

# СИСТЕМА ПРОВЕРКИ ОГРАНИЧЕНИЙ СУЩЕСТВОВАНИЯ ДЛЯ МОЛЕКУЛЯНОЙ СПЕКТРОСКОПИИ

Н.А. Лаврентьев, М.М. Макогон, А.З.Фазлиев

*Институт оптики атмосферы имени В.Е. Зуева СО РАН (Томск), Россия  
lnick@iao.ru*

## **Аннотация**

В докладе представлено обоснование необходимости применения ограничения существования спектральных данных по отношению к экспертным массивам. Применение этого ограничения возможно только в том случае, когда декомпозиция проводится по полному набору данных из всех опубликованных работ в предметной области, относящимся к решениям задач по определению значений исследуемой физической величины.

На примере экспертных массивов спектральных данных, используемых в задачах, входными данными которых являются параметры спектральных линий (задачи астрономии, оптики атмосферы, переноса излучения в атмосфере и т.д.) показано, что имеет место включение в эти массивы значений данных, которые не были опубликованы и достоверность их не документирована.

## **1. ВВЕДЕНИЕ**

Пользователь, использующий на практике экспертные массивы спектральных данных, поставлен перед проблемой выбора: какой из массивов ему выбрать для того, что проводимый им расчет точнее совпадал с результатами измерений. В исследовании потоков приходящего и уходящего излучения в атмосфере влияние достоверности большей части данных слабо связано с корректностью отнесения переходов к квантовым числам,  $a$ , при расчете интегральных потоков, и с точностью значительного числа слабых линий. Однако при расчете потоков излучения в узком спектральном диапазоне существенным становится наличие некорректности в исходном массиве спектральных данных.

С какими ограничениями к массиву спектральных данных должен подходить пользователь? Есть две группы ограничений: ограничения на значения физических величин входящих в экспертные массивы и ограничение на существование спектральных данных.

Первая группа ограничений включает в себя проверку типов данных, к которым относятся значения физических величин (например, большая часть квантовых чисел должна быть целыми неотрицательными числами), интервалов изменения физических величин (ограничителем может являть энергия ионизации молекулы), правила отбора, и ряд других ограничений связанных с базовыми математическими отношениями (транзитивности, рефлексивности и т.д.). Как показывают проверки публикаций,

существует небольшой процент ошибок связанных с правилами отбора, тогда как ошибки, связанные с нарушениями отношения рефлексивности, составляют уже на порядок большее число. Эти ограничения проверяются при публикации оригинальных работ.

Вторая группа ограничений связана с публикациями работ, в которых авторы предлагают отобранные по некоторым критериям, как правило, в явном виде не описанным, составные массивы спектральных данных. Эти массивы, содержащие несколько десятков миллионов значений физических величин, снабжаются библиографией, относящейся почти к каждому из значений такого массива. Проверить соответствие библиографии и значений, размещенных в экспертном массиве, не используя компьютерные приложения, практически невозможно.

Отметим, что в последние несколько лет выполняется проект IUPAC одной из целей которого является систематизация всех опубликованных характеристик переходов молекулы воды и ее изотопмеров [1]. В рамках проекта собраны практически все публикации, относящиеся к спектральным характеристикам молекулы воды [2,3]. Наличие этого набора публикаций, данные из которых размещены в информационной системе W@DIS, позволило ответить на вопрос, какие из частей существующих экспертных массивов, относящиеся к изотопмерам молекулы воды, содержат опубликованные данные, а какие содержат данные ранее не публиковавшиеся.

В качестве исследуемых массивов были выбраны части массивов спектральных данных Nitran [4] и Geisa [5], содержащие сведения об изотопмерах  $\text{H}_2^{17}\text{O}$ ,  $\text{H}_2^{18}\text{O}$ ,  $\text{HDO}$ ,  $\text{HD}^{17}\text{O}$ ,  $\text{HD}^{18}\text{O}$ .

Для решения задачи декомпозиции экспертного массива данных на части, содержащие опубликованные и неопубликованные данные, в информационной системе W@DIS (<http://wadis.saga.iao.ru>) был создан сервис, проводящий такое разложение в соответствие с ограничениями на допустимую величину отклонения сравниваемых величин.

## 2. ДОСТОВЕРНОСТЬ СПЕКТРАЛЬНЫХ ДАННЫХ

Проверка достоверности данных в предметных областях естественных наук является одной из важнейших задач. Понятие достоверность трактуется при этом разными способами. Однако тем, что объединяет эти трактовки, является общий принцип, а именно, поскольку при описании свойств объектов используются их математические модели, в основании которых лежат базовые отношения математики (такие как рефлексивность, транзитивность и т.д.), то результаты, получаемые при использовании этих моделей, должны удовлетворять ограничениям, накладываемым ими. В спектроскопии к числу наиболее известных ограничений относятся правила отбора, определяющие отношения между квантовыми числами. Проверка таких ограничений тесно связана с проверкой декларируемого утверждения. Например, если  $\{x\}$  – набор квантовых чисел, то он может описывать состояние (переход) только в том случае, если их значения удовлетворяют правилам отбора. Проверки такого типа ограничений на значения физических величин назовем проверками на ограничения значений. Подавляющее большинство опубликованных спектральных данных удовлетворяет таким ограничениям.

Существует другой тип ограничений, связанный с проверкой существования значений физических величин. Здесь термин «существование» трактуется как публикация значений физических величин в журналах, сети Интернет, диссертациях и иных средствах коммуникаций, перечень которых определяется научным сообществом, исследующим ту или иную предметную область. Значимость этого типа ограничений возрастает при автоматизированной проверке достоверности данных требующей представления данных в явном виде, допускающем компьютерную обработку.

На практике проверки такого типа важны в силу разных причин. Во-первых, при решении задач спектроскопии важны условия, при которых проводились измерения, или приближения, используемые при вычислениях. Табличное представление спектральных данных использует библиографические ссылки на публикации, в которых описаны условия экспериментов или приближения, при которых построены модели. Автоматизированная обработка таких ссылок в настоящее время не представляется возможной. Более того в некоторых работах данные публикуются без подобных описаний. Последнее утверждение относится в первую очередь к тем работам, в которых публикуются синтетические массивы спектральных данных.

В данной статье представлены результаты проверки ограничения существования на примере двух массивов спектральных данных HITRAN [3] и GEISA [4]. Эти массивы позиционируются как спектральные данные для применения в задачах атмосферной радиации, астрономии и атмосферной оптики. Они иногда использовались и для решения внутренних задач спектроскопии, например, определения уровней энергии молекул. Анализ данных позволил ответить на вопрос, что общего между этими массивами данных, чем они качественно отличаются и почему их использование для решений внутренних задач спектроскопии возможно лишь при тщательной проверке накопленных в этих массивах данных.

### **3. ФИЗИЧЕСКИЕ ОСНОВАНИЯ ДЛЯ СРАВНЕНИЯ**

При проведении массовых спектроскопических измерений положения центров линий поглощения и излучения газов и паров в настоящее время принята следующая схема редукции и анализа данных:

1. из экспериментальных измерений определяются центры линий;
2. линеаризуется развертка частотной шкалы;
3. осуществляется абсолютная «привязка» частотной шкалы по положению точно определенных линий некоторых веществ, выполняющих, по существу, роль вторичных эталонов.

Критерием надежности (точности, воспроизводимости в эксперименте) данных служит степень совпадения с данными расчетов или других измерений. Для корректного сопоставления разных данных необходим учет зависимости частот центров линий от температуры, давления и состава смеси анализируемых газов. Такая схема применяется при любых измерениях; вместе с тем каждый метод измерений вследствие различных физических принципов, положенных в его основу, накладывает на нее свой специфический «отпечаток».

Наиболее часто в последние 10-15 лет для определения центров линий поглощения применяются спектрофотометрические методы, метод внутривибрационной лазерной спектроскопии и метод «*ging down spectroscopy*». По спектрофотометрическому методу измерения проводят с помощью Фурье-спектрометров (в таком приборе источник и приемник излучения «объединены») и спектрометров, в которых источником излучения является перестраиваемый лазер или лазерная система. Центры линий излучения измеряют с помощью Фурье-спектрометров.

### ***1. Определение центров линий***

Для нахождения центра линии измеренный спектр аппроксимируется кривой, описывающей форму контура (доплеровской, лоренцевской, фойгтовской или более сложной) в условиях конкретного эксперимента, и положение максимума кривой принимается за центр линии. Погрешность «измеренного» таким образом положения центра будет зависеть от соотношения между спектральным разрешением используемого прибора и шириной анализируемой линии, отношения сигнал/шум при записи спектра, правильности выбора формы контура, присутствия в области анализируемого (подгоняемого) спектра других, даже существенно более слабых линий, наклона нулевой линии и некоторых других параметров. Если имеется несколько близких линий, расположенных так, что спектрального разрешения прибора недостаточно для их разделения, то найденное по этой процедуре положение «центра линии» даст какое-то усредненное значение для всей группы линий. Оценить рассматриваемую погрешность в общем случае не представляется возможным.

Вместе с этим практика показывает, что спектроскопические измерения проводятся, как правило, с разрешением на порядок лучшим, чем ширина линии, а центры находятся с точностью, которая на 1-2 порядка лучше, чем спектральное разрешение. Таким образом, погрешность определения центров линий поглощения газов в атмосферных условиях для видимого и близким к нему диапазонах составляет около  $10^{-3}$ - $10^{-4}$  см<sup>-1</sup>. При продвижении в ИК и микроволновую область спектра ширины линий уменьшаются, и вместе с ними уменьшается абсолютная величина этой погрешности.

### ***2. Линеаризация развертки частотной шкалы***

Нелинейность развертки частотной шкалы имеет разные физические причины. В дифракционных спектрометрах, широко используемых для регистрации ВРЛ-спектров, она вызывается несовпадением фокальной поверхности с плоскостью регистратора (фотопластинка, ПЗС-линейка). В лазерных спектрометрах на перестраиваемых диодных лазерах – нелинейной зависимостью частоты излучения от тока накачки. В спектрометрах с лазерами на красителях эта нелинейность обусловлена сложной зависимостью частоты излучения от положения перестроечных элементов.

### ***3. Абсолютная привязка частотной шкалы***

Приборы, используемые в измерениях положения линий, как правило, не имеют абсолютной привязки частотной шкалы. Только в спектрофотометрах терагерцового диапазона частота излучения, формируемого стабилизированными СО<sub>2</sub>-лазерами и микроволновыми генераторами, известна абсолютно. Определенным исключением

являются также Фурье-спектрометры, в составе которых имеются стабилизированные лазеры. Однако поскольку положения восстановленных из записанной интерферограммы линий зависят от настройки прибора, диаметров установленных в нем диафрагм, ширины аппаратной функции и ряда других параметров, то абсолютная погрешность их определения оказывается на уровне  $10^{-3} \text{ см}^{-1}$ , хотя длина волны лазера известна с погрешностью порядка  $10^{-5} \text{ см}^{-1}$ .

Поэтому в ходе эксперимента при одной и той же настройке аппаратуры дополнительно записываются спектры веществ, положение отдельных линий которых принимается за эталонное.

В таблице приведены данные о характерных (типичных) измерениях положения центров линий водяного пара, выполненных в последние 10-15 лет. Поскольку погрешность определения центров линий в отдельных работах очень сильно варьируется, то в этих случаях приведены округленные значения границ диапазона изменения, и надо иметь в виду, что подавляющее большинство центров определено с погрешностью в середине этого диапазона.

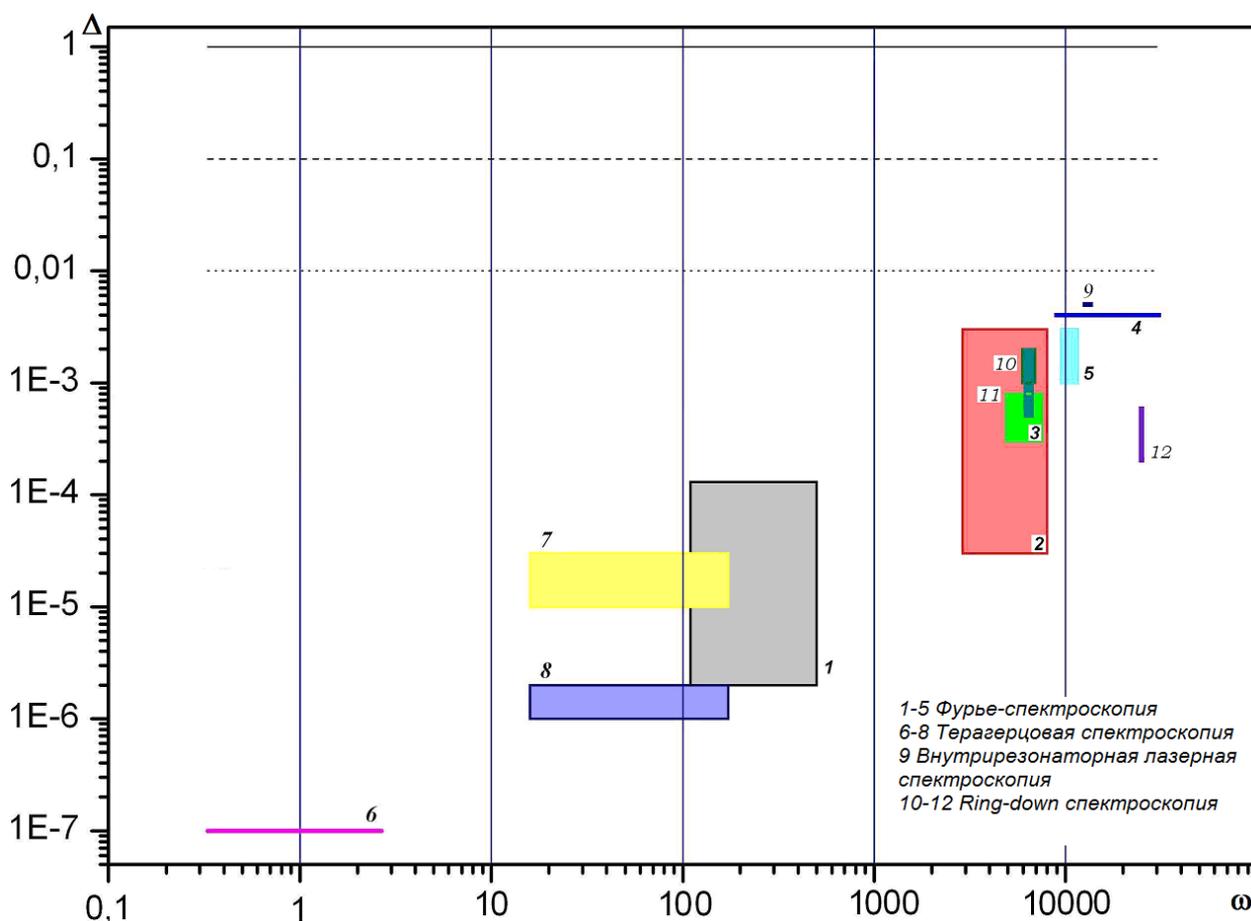


Рис.1. Погрешность определения центров линий в ряде публикаций из таблицы 1.

Прямоугольные области иллюстрируют спектральный диапазон измерений и область вариации погрешностей; в ряде случаев эти области в масштабе графика вырождаются в прямые линии. Метод измерений отображается шрифтом цифр.  $\Delta$  - погрешности определения центров спектральных линий ( $\text{см}^{-1}$ ,  $\omega$  - вакуумные волновые числа ( $\text{см}^{-1}$ )).

Таблица 1.

Метод измерений	Спектральный диапазон, см <sup>-1</sup>	$\delta$ (см <sup>-1</sup> )	Линеаризация	Привязка шкалы по веществу	$\Delta$ (10 <sup>-3</sup> см <sup>-1</sup> )	Литература	Номер области
Спектрофотометрический							
Фурье-спектрометрия	20-12000	0,0056			0,01-10		
	110-500	0,0025		OCS	0,002-0,13	[27]	<b>1</b>
	2900-8000	0,011		NO <sub>2</sub> , CO	0,03-3	[16]	<b>2</b>
	4880-7550	0,02			0,3-0,8	[10]	<b>3</b>
	8600-15000	0,03				[8]	
	9000-30000	0,06		J <sub>2</sub>	4	[29]	<b>4</b>
	9250-13000	0,03 и 0,06	3			[12]	
	9500-11500	0,01-0,02			1-3	[9]	<b>5</b>
	13000-26000	0,06		J <sub>2</sub>		[11]	
Лазерная спектроскопия терагерцового диапазона	0,33-2,67 (10-80 ГГц)				10 <sup>-4</sup> (3 кГц)	[24]	<b>6</b>
	3,33-20(100-600 ГГц)					[25]	
	16-173 (0,5-5 ТГц)	35 кГц			0,01-0,03 0,0013 (40 кГц)	[20] [23]	<b>7</b>
	16-173 (0,5-5 ТГц)				0,001-0,002	[26], [29]	<b>8</b>
ВРЛС	11335-12843	0,015	есть	H <sub>2</sub> O	4	[18]	
	12746-13558	F(v)			5	[22]	<b>9</b>
Ринг-даун спектроскопия	6025-6870	0,0001	есть	H <sub>2</sub> O	1-2	[21]	<b>10</b>
	6130-6750		есть	H <sub>2</sub> O	0,5-2		<b>11</b>
	13312-13377		есть	J <sub>2</sub>	5	[15]	
	16570-17125	0,06				[19]	
	16600-18000	0,05		J <sub>2</sub>	4	[7]	
	25195-25340		есть	J <sub>2</sub>	0,2-0,6	[14]	<b>12</b>
Фурье-спектрометрия	50-500	0,005			0,9	[13]	
	373-933	0,01		H <sub>2</sub> O		[6]	
	500-13000	0,03 – 0,05		H <sub>2</sub> O		[17]	
Измеряемая величина			$\delta$ - спектральное разрешение, $\Delta$ - погрешность определения центров линий, F(v)- ограничено доплеровским уширением				
Ослабление		Излучение					

#### 4. БАНКИ ДАННЫХ В АТМОСФЕРНЫХ ПРИЛОЖЕНИЯХ

Спектральные данные в молекулярной спектроскопии собираются и классифицируются разными группами исследователей. В прикладных атмосферных науках используется только часть этих данных, связанная с молекулами, участвующими в процессах поглощения солнечной радиации. Наиболее используемыми среди таких массивов данных являются массивы HITRAN и GEISA. Применяемые на практике версии этих массивов для ряда прикладных задач дают удовлетворительные результаты.

Процедура составления этих массивов является непрозрачной, хотя значение каждой величины, например, в массиве HITRAN, связано с публикацией. К сожалению, среди этих ссылок присутствуют ссылки на частные сообщения исследователей авторам массивов данных и ссылки на другие составные массивы данных, вплоть до циклических ссылок. Проверить корректность ссылок практически невозможно. Это обстоятельство объясняет необходимость постановки задачи автоматического разложения экспертного массива данных по наборам данных из первичных публикаций, содержащих результаты экспериментальных и теоретических работ. Результаты такого разложения экспертных массивов можно сравнивать между собой. В первую очередь, такое сравнение позволит выяснить какие части двух разных экспертных массивов совпадают, какие различаются и в чем состоят различия. Знание об этих свойствах экспертных массивов может служить основанием, с одной стороны, для выбора пользователем требуемого ему по качеству набора данных, а с другой стороны, для постановки экспериментов для проверки недостоверных характеристик переходов.

Выше были описаны спектральные интервалы и точности измерения в них. Ниже приведена Таб.2 в которой для ряда изотопомеров воды представлено число переходов в этих спектральных интервалах, число переходов с некорректными квантовыми числами и число переходов не найденных в массиве опубликованных данных (процент который это число составляет относительно общего количества переходов) собранных в ИС W@DIS. В этой ИС размещены опубликованные данные, относящиеся к обработке измеренных величин. Эти данные собраны и прокомментированы в [2, 3]. Данные теоретических расчетов взяты из публикаций последних 14 лет.

**Таблица 2.** Сравнение количества переходов по интервалам частот и число частот, не удовлетворяющих ограничению существования для массивов данных HITRAN & GEISA .

	Интервалы значений вакуумных волновых чисел (см <sup>-1</sup> )						0-20000 см <sup>-1</sup>		
	0-10	10-30	30-50	50-200	200-10000	10000-20000	0-20000	HITRAN	GEISA
<b>H<sub>2</sub><sup>17</sup>O</b>	0	10	25	279	4762	1916	6992 (18,0,0)	<i>403</i>	<i>369</i>
	1	10	25	279	4828	325	5468 (0,8,0)	6%	7%
<b>H<sub>2</sub><sup>18</sup>O</b>	31	15	33	342	6889	2443	9753 (1,6,0)	<i>710</i>	<i>738</i>
		15	33	342	7106	863	8360 (0,20,0)	7%	9%
<b>HDO</b>	65	76	76	599	8825	3483	13238 (2,0,0)	<i>44</i>	<i>128</i>
	66	76	76	599	11045	119	11981(121,88,31)	0%	1%
<b>HD<sup>17</sup>O</b>	0	0	0	0	175	0	175	<i>0</i>	<i>0</i>
	0	0	0	0	175	0	175	0%	0%
<b>HD<sup>18</sup>O</b>	53	61	45	544	908	0	1611	<i>189</i>	<i>154</i>
	0	0	0	0	669	0	659	12%	23%

В части таблицы, представляющей распределение данных по интервалам значений вакуумных волновых чисел, для каждого изотопомера, серым цветом отмечено количество переходов в массиве данных GEISA. Не помеченные цветом поля таблицы относятся к массиву данных HITRAN.

Отметим, что в столбце, характеризующем количество переходов во всем интервале (0-20000), представлены число переходов без определенных значений квантовых чисел, число переходов с совпадающими квантовыми числами и число переходов не удовлетворяющих правилам отбора.

Вторая часть таблицы представляет результаты разложения двух экспертных массивов данных по набору первичных хранящихся в ИС W@DIS. Для каждого вещества приведено число переходов из экспертного массива не обнаруженных в оригинальных публикациях, представляющих результаты вариационных расчетов и обработки измерений. Ниже указано какой процент составляет это число недокументированных переходов относительно общего числа переходов. Более детальные результаты анализа представлены в разделе 6. В Приложении для всех интервалов указаны ссылки на работы, в которых опубликованы данные, удовлетворяющие критериям заданным при анализе.

Для понимания деталей результатов последующего анализа обратимся к описанию технологии декомпозиции экспертных массивов, использованной в ИС W@DIS.

## **5. ПРОВЕРКА ЭКСПЕРТНОГО МАССИВА ПО КРИТЕРИЮ СУЩЕСТВОВАНИЯ**

Для проведения проверки существования данных собраны публикации за период более 70 лет по ряду молекул ( $H_2O$ ,  $H_2S$ ,  $CO_2$ ). Часть данных о вакуумных волновых числах воды была обработана группой данных IUPAC [2,3] и выделены данные, содержащие не корректные отнесения квантовых чисел. Эти скорректированные решения задач используются ниже при проверке ограничения существования информационных ресурсов об изотопомерах молекулы воды в массиве данных HITRAN и GEISA. Критерий для проверки ограничения существования, используемый в спектроскопии, основан на допущении, что проверяемые данные должны отличаться от опубликованных данных не более чем на величину ошибок измерений, характерных для соответствующего диапазона вакуумных волновых чисел.

При проведении работ по систематизации спектральных данных по изотопомерам воды создана информационная система W@DIS в которой аккумулированы все первичные данные, собранные в рамках проекта IUPAC.

**Поиск источников информации**

Выбор задачи:  Уровни энергии  Профили линий  Переходы

Вещество: H2O

Диапазон вакуумных частот (см<sup>-1</sup>): 0 - 30000

Слова для поиска источников данных по контексту, содержащемуся в аннотации или ссылке на публикацию. (Фамилии авторов публикаций, журнал, год публикации, слова из названий статьи)

Искать источники информации

Выбрать источник информации

Показать 40 строк от 0 Всего строк 115 > >> Настройки

Выбор	Название Вычисления/Эксперимент	Число записей [z-a]	Публикация
<input type="radio"/>	2006_BaTeHaTo_H2O	1791456	R.J.Barber, J. Tennyson, G.J. Harris, R.N. Tolchenov, A high accuracy computed water line list - BT2. // Mon. Not. R. Astron. Soc., 2006, Volume 368, Pages 1087-1094, DOI: 10.1111/j.1365-2966.2006.10184.x.
<input type="radio"/>	2007_ScPaTa_b_H2O	1027079	Schwenke D.W., H. Partridge, Tashkun S.A., Schwenke-Partridge linelists (PS-2007-1000) for H <sub>2</sub> <sup>16</sup> O. <a href="http://spectra.iao.ru">http://spectra.iao.ru</a>

Рис.2. Интерфейс для выбора источника данных с целью его декомпозиции.

Собранные первичные данные были отнесены к публикации из которой они извлечены и размещены в базах данных ИС W@DIS. Наряду с первичными данными в системе размещены составные (экспертные) массивы HITRAN и GEISA.

Для пользователя процедура декомпозиции массива данных состоит из следующей последовательности действий:

1. выбора данных для последующей декомпозиции (см. рис.2)
2. задание допустимой точности измерения ( $\delta$ , см. Таб.3) (см. Рис.3)
3. задание интервала вакуумных волновых чисел
4. указание на то по какой части первичных данных (расчетной или измерительной) проводить декомпозицию просмотра переходов с одинаковыми квантовыми числами, удовлетворяющими ограничению  $\delta$ .
5. и просмотр остатка, т.е. переходов не удовлетворяющих ограничению  $\delta$ .

<b>2009_RoGoBaBe_H2O</b>	L.S. Rothman, I.E. Gordon, A. Barbe, D.Chris Benner, P.F. Bernath, M. Birk, V. Boudon, L.R. Brown, A. Campargue, J.-P. Champion, K. Chance, L.H. Coudert, V. Dana, V.M. Devi, S. Fally, J.-M. Flaud, R.R. Gamache, A. Goldman, etc, The HITRAN 2008 molecular spectroscopic database. // Journal of Quantitative Spectroscopy and Radiative Transfer, 2009, Volume 110, Issue 9, Pages 533-572, DOI: 10.1016/j.jqsrt.2009.02.013.
--------------------------	---

<input checked="" type="checkbox"/> Ошибка определения частоты (см <sup>-1</sup> )	1
<input checked="" type="checkbox"/> Ограничения на частоту (см <sup>-1</sup> )	0 - 20000
<input checked="" type="checkbox"/> Декомпозиция по эксперименту	
<input type="checkbox"/> Декомпозиция по расчету	
Провести декомпозицию	

Показать 40 строк от 0	Всего строк 82	> >>	Настройки
------------------------	----------------	------	-----------

Название Вычисления/Эксперимент	Число записей	Число совпадений [z-a]	Показать	Публикация
<b>2005_ToNaZoSh_H2O</b>	16990	11922	Показать	Tolchenov R.N., O.Naumenko, N.F.Zobov, S.V.Shirin, O.L.Polyansky, J.Tennyson, M.Carleer, P.-F.Coheur, S.Fally, A.Jenouvrier and A.C.Vandaele, Water vapour line assignments in the 9250-26 000 cm <sup>-1</sup> frequency range. // Journal of Molecular Spectroscopy, 2005, v. 233, no. 1, p. 68-76.

Рис.3. Интерфейс для задания условий декомпозиции и представления ее результатов.

При просмотре результатов декомпозиции (см. Рис.2) для каждого первичного источника данных указывается число переходов, удовлетворяющих заданным ограничениям, и в колонке «Показать» размещена ссылка на таблицы, в которых размещены переходы из первичного источника данных, соответствующие указанным в интерфейсе ограничениям.

Information source left								Information source right							
<b>2009_RoGoBaBe_H2O</b>								<b>2005_ToNaZoSh_H2O</b>							
L.S. Rothman, I.E. Gordon, A. Barbe, D.Chris Benner, P.F. Bernath, M. Birk, V. Boudon, L.R. Brown, A. Campargue, J.-P. Champion, K. Chance, L.H. Coudert, V. Dana, V.M. Devi, S. Fally, J.-M. Flaud, R.R. Gamache, A. Goldman, etc, The HITRAN 2008 molecular spectroscopic database. // Journal of Quantitative Spectroscopy and Radiative Transfer, 2009, Volume 110, Issue 9, Pages 533-572								Tolchenov R.N., O.Naumenko, N.F.Zobov, S.V.Shirin, O.L.Polyansky, J.Tennyson, M.Carleer, P.-F.Coheur, S.Fally, A.Jenouvrier and A.C.Vandaele, Water vapour line assignments in the 9250-26 000 cm <sup>-1</sup> frequency range. // Journal of Molecular Spectroscopy, 2005, v. 233, no. 1, p. 68-76.							
Show 40 rows starting from 0								In all rows 11922							
								Control							
$\omega_1$ (cm <sup>-1</sup> )	$\nu_1^{up}$	$\nu_2^{up}$	$\nu_3^{up}$	$J^{up}$	$K_a^{up}$	$K_c^{up}$	$\nu_1^{low}$	$\nu_2^{low}$	$\nu_3^{low}$	$J^{low}$	$K_a^{low}$	$K_c^{low}$	$\omega_2$ (cm <sup>-1</sup> )	$\Delta\omega$ (cm <sup>-1</sup> )	
9251.2391	1	1	1	9	4	6	0	0	0	8	2	7	9251.2408	1.7000e-3	
9253.5825	0	1	2	6	4	2	0	0	0	5	3	3	9253.5841	1.6000e-3	
9256.5636	2	1	0	7	5	2	0	0	0	7	0	7	9256.5661	2.5000e-3	
9257.0917	0	1	2	5	5	0	0	0	0	4	4	1	9257.0921	4.0000e-4	
9257.0917	0	1	2	5	5	1	0	0	0	4	4	0	9257.0624	2.9300e-2	

Рис.4. Сравнение вакуумных волновых чисел массива HITRAN и одного из источников данных, удовлетворяющих ограничениям пользователя.

На рис.4. показано сравнение удовлетворяющих критериям пользователя переходов.

## 6. Сравнение массивов HITRAN и GEISA по результатам декомпозиции

В таблицах 3а и 3б приведены результаты декомпозиции экспертных массивов данных. В таб.3а представлены результаты декомпозиции в трех интервалах от 0 до 50 см<sup>-1</sup>, а в таб.3б – в двух интервалах от 5 до 20000 см<sup>-1</sup>. В таблице каждому изотопимеру

соответствуют две строки. В верхней строке представлены результаты декомпозиции экспертного массива по публикациям собранным группой данных IUPAC, а в нижней строке – по вариационным расчетам. Отношение  $n/m$  содержит  $n$  - число переходов не удовлетворяющих критериям сравнения заданного при декомпозиции, а  $m$  – число переходов экспертного массива в данном интервале. Какой процент переходов экспертного массива не удовлетворяет критериям сравнения, заданным при декомпозиции, показан в нижней части строки. Точность, с которой проводилось сравнение данных измерений, в Таб. 3а и 3б задана шрифтом. Подчеркнутые цифры означают, что сравнение проведено с точностью  $\delta = 0.01 \text{ см}^{-1}$ , цифры, выделенные курсивом -  $\delta = 0.1 \text{ см}^{-1}$ , и цифры, выделенные жирным шрифтом –  $\delta = 1 \text{ см}^{-1}$ .

Как следует из количественного сравнения массивов HITRAN и GEISA в диапазоне 0-200  $\text{см}^{-1}$  число переходов в них практически одинаково и число опубликованных переходов также одинаково. Исключением является диапазон 10-30  $\text{см}^{-1}$  в котором для молекулы HDO данные GEISA лучше соответствуют набору данных IUPAC.

Как видно из анализа обеих Таб. (3а и 3б) в диапазоне 0-10000  $\text{см}^{-1}$  процент неопубликованных данных в Hitran и GEISA в случае декомпозиции по данным вариационных расчетов ниже по сравнению результатом декомпозиции по данным измерений. В диапазоне 10000-20000  $\text{см}^{-1}$  тенденция обратная, а именно, число переходов взятых из экспериментальных работ в экспертные массивы существенно больше числа переходов из вариационных расчетов, не подтвержденных измерениями. Этот факт, по-видимому, указывает на некорректное приписывание квантовых чисел в этом диапазоне при вариационных расчетах.

**Таблица 3а. Сравнение массивов спектральных данных Hitran и GEISA в диапазоне 0-50  $\text{см}^{-1}$ .**

	0-10 $\text{см}^{-1}$		10-30 $\text{см}^{-1}$		30-50 $\text{см}^{-1}$	
	HITRAN	GEISA	HITRAN	GEISA	HITRAN	GEISA
<b>H<sub>2</sub><sup>17</sup>O</b>		<u>0/1</u> 0%		<u>0/10</u> 0%	9/25 36%	9/29 36%
		<u>0/1</u> 0%		<u>0/10</u> 0%	<u>0/25</u> 0%	<u>0/29</u> 0%
	<i>28/31</i> 90%	<u>0/1</u> 0%	<u>0/15</u> 0%	<u>0/15</u> 0%	<u>17/33</u> 52%	<u>17/33</u> 52%
<b>H<sub>2</sub><sup>18</sup>O</b>	<i>8/31</i> 26%	<u>0/1</u> 0%	<u>0/15</u> 0%	<u>0/15</u> 0%	<u>0/33</u> 0%	<u>0/33</u> 0%
	<b>20/65</b> 31%	<b>20/66</b> 30%	<b>39/76</b> 51%	<b>20/76</b> 26%	<b>60/76</b> 79%	<b>60/76</b> 79%
	<i>0/65</i> 0%	<i>0/66</i> 0%	<i>0/76</i> 0%	<i>0/76</i> 0%	<i>2/76</i> 3%	<i>2/76</i> 3%
<b>HD<sup>18</sup>O</b>	<i>26/53</i> 49%		<i>52/61</i> 85%		<i>28/45</i> 62%	
	<i>8/53</i> 15%		<i>3/61</i> 5%		<i>0/45</i> 0%	

Особое внимание следует обратить на факт отсутствия данных из экспертных массивов в вариационных расчетах. Это относится к диапазону 0-10  $\text{см}^{-1}$  для изотопмеров H<sub>2</sub><sup>18</sup>O и HDO, а для изотопомера HD<sup>18</sup>O к диапазону 0-50  $\text{см}^{-1}$ . Это несоответствие ставит вопрос о необходимости постановки экспериментов в указанных диапазонах.

Таблица 3б. Сравнение массивов спектральных данных Hitran и GEISA в диапазоне 50-20000 см<sup>-1</sup>.

	50-200 см <sup>-1</sup>		200-10000 см <sup>-1</sup>		10000-20000 см <sup>-1</sup>	
	HITRAN	GEISA	HITRAN	GEISA	HITRAN	GEISA
<b>H<sub>2</sub><sup>17</sup>O</b>	<i>147/279</i>	<i>147/279</i>	<i>1309/4762</i>	<i>1383/4828</i>	<i>118/1916</i>	<i>9/325</i>
	53%	53%	28%	29%	6%	3%
	<i>0/279</i>	<i>0/279</i>	<i>669/4762</i>	<i>119/4828</i>	<i>639/1916</i>	<i>107/325</i>
	0%	0%	14%	2%	33%	33%
<b>H<sub>2</sub><sup>18</sup>O</b>	<i>205/342</i>	<i>205/342</i>	<i>1529/6889</i>	<i>1380/7109</i>	<i>164/2443</i>	<i>23/863</i>
	60%	60%	22%	20%	7%	3%
	<i>2/342</i>	<i>0/342</i>	<b>302/6889</b>	<b>317/7109</b>	<b>208/2443</b>	<b>49/863</b>
	1%	1%	4%	4%	9%	6%
<b>HDO</b>	<i>370/599</i>	<i>370/599</i>	<b>1674/8825</b>	<b>2980/11045</b>	<i>8/3483</i>	<i>8/119</i>
	62%	62%	19%	27%	0%	7%
	<b>2/599</b>	<b>0/599</b>	<b>5/8825</b>	<b>27/11045</b>	<b>94/3483</b>	<b>0/119</b>
	0%	0%	0%	3%	0%	
<b>HD<sup>17</sup>O</b>			7/175	7/175		
			4%	4%		
			0/175	0/175		
			0%	0%		
<b>HD<sup>18</sup>O</b>	<i>378/544</i>		<i>434/908</i>	<i>195/679</i>		
	69%		48%	29%		
	<i>16/544</i>		<b>221/908</b>	<b>221/679</b>		
	3%	24%	33%			

## 7. ЗАКЛЮЧЕНИЕ

В докладе представлено обоснование необходимости применения ограничения существования спектральных данных по отношению к экспертным массивам. Применение этого ограничения возможно только в том случае, когда декомпозиция проводится по полному набору данных из всех опубликованных работ в предметной области, относящимся к решениям задач по определению значений исследуемой физической величины.

На примере экспертных массивов спектральных данных, используемых в задачах, входными данными которых являются параметры спектральных линий (задачи астрономии, оптики атмосферы, переноса излучения в атмосфере и т.д.) показано, что имеет место включение в эти массивы значений данных, которые не были опубликованы и достоверность их не документирована.

## ЛИТЕРАТУРА

- [1]. IUPAC project N 2004-035-1-100 «A database of water transitions from experiment and theory». <http://www.iupac.org/web/ins/2004-035-1-100>.
- [2]. J.Tennyson, P.F.Bernath, L.R.Brown, et al., IUPAC Critical Evaluation of the Rotational-Vibrational Spectra of Water Vapor. Part I. Energy Levels and Transition Wavenumbers for H<sub>2</sub><sup>17</sup>O and H<sub>2</sub><sup>18</sup>O // J. Quant. Spectrosc. Radiat. Transfer, 2009, V. 110, Issue 9, P. 573-596.
- [3]. J. Tennyson, P. F. Bernath, L.R. Brown, et al., IUPAC critical evaluation of the rotational-vibrational spectra of water vapor. Part II: Energy levels and transition wavenumbers for HD<sup>16</sup>O, HD<sup>17</sup>O, and HD<sup>18</sup>O // J. Quant. Spectrosc. Radiat. Transfer, 2010, V. 111, Issue 15, P. 2160-2184.

- [4]. L.S. Rothman, I.E. Gordon, A. Barbe, et al., The HITRAN 2008 molecular spectroscopic database. // *J. Quant. Spectrosc. Radiat. Transfer*, 2009, V. 110, Issue 9, P. 533-572.
- [5]. N. Jacquinet-Husson, N.A. Scott, A. Chédin, et al., The GEISA spectroscopic database: Current and future archive for Earth and planetary atmosphere studies, // *J. Quant. Spectrosc. Radiat. Transfer*, 2008, V. 109, Issue 6, P. 1043-1059.
- [6]. P. De Natale, L. Lorini, M. Inguscio, I. G. Nolt, Jae H. Park, G. Di Lonardo, L. Fusina, Peter A. R. Ade, and A. G. Murray, Accurate frequency measurements for H<sub>2</sub>O and <sup>16</sup>O<sub>3</sub> in the 119-cm<sup>-1</sup> OH atmospheric window, *Appl. Opt.*, 1997, V. 36, no. 33, P. 8526-8532.
- [7]. H. Naus, W. Ubachs, P. F. Levelt, O. L. Polyansky, N. F. Zobov and J. Tennyson, Cavity-Ring-Down Spectroscopy on Water Vapor in the Range 555-604 nm. *J. Mol. Spectrosc.*, 2001, V.205, P.117-121.
- [8]. Schermaul, R., Learner, R.C.M., Newnham, D.A., Williams, R.G., Ballard, J., Zobov, N.F., Belmiloud, D., Tennyson, J., The water vapor spectrum in the region 8600-15000 cm<sup>-1</sup>: experimental and theoretical studies to a new spectral line database. *J. Mol. Spectrosc.*, 2001, V. 208, P. 32-42
- [9]. L. R. Brown, R. A. Toth and M. Dulick, Empirical Line Parameters of H<sub>2</sub><sup>16</sup>O near 0.94 μm: Positions, Intensities and Air-Broadening Coefficients. *J. Mol. Spectrosc.*, 2002, V. 212, P. 57-82
- [10]. Tereszchuk K., P.F.Bernath, N.F.Zobov, S.V.Shirin, O.L.Polyansky, N.I.Libeskind, J.Tennyson and L.Wallace, Laboratory Spectroscopy of Hot Water Near 2-μm and Sunspot Spectroscopy in the H-Band Region. *The Astrophysical Journal*, 2002, V. 577, P. 496-500.
- [11]. P.-F. Coheur, S. Fally, M. Carleer, C. Clerbaux, R. Colin, A. Jenouvrier, M.-F. Mérianne, C. Hermans, A. C. Vandaele, New water vapor line parameters in the 26000–13000 cm<sup>-1</sup> region, *J. Quant. Spectrosc. Radiat. Transfer*, V. 74, Issue 4, 2002, P. 493-510.
- [12]. M.-F. Mérianne, A. Jenouvrier, C. Hermans, A. C. Vandaele, M. Carleer, C. Clerbaux, P.-F. Coheur, R. Colin, S. Fally, M. Bach, Water vapor line parameters in the 13 000–9250 cm<sup>-1</sup> region, *J. Quant. Spectrosc. Radiat. Transfer*, V. 82, Issues 1-4, 2003, P. 99-117.
- [13]. L.H.Coudert, O.Pirali, M.Vervloet, R.Lanquetin and C.Camy-Peyret, The eight first vibrational states of the water molecule: measurements and analysis. *J. Mol. Spectrosc.*, 2004, V. 228, p. 471-498.
- [14]. Dupre P., T.Germain, N.F.Zobov, R.N.Tolchenov and J.Tennyson, Continuous Wave - Cavity ring down near ultraviolet rotation-vibration spectrum of water. *J. Chem. Phys.*, 2005, V. 123, p. 154307.
- [15]. S. Kassi, P. Macko, O. Naumenko, A. Campargue, The absorption spectrum of water near 750 nm by CW-CRDS: contribution to the search of water dimer absorption. *Phys. Chem. Chem. Phys.*, 2005, V. 7, no. 12, p. 2460-2467
- [16]. Toth R.A., Measurements of positions, strengths and self-broadened widths of H<sub>2</sub>O from 2900 to 8000 cm<sup>-1</sup>: line strength analysis of the 2-nd triad bands. *J. Quant. Spectrosc. Radiat. Transfer*, 2005, V. 94, P. 51-107.
- [17]. P.-F. Coheur, P. F. Bernath, M. Carleer, R. Colin, O. L. Polyansky, N. F. Zobov, S.V. Shirin, R.J. Barber and J. Tennyson, A 3000K laboratory emission spectrum of water, *J. Chem. Phys.* 122, 074307 2005.
- [18]. F. Mazzotti, O.V. Naumenko, S. Kassi, A.D. Bykov, and A. Campargue, ICLAS of weak transitions of water between 11300 and 12850 cm<sup>-1</sup>. Comparison with FTS databases. *J. Mol. Spectrosc.*, 2006, V. 239, no. 2, P. 174-181.
- [19]. O.Naumenko, M.Sneep, M.Tanaka, S.V.Shirin, W.Ubachs and J.Tennyson, Cavity ring-down spectroscopy of H<sub>2</sub><sup>17</sup>O in the range 16570-17125 cm<sup>-1</sup>. *J. Mol. Spectrosc.*, 2006, V. 237, no. 1, P. 63-69.
- [20]. F. Matsushima, N. Tomatsu, T. Nagai, Y. Moriwaki, K. Takagi, Frequency measurement of pure rotational transitions in the ν<sub>2</sub> = 1 state of H<sub>2</sub>O. *J. Mol. Spectrosc.*, 2006, V. 235, P. 190-195. S. N. Mikhailenko, Wang Leb, S. Kassi and A. Campargue, *Weak water absorption lines around 1.455 and 1.66 μm by CW-CRDS*. *J. Mol. Spectrosc.*, 2007, V. 244, no. 2, P. 170-178.

- [21]. Campargue, S. Mikhailenko, A.W. Liu, ICLAS of water in the 770 nm transparency window 12 746–13 558  $\text{cm}^{-1}$ . Comparison with current experimental and calculated databases, *J. Quant. Spectrosc. Radiat. Transfer*, V. 109, 2008, P. 2832-2845.
- [22]. T. Furtenbacher, A. G. Császár, *On employing  $\text{H}_2^{16}\text{O}$ ,  $\text{H}_2^{17}\text{O}$ ,  $\text{H}_2^{18}\text{O}$ , and  $\text{D}_2^{16}\text{O}$  lines as frequency standards in the 15–170  $\text{cm}^{-1}$  window*, *J. Quant. Spectrosc. Radiat. Transfer*, V. 109, Issue 7, P. 1234-1251.
- [23]. T. Amano and F. Scappini, Millimeter-wave spectrum of rotationally excited  $\text{H}_2\text{O}$ . *Chemical Physics Letters*, 1991, V. 182, no. 2, P. 93-95.
- [24]. J.C.Pearson, T. Anderson, E. Herbst, F. C. De Lucia, and P. Helminger, Millimeter- and submillimeter-wave spectrum of highly excited states of water, *The Astrophys. J.*, V.379, L41-L43, 1991.
- [25]. F. Matsushima, H. Odashima, T. Iwasaki, S. Tsunekawa and K. Takagi, Frequency measurement of pure rotational transitions of  $\text{H}_2\text{O}$  from 0.5 to 5 THz, *J. Molec. Structure*, 1995, v. 352, p. 371-378.
- [26]. R. Paso and V.-M. Horneman, High-resolution rotational absorption spectra of  $\text{H}_2^{16}\text{O}$ ,  $\text{HD}^{16}\text{O}$ , and  $\text{D}_2^{16}\text{O}$  between 110 and 500  $\text{cm}^{-1}$ , *J. Opt. Soc. Am., B*, 1995, V. 12, no. 10, P. 1813-1838.
- [27]. O. L. Polyansky, J. R. Busler, B. Guo, K. Zhang and P. F. Bernath, The Emission Spectrum of Hot Water in the Region between 370 and 930  $\text{cm}^{-1}$ , *J. Mol. Spectrosc.*, V. 176, Issue 2, P. 305-315.