослойной разорванной облачности», доступная в сети Интернет // Оптика атмосфер. и океана. 2002. Т. 15. № 8. С. 699–701.
22. Зуев В.Е., Комаров В.С. Статистические модели температуры и газовых компонент атмосферы. Л.: Гидрометеониздат, 1986. 264 с.
23. Комаров В.С., Креминский А.В., Синева К.Я. Компьютерная информационная база региональных климатических моделей температуры и ветра для пограничного слоя атмосферы // Оптика атмосфер. и океана. 1996. Т. 9. № 4. С. 484–488.
24. Афонин С.В., Гендрин А.Г. Информационно-программное обеспечение моделирования переноса ИК-радиации в атмосфере // Информационно-программное обеспечение задач атмосферной оптики. Новосибирск: Наука, 1988. С. 38–65.
25. Головкин В.Ф., Поздняков А.А., Тютерев Вл.Г., Чурсин А.А. Графическое программное обеспечение для информационной спектроскопической компьютерной системы // Оптика атмосфер. и океана. 1992. Т. 5. № 10. С. 1088–1095.
26. Войцеховская О.К., Кузнецов С.В. Программная оболочка открытой многофункциональной информационной системы по спектроскопии высокого разрешения для IBM-совместимых компьютеров // Оптика атмосфер. и океана. 1994. Т. 7. № 7. С. 998–1005.
27. Катаев М.Ю., Мицель А.А., Пташник И.В., Фирсов К.М. Методы и системы автоматизации. Обработка данных дистанционного зондирования. Пакет программ для моделирования и обработки данных дистанционного зондирования газового состава атмосферы // Оптика атмосфер. и океана. 1996. Т. 9. № 10. С. 1412–1417.
28. Комаров В.С., Мицель А.А., Михайлов С.А., Пономарев Ю.Н., Руденко В.П., Фирсов К.М. Программное и информационное обеспечение задач оптики атмосферы // Оптика атмосфер. 1988. Т. 1. № 5. С. 84–89.
29. De Roure D., Jennings N., Shadbolt N. A Future e-Science Infrastructure. Report commissioned for EPSRC/DTI Core e-Science Programme. 2001. 783 p.
30. Зуев В.Е., Креков Г.М. Оптические модели атмосферы. Л.: Гидрометеониздат, 1986. 256 с.
31. Несмелова Л.И., Родимова О.Б., Творогов С.Д. Контуры спектральной линии и межмолекулярное взаимодействие. Новосибирск: Наука, 1986. 216 с.
32. Алексеев В.А., Володин Е.М., Галин В.Я., Дымников В.П., Лысков В.Н. Моделирование современного климата с помощью атмосферной модели ИВМ РАН. М.: ВИНТИ, 1998. 120 с.
33. Фролькис В.А., Киселев А.А., Кароль И.Л. Диагностика антропогенных изменений климата моделированием радиационно-фотохимических процессов в атмосфере // Изв. РАН. Физ. атмосфер. и океана. 1999. Т. 35. № 4. С. 444–456.

A.Z. Fazliev. Development of the information systems at Institute of Atmospheric Optics SB RAS.

A brief review of information systems development during the last 25 years is presented. The evolution of molecular spectroscopy information systems is described from different points of view: structural, functional, and technical implementation.