

Структура Венской базы атомных параметров спектральных линий (VALD3) и ее интеграция в виртуальный центр атомных и молекулярных данных (VAMDC).

© Ю.В. Пахомов

Институт астрономии
Российской академии наук,
Москва, Россия
pakhomov@inasan.ru

© Н.Е. Пискунов

Отделение физики и
астрономии, Уппсальский
университет, Уппсала,
Швеция
piskunov@fysast.uu.se

© Т.А. Рябчикова

Институт астрономии
Российской академии наук,
Москва, Россия
ryabchik@inasan.ru

© Р.Л. Куруц

Гарвард-Смитсоновский
центр астрофизики,
Гарвардский университет,
Кэмбридж, США

© Г.К. Стемпельс

Отделение физики и астрономии, Уппсальский
университет, Уппсала, Швеция

© У. Хейтер

Аннотация

Представлено описание структуры Венской базы атомных параметров спектральных линий VALD, возможности ее использования в физике и астрофизике. Также описана интеграция VALD в европейский проект VAMDC - Виртуальный центр атомных и молекулярных данных.

1 Венская база атомных параметров спектральных линий

Венская база атомных параметров спектральных линий (VALD) была создана в 1995 году группой астрофизиков из Австрии, Швеции и России [1] для исследования химического состава и структуры космических объектов - комет, планет, звезд, галактик, межзвездной среды методами спектроскопии. Большое количество информации по спектральным линиям, представленной, в основном, в публикациях в различных форматах и физических единицах, было необходимо преобразовать в единый формат для дальнейшего использования в различных областях спектроскопии. Поэтому основа VALD – это компиляция данных из различных источников. Ее вторая версия VALD2 появи-

лась в 1999 году [2]. Обе версии содержали параметры только атомарных спектральных линий.

VALD – постоянно развивающаяся база данных. Начиная с 2012 года начала функционировать третья версия VALD, которая существенно расширена за счет включения новых атомарных линий и, главным образом, многочисленных линий двухатомных молекул TiO, CN, CH, C₂, O₂, SiH, FeH, и других, а также H₂O. Атомарные данные основаны на новых теоретических расчетах, выполненных Р.Куруцем [3], которые позволили увеличить количество спектральных линий в 7-10 раз по сравнению с VALD2. VALD3 содержит около 250 млн спектральных линий атомов и ионов до девятой стадии ионизации и около 1.5 млрд линий двухатомных молекул и воды. Каждая спектральная линия в базе данных характеризуется полным набором параметров, необходимых для любых видов спектроскопических исследований.

VALD – это разработанная авторами специализированная реляционная база данных с индексацией и сжатием данных. Основное отличие от баз SQL в том, что уникальной является не каждая фактическая запись, а основной объект информации - спектральная линия. Каждой спектральной линии, которая характеризуется длиной волны, силой осциллятора, потенциалом возбуждения нижнего уровня и др., может соответствовать много записей, параметры которых будут отличаться, так как получены из различных источников. Но пользователю должна быть отправлена только одна запись. И главная задача, которую

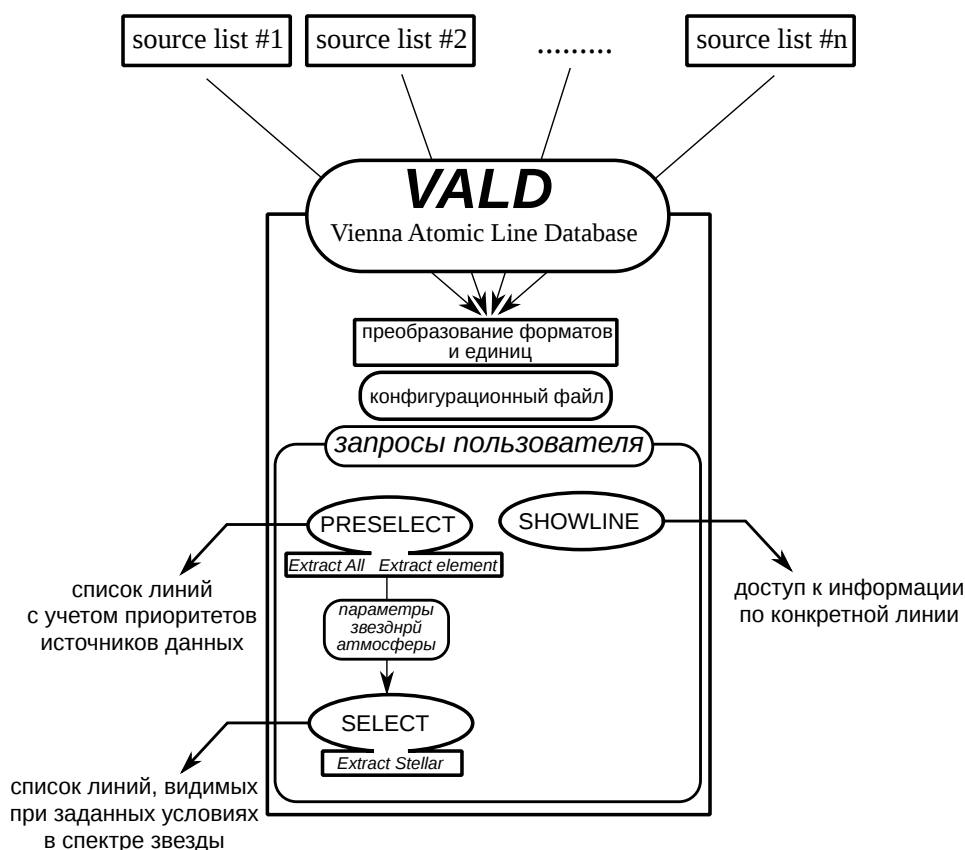


Рис. 1: Структура VALD

решает VALD – идентификация спектральных линий из различных источников данных и выборка этих данных в зависимости от приоритета источника, либо в зависимости от текущей конфигурации базы данных. Для реализации на SQL потребовался бы очень сложный запрос с множественными подзапросами и группировками, либо написание SQL скрипта. Однако оба варианта значительно проигрывают VALD как по скорости обработки и выдачи информации, так и по потребляемым ресурсам, прежде всего памяти и процессора.

В VALD данные отсортированы по длинам волн спектральных линий в вакууме. Пользователь имеет возможность получать значения длин волн в одной из трех систем: в вакууме, в воздухе или в виде волновых чисел. В отличие от VALD2, где энергии уровней хранились в электрон-вольтах, в VALD3 значения энергии даны в см^{-1} , что обеспечивает более высокую точность. Также в отличие от предыдущих версий VALD3 содержит полное описание энергетических уровней и термов, соответствующее стандартам МАГАТЭ. Для каждого перехода приводится точность значения вероятности перехода, если она приводится в оригинальной публикации. Для удобства пользователей предусмотрена расширенная система ссылок на оригинальные публикации, которая выдается пользователю в формате BibTeX вместе с запрашиваемы-

ми данными.

Вновь поступившие данные по атомным параметрам тщательно анализируются, и им присваиваются приоритеты по всем позициям строки записи. Все приоритеты заносятся в конфигурационный файл, на основании которого производится окончательная выборка данных для пользователя. Команда VALD предоставляет рекомендованный экспертами конфигурационный файл, однако для каждого пользователя, есть возможность изменять этот файл (корректируя приоритеты), что позволяет использовать базу данных VALD для различных задач.

Параметры конкретных атомных спектральных линий по данным различных источников можно получить по запросу “SHOWLINE”, который доступен в online режиме на сайте VALD. Пользователю предоставляется возможность сравнивать данные различных авторов, увидеть приоритеты каждого источника данных и окончательный список спектральных линий, сформированный с учетом этих приоритетов.

По запросу пользователя “EXTRACT ALL” или “EXTRACT ELEMENT” происходит поиск необходимой части базы в сжатом виде с использованием созданного индекса и выдача данных в заданном диапазоне длин волн в соответствии с приоритетами конфигурационного файла. За это от-



Рис. 2: Структура VAMDC

вечает внутренняя программа PRESELECT. Кроме того, VALD предоставляет пользователю ряд инструментов для выполнения запросов, которые необходимы в спектроскопических исследованиях. Разработана программа SELECT для работы с базой VALD, особенностью которой является возможность не столько простой экстракции данных по заданным параметрам, но выдача тех спектральных линий, которые будут наблюдаться в определенных физических условиях (запрос “SELECT STELLAR”). Например, нет необходимости выдавать информацию по высокоионизованным атомам в условиях атмосфер холодных звезд, так как их там просто не будет. Также не нужно выдавать параметры молекул в горячей плазме, где они отсутствуют.

Обработка запросов происходит последовательно в соответствии с очередью по времени поступления запроса. Результат можно получить через электронную почту, для выборки не более 1000 спектральных линий, или через FTP, что более предпочтительнее для больших объемов данных.

В настоящее время более 3000 пользователей из 52 стран мира зарегистрированы в базе данных VALD3, которая имеет три официальных зеркала:

в Вене ¹, Москве ² и Уппсале ³

О популярности базы данных VALD среди спектроскопистов свидетельствует индекс цитируемости статей – более 1500 ссылок.

2 Интеграция VALD в европейский проект VAMDC

Несмотря на большой объем информации, VALD3 не содержит ряда данных, необходимых для детального анализа химических процессов, например, скорости различных химических реакций, сечения столкновений и т.д. Эти данные содержатся в других базах, однако форматы записи и хранения данных, различные пользовательские интерфейсы существенно усложняют для пользователя работу с этими базами. В связи с этим возникла идея создания инфраструктуры, которая позволила бы пользователю свободно общаться со всеми базами, содержащими необходимые данные для его научных исследований. Эта идея была реализована в проекте VAMDC – виртуальный центр атомных и молекулярных данных, поддерживаемый Европейским Сообществом. В проект VAMDC, ру-

¹ <http://vald.astro.univie.ac.at/~vald3/php/vald.php>

² <http://vald.inasan.ru/~vald3/php/vald.php>

³ <http://vald.astro.uu.se/~vald/php/vald.php>

ководимый Центром научных исследований Франции (CNRS), входят представители шести стран ЕС (Франция, Германия, Швеция, Великобритания, Австрия, Италия) и трех стран-партнеров (Венесуэла, Россия, Сербия) [4]. VALD3 является одной из 30 баз данных проекта VAMDC ⁴.

Центральный портал VAMDC находится в Париже и выполняет роль посредника между пользователем и отдельными базами данных. Через специальные VAMDC интерфейсы с помощью разработанных протоколов центральный портал общается, с одной стороны, с пользователем (интерфейс пользователя) и, с другой стороны, с базами данных (интерфейс баз данных).

Данные по атомным и молекулярным параметрам поддерживаются их авторами, во-первых, в своем внутреннем формате (ASCII файлы, таблицы и т.д.) для работы в своей лаборатории, институте, и во-вторых, для проекта VAMDC в виде SQL базы данных с единой формой наименования таблиц и их полей. В качестве SQL баз данных могут быть использованы любые коммерческие (MSSQL, Oracle, Access и др.) или некоммерческие СУБД (MySQL, MariaDB, PostgreSQL, SQLite и др.), для которых существуют драйвера для языков программирования Python или JAVA. Для создания нового узла VAMDC на сервере лаборатории или института - производителя данных, устанавливается программное обеспечение Node-Software. В центральном портале VAMDC собрана вся информация о содержимом зарегистрированных баз данных в едином формате (регистр баз данных), что обеспечивает эффективность обработки запроса: выбираются только те базы, которые имеют запрашиваемые пользователем данные.

Запрос пользователя может быть осуществлен как из Портала Пользователя VAMDC⁵, так и минуя формы Портала Пользователя из любой программы с помощью WWW-запроса, используя протокол TAP (Transport Access Protocol), который в некоторой степени совместим с SQL запросами, например: `select * where (RadTransWavelength >= 5000.0 AND RadTransWavelength <= 5001.0) AND ((AtomSymbol = 'Fe'))` Запрос поступает в центральный портал, где происходит его интерпретация и отправка в базы, содержащие требуемые в запросе данные. Пользователь также имеет возможность послать тот же запрос напрямую в конкретную базу данных. Однако использование портала более предпочтительнее, поскольку отслеживаются все обновления в базах данных, их работоспособность в настоящее время и производится оценка объема запрашиваемых данных. Центральный портал в этом случае обеспечивает сбор и отправку результатов запроса пользователю, от-

слеживая целостность предоставляемых данных, а также отвечает за надежность работы всей инфраструктуры.

Результаты запроса выдаются в виде XML файла, в котором кроме непосредственно данных содержится их описание, единицы измерения и другая необходимая информация. Файл соответствует схеме XSAMS ⁶, принятой МАГАТЭ в качестве стандарта для описания атомных и молекулярных параметров. Структурно XSAMS состоит из основных объектов: библиографический блок ссылок на используемые данные, описание методов, описание параметров среды (температура, давление и т.д.), параметры атомов, молекул или частиц, параметры процессов (взаимодействия, переходы и т.д.):

```
<XSAMSData xsi:schemaLocation=
"http://vamdc.org/xml/xsams/0.3">
  <Sources></Sources>
  <Methods></Methods>
  <Environments></Environments>
  <Species></Species>
  <Processes></Processes>
</XSAMSData>
```

Портал Пользователя предоставляет возможность конвертирования полученных результатов из XSAMS в другие форматы, в том числе в формат VALD, который используется для расчета синтетического спектра звезд и содержаний химических элементов в звездных атмосферах.

Однако при использовании VAMDC для запросов в VALD пользователь ограничен возможностью только простой выборки спектральных линии и атомных данных, что соответствует запросам в VALD "SELECT ALL" и "SELECT ELEMENT". При этом отсутствует доступ к редактированию конфигурационного файла VALD, и используется стандартный рекомендованный файл. Вследствие этого отсутствует функционал запроса "SHOWLINE", то есть пользователь получит атомные данные только из одного источника и не сможет их сравнить с данными других авторов. Также нет и подобия запроса "SELECT STELLAR", пользователю выдается весь набор спектральных линии в заданном диапазоне длин волн без анализа интенсивностей этих линии в звездных атмосферах.

Вхождение VALD в состав проекта VAMDC ни в коей мере не уменьшает его функционал и возможности запросов для пользователя. Специалисты по астрофизике и звездной спектроскопии смогут продолжать пользоваться стандартными запросами в VALD, организованными через сайт или email. Тогда как VAMDC открывает одновременно доступ к VALD и другим базам данных многим другим специалистам.

⁴<http://vamdc.eu>

⁵<http://portal.vamdc.eu>

⁶<http://www-amdis.iaea.org/xsams/index.html>

Список литературы

- [1] Kupka F., Piskunov N., Ryabchikova T.A., Stempels H.C., Weiss W.W. VALD-2: Progress of the Vienna Atomic Line Data Base // *Astronomy & Astrophysics Supplement Serie.* – 1999. – V. 138. – P. 119–133.
- [2] Piskunov N.E., Kupka F., Ryabchikova T.A., Weiss W.W., Jeffery C.S. VALD: The Vienna Atomic Line Data Base // *Astronomy & Astrophysics Supplement Serie.* – 1995. – V. 112. – P. 525–535.
- [3] Kurucz R.L. Online materials <http://cfaku5.cfa.harvard.edu/atoms.html>
- [4] Dubernet M.L., Boudon V., Culhane J.L., et al. Virtual atomic and molecular data centre // *Journal of Quantitative Spectroscopy & Radiative Transfer.* – 2010. – V. 111, No. 15, – P. 2151–2159.