

УДК 504.064.36

Р.А. Родригес Залепинос, аспирант  
Донецкий национальный технический университет, г. Донецк, Украина  
rodrigues@csm.donntu.edu.ua

## Новые методы и способы доступа к данным дистанционного зондирования Земли, их организации и визуализации в системах экологического мониторинга окружающей природной среды

*От оперативности и ясности данных, предоставляемых системами экологического мониторинга, напрямую зависит эффективность проводимых мероприятий по улучшению экологической обстановки. Данные дистанционного зондирования Земли предоставляют беспрецедентную возможность получать наборы экологических показателей с большим разрешением, достоверностью и частотой. На сегодняшний день их использование крайне ограничено по причине больших объемов этих данных и сложных форматов их хранения. В статье описаны новые методы и способы, позволяющие: а) организовать данные мониторинга различных форматов и ведомств в единую информационную систему; б) обеспечить постоянный доступ ко всему имеющемуся объему данных в реальном времени; в) выполнить интерактивную трехмерную визуализацию данных. Специалисты в области природоохранной деятельности получают возможность более оперативно и наглядно анализировать большие объемы данных мониторинга, что повышает эффективность поддержки принятия решений в сфере экологической безопасности.*

**Ключевые слова:** данные дистанционного зондирования Земли, доступ в реальном времени, трехмерная визуализация, человеко-машинное взаимодействие, экологический мониторинг.

### Введение

Экологический мониторинг является неотъемлемой частью государственной деятельности по охране окружающей среды, сохранению национального богатства и здоровья населения. На сегодняшний день, данные дистанционного зондирования Земли (ДЗЗ) не находят широкого применения для решения задач мониторинга. Основным препятствием в широком использовании данных ДЗЗ играет не столько сложность алгоритмов их анализа, сколько сама организация доступа к ним.

Описанные в статье методы и способы, позволяют существенно упростить доступ к данным ДЗЗ и их анализ. Данные организуются в единое иерархическое пространство имен и интерактивно визуализируются в 3D.

При этом доступ к данным осуществляется в реальном времени: они автоматически неприменно для пользователя подгружаются по мере надобности с удаленного ЦОД. У пользователя появляется возможность исследовать весь большой объем доступных данных, причем единообразно, не имея представления ни об их типах, ни о форматах.

Представленные методы и способы позволяют организовать и получить доступ к данным ДЗЗ с довольно крупным разрешением (до  $1^\circ \times 1^\circ$ ). Однако, они могут быть модифицированы для

работы с данными более высокого разрешения. Их главный вклад состоит в том, что они упрощают схему доступа к данным ДЗЗ и проведения экологического мониторинга на их основе. Это является существенным шагом вперед к установлению новых стандартов предоставления данных широкому кругу пользователей.

### Проблематика

Основные причины низкой популярности данных дистанционного зондирования Земли:

- **отсутствие систематизированного перечня доступных данных:** существует множество спутников ДЗЗ, продуктов на их основе и ведомств, их распространяющих. Каждая группа создает собственные Интернет ресурсы, непохожие друг на друга, со сложным набором форм для выбора необходимых параметров;

- **данные ДЗЗ необходимо заказывать**, а потом загружать, используя протоколы HTTP либо FTP, что каждый раз отнимает время и на очередной выбор параметров и на саму загрузку;

- **заказанные данные ДЗЗ выдаются в виде порций**, разбитых по времени и пространству - наборы файлов с длинными и сложными для восприятия человеком именами;

- **отсутствуют средства организации набора файлов ДЗЗ:** каждый файл нужно явно открывать и исследовать отдельно от других;

- **файлы имеют сложные научные форматы**, использующие сжатие и хранящие массу метаданных. Часто требуется разработать дополнительный программный код чтения из них для передачи в последующие стадии анализа.

Отсюда следует ряд организационных, и даже психологических барьеров:

- **низкая информированность**: специалисты в области экологического мониторинга зачастую даже не догадываются о существовании многих полезных продуктов на основе данных ДЗЗ;

- **большие объемы данных**: на одной рабочей станции можно проанализировать лишь ограниченный временной отрезок. При необходимости исследования большего объема данных, текущие нужно удалять, а для новых повторять всю процедуру заказа с самого начала;

- **потребность вспомогательных данных**: векторных картах государств и материков для повышения информативности данных ДЗЗ, поиск которых не менее трудоемок;

- **сложность ориентации в данных**: по причине их объемов, разрозненности, отсутствия систематизации и сложности именования.

Специалисты не только расходуют довольно много времени на подготовку к исследованию, но и на сам анализ по причине слабой развитости средств визуализации данных ДЗЗ. Представленные на рынке продукты могут отображать на мониторе отдельные файлы в двумерном режиме. Даже привязка изображения к векторной карте местности уже требует дополнительных усилий. Это отвлекает от первостепенных целей, вносит погрешности и задерживает получение результатов.

### Данные дистанционного зондирования Земли в экологическом мониторинге

Данные MODIS L3 Атмосфера доступны с разрешением  $1^\circ \times 1^\circ$  для всей территории планеты и интервалом в один день с 2000 года по настоящее время [1]. Данные содержат в себе:

- для аэрозолей: типы, оптическую плотность, распределение по размерам примесей, концентрацию масс, оптические свойства и радиационную силу;

- для водяного пара: концентрацию и осадки;

- для облаков: физические и радиационные свойства, включая информацию о переходных фазах (лед – вода, облака – снег), радиус и дисперсность составных частиц облаков, оптическую плотность облаков, температуру их поверхностей, эффекты затенения облаками, высоту верхней области облаков, фазу облаков, долю облаков в дневное либо ночное время;

- количество и распределение озона над территорией.

В данные повторного анализа климата (climate reanalysis) входят все измерения метеорологических спутников, а также приборов на суше, кораблях, вертикального зондирования атмосферы (с помощью водородных баллонов), самолетах, интерполированные на глобальную регулярную широтно-долготную решетку с помощью ассимиляции данных.

NCEP/NCAR Reanalysis версии 1 (R1) [2] впервые выпущен в 1996 г. Он регулярно обновляется и содержит данные с 1942 года по настоящее время. Вторая версия под названием NCEP–DOE AMIP R2 отличается исправленными ошибками и использованием новых систем ассимиляции данных [3]. При этом он содержит данные для более короткого срока – с 1979 г. Повторный анализ ERA-Interim содержит данные с 1979 года по настоящее время.

К показателям относятся температура, влажность, геопотенциал, атмосферное давление на поверхности и среднем уровне моря, скорость и направление ветра. Для каждого момента времени с шагом 6 часов значения показателя размещены в узлах регулярной широтно-долготной решетки с разрешением  $1.5^\circ \times 1.5^\circ$  (ERA-Interim) и  $2.5^\circ \times 2.5^\circ$  (R1 и R2).

### Организация данных

Описанные данные ДЗЗ, получены от различных ведомств. Их перечень организован в единую иерархическую систему имен (рис. 1).

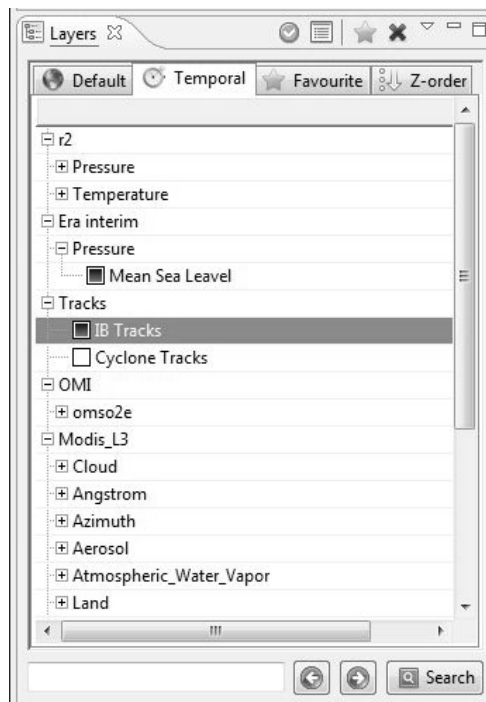


Рисунок 1 – Иерархия доступных данных ДЗЗ

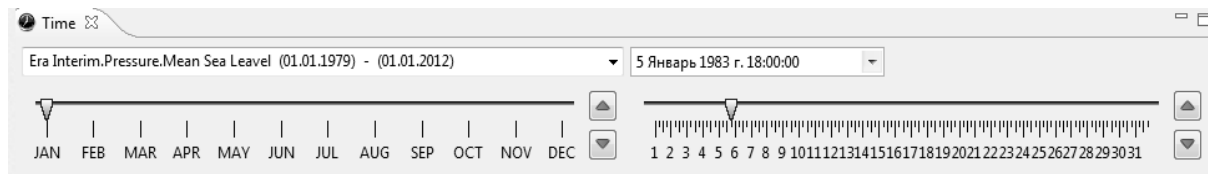


Рисунок 2 – Временная шкала для доступных данных ДЗЗ

Данные ДЗЗ размещаются на высокопроизводительном компьютерном кластере. Для каждого набора данных визуализируется весь доступный интервал в виде простого временной шкалы (рис. 2). С помощью нее можно перемещаться во времени, используя мышь. При этом шаг на шкале каждый раз подстраивается под временное разрешение выбранного показателя.

При описанном подходе любые данные легко найти в общем списке и перемещаться в них одним и тем же способом. Любой момент времени всегда доступен для немедленной визуализации.

**Визуализация данных**

Важность эффективной визуализации для понимания данных трудно переоценить. Она предоставляет гораздо больше информации, чем просто текст либо числа. Трехмерная визуализация данных ДЗЗ на цифровой модели планеты Земля позволяет лучше понять географический контекст исследуемых данных.

По значениям в узлах регулярной широтно-долготной решетки выполняется визуализация

трех видов. Исходные данные отображаются в виде кружков в узлах решетки, размер и цвет которых пропорциональны значениям данных, трехмерная поверхность и изолинии (рис. 3).

Многие типы географических данных моделируются либо интерполируются на регулярную широтно-долготную решетку. Изолинии являются ключевой графикой для понимания данных, расположенных в узлах решетки с географической координатной привязкой. Метод [4] включает в себя эвристики для обработки особенностей координатной системы и реальных данных. Он достаточно быстр и подходит для построения изолиний на лету. Также он представляет изолинии в виде полигонов и гарантирует, что все полигоны замкнуты. Это предоставляет возможность выполнять на них ГИС вычисления, например, расчет занимаемой площади и обнаружение объектов которые покрываются либо пересекаются изолинией. Реализована интерактивная 3D визуализация изолиний для данных повторного анализа. Изолинии визуализируются с огибанием рельефа.

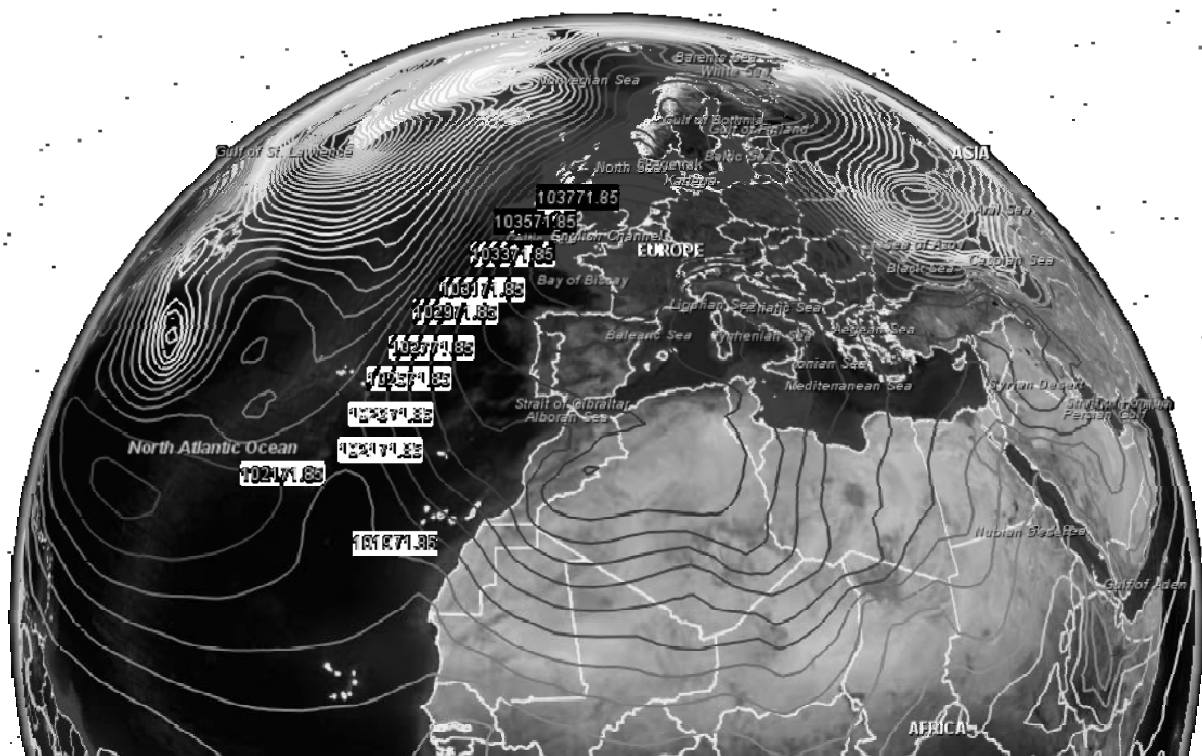


Рисунок 3 – Изобары, построенные по данным повторного анализа ERA-Interim

**Доступ к данным в реальном времени**

Разработана програмная високопроизводительная распределенная система [5], которая работает на компьютерном кластере, построенном на оборудовании широкого потребления. Система обладает масштабируемостью, высокой доступностью и отказоустойчивостью. Она создает на основе уже существующих больших объемов данных деятельный интеллектуальный продукт не изменяя исходные файлы. ChronosServer обнаруживает файлы на узлах кластера, анализирует их структуру и предоставляет независимую от формата SQL-подобную модель запросов для доступа к их содержимому. Он способен напрямую читать сжатые данные из различных форматов, включая NetCDF, GeoTIFF, GRIB, HDF и многих других. Это полностью сохраняет метаданные, хранящиеся в файле, в оригинальном виде, что необходимо для их корректной интерпретации и обработки другим программным обеспечением. Новые данные добавляются в систему прозрачным plug-and-play образом простым копированием их на узел кластера, сокращая затраты на администрирование. Это позволяет существующему программному обеспечению, например, ГИС системам либо статистическим пакетам напрямую оперировать с файлами, которые используются ChronosServer, а также не изменять старые коды генерации данных. ChronosServer сохраняет действующую на данный момент инфраструктуру неизменной, избегая болезненные, трудоемкие и подверженные ошибкам процедуры конвертации файлов данных,

предоставляя в то же время дополнительные возможности для их анализа.

Система состоит из трех компонент: клиентов, рабочих узлов компьютерного кластера и шлюза. Шлюз и узлы кластера объединены в локальную сеть. Клиенты устанавливают соединение со шлюзом через Интернет и инициируют запросы (рис. 4).

Общая схема выполнения запроса работает следующим образом (рис. 5). При запуске приложения автоматически устанавливается соединение (1). Автоматически формируется запрос (2) в виде строки с SQL-подобным синтаксисом.

Например, «SELECT DATA FROM r2.pressure.msl WHERE TIME = 01.01.2003 00:00» вернет регулярную широтно-долготную решетку давления на среднем уровне моря за 1 января 2003 года 00:00 часов.

Шлюз получает строку, разбирает ее (3), определяет рабочие узлы с данными (4), выбирает один из них (5) и отправляет параметры запроса (6). Узел считывает данные с диска (7). Когда данные готовы, рабочий узел асинхронно передает их на шлюз (8), который в свою очередь упаковывает их и отправляет клиенту (9). Клиент принимает данные и кэширует их для последующего использования. За один запрос может вернуться данные для одной решетки. Если необходимо несколько решеток, то нужно выполнить несколько запросов. При таком подходе система лучше масштабируется для большого количества клиентов.

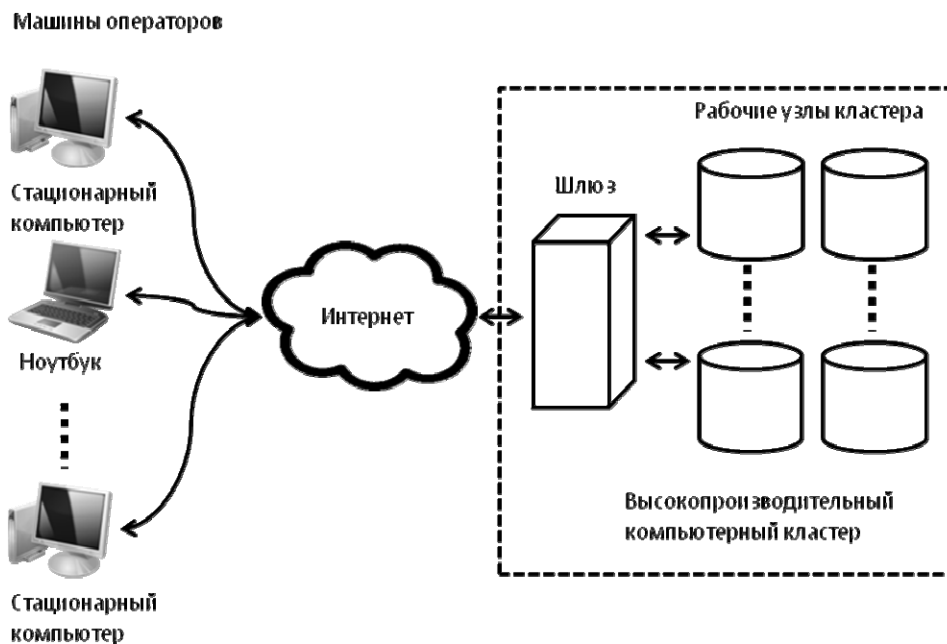


Рисунок 4 – Общая схема доступа к данным в реальном времени множеством одновременных клиентов

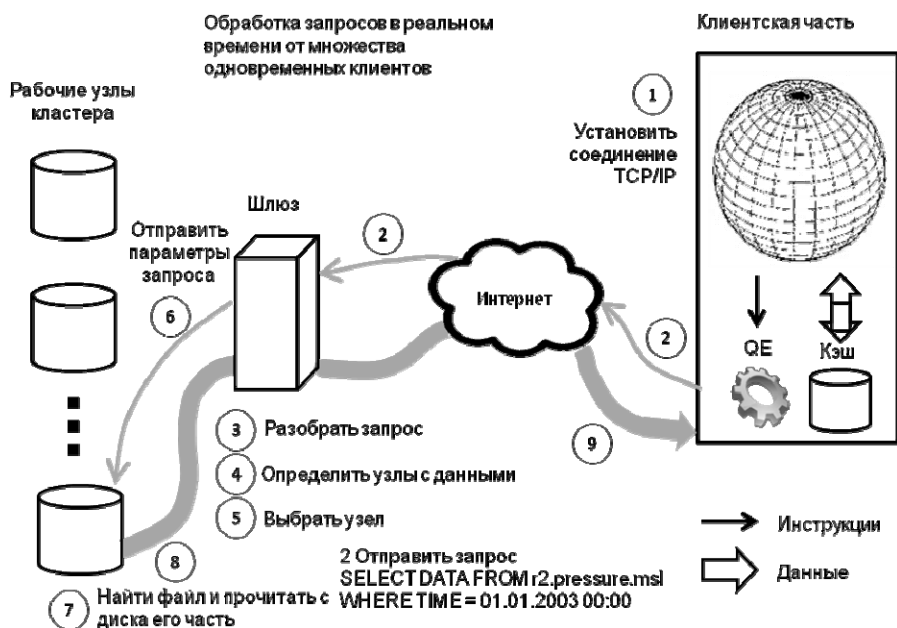


Рисунок 5 – Схема выполнения запроса в реальном времени

Для тестирования системы в режиме повышенной нагрузки, выполнялись синтетические тесты. Для моделирования большого количества клиентов была разработана многопоточная программа. Каждый поток моделирует одного клиента. Шлюз, рабочие узлы и машины клиентов находились в одной 1 Гбит/с локальной сети. Шлюз содержит 8 GB оперативной памяти, CentOS 6.0, Core i7 (8 ядер, 3.46 ГГц). Использовались 6 машин с Windows 7, 4 GB ОП, Intel Core Quad 2 (2.66ГГц) - четыре для клиентов, на двух рабочих узлах виртуальные машины (Ubuntu 10.04, 1GB ОП и 1 ядро). Рабочие узлы были заполнены ранее описанными данными. Для обработки запросов от клиентов на шлюзе было запущено 64 потока и по 32 на рабочих узлах.

На клиентских машинах запускалось 512 потоков, что, в общем, составляет 2048 клиентов. Каждый поток случайным образом выбирает тип данных и время, для которого составляет запрос и инициирует его выполнение. Поток генерирует запросы непрерывно, один за другим. Следующий запрос не генерируется, пока не получен результат выполнения предыдущего запроса. Моделирующие программы работали в таком режиме в течение 5 минут. За это время система успела ответить на 66282 запроса, т.е. 220 запросов в секунду. В реальной ситуации это значение может быть еще выше, поскольку большинство времени для выполнения запроса расходуется на клиенте из-за задержек, связанных с большим числом потоков. Статистика приведена для всех этапов выполнения запроса, начиная от генерации строки, передачи данных до его распаковки на стороне клиента.

### Предыдущая работа

Google Планета Земля [6] визуализирует фотографии Земли в высоком разрешении и реальном времени, полученные коммерческими спутниками дистанционного зондирования Земли. Однако доступны только фотографии в RGB и только для одного момента во времени. Google Планета Земля не поддерживает создание дополнительных модулей (plug-ins). Приложение Unidata IDV [7] предоставляет доступ к данным НАСА и другим серверам. Оно способно выполнять их трехмерную визуализацию. Однако при этом доступны только выборка данных в качестве примеров за определенные моменты времени. Также трехмерная визуализация обладает довольно ограниченным набором интерактивных возможностей. Кроме того, используются WMS сервера, поэтому задержка при загрузке данных довольно заметна. Для некоторых данных, не хранящихся на серверах WMS, предлагается загрузить весь их объем в случае их визуализации за желаемый промежуток времени, что может составлять сотни мегабайт. НАСА Giovanni [8] предоставляет доступ ко всем данным спутников НАСА, однако не имеет интерактивности и визуализирует их в двумерном режиме. Вначале нужно заполнить Интернет формы, основываясь на параметрах которых происходит извлечение данных и их визуализация, на что требуется несколько минут.

### Заключение

Впервые данные мониторинга различных форматов и ведомств организованы в единую

информационную систему. Данные находятся в разных форматах и получены от различных международных агентств по мониторингу окружающей природной среды.

Особенности организации доступа к данным позволяют пользователю не загружать и не

хранить на своей машине файлы с данными. Данные доставляются в реальном времени по Интернет с удаленного ЦОД по мере надобности.

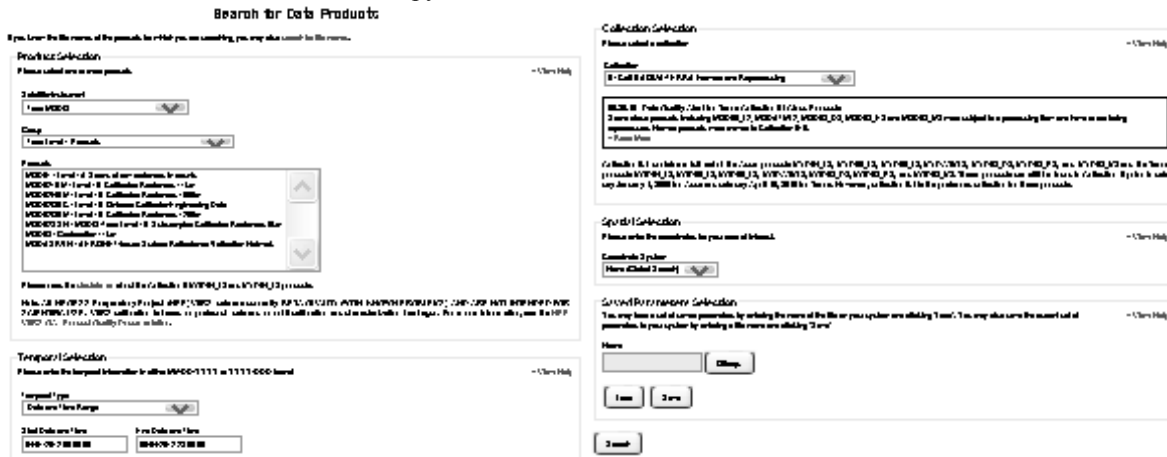


Рисунок 6 – Экранный снимок сервиса НАСА по заказу файлов данных продуктов MODIS [9]

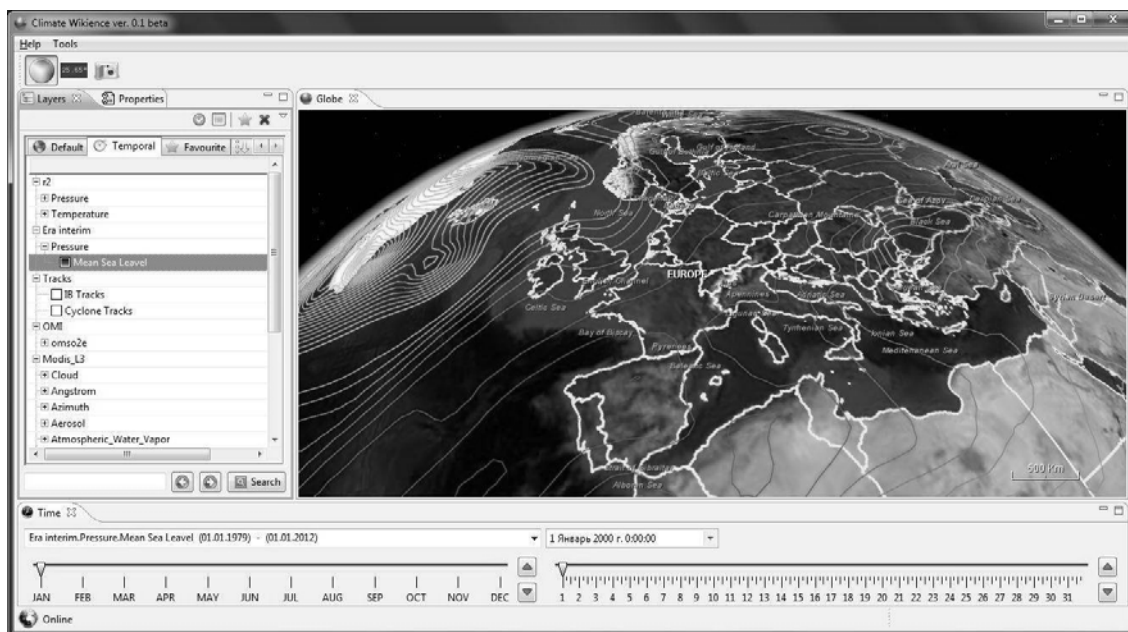


Рисунок 7 – Экранный снимок графического интерфейса системы Climate Wikience [10]

Все моменты времени всегда доступны для немедленной визуализации по щелчку мыши. Это устраняет чувство «тяжелых» данных несмотря на их объемы и сложные форматы. Впервые их исследование становится настолько простым, как просмотр фотографий на локальном компьютере. Это существенно повышает оперативность доступа к данным дистанционного зондирования Земли в системах экологического мониторинга окружающей природной среды.

Описанные методы и способы реализованы в системе Climate Wikience [10], которая доступна для свободной загрузки и использования.

Экранные снимки одного из сервисов НАСА по заказу файлов данных ДЗЗ (рис. 6) и разработанной системы (рис. 7) наглядно демонстрируют преимущество последней.

**Благодарности**

Эта работа была поддержана грантом No. UKM1-2973-DO-09 Фонда гражданских исследований и развития США (CRDF). Любые мнения, анализы, выводы либо рекомендации, выраженные в этой статье, принадлежат автору, и не обязательно отражают официальные взгляды либо мнения CRDF.

**Список использованной литературы**

1. MODIS Atmosphere: Products: [El. resource]. – URL: <http://modis-atmos.gsfc.nasa.gov/products.html> (14.05.2012).
2. The NCEP/NCAR 40-Year Reanalysis Project / Kalnay E. et al. // Bull. Amer. Meteor. Soc. – 1996. – 77. – P. 437–472.
3. NCEP-DEO AMIP-II Reanalysis (R-2) / Kanamitsu M., Ebisuzaki W., Woollen J. et al. // Bulletin of the American Meteorological Society. – 2002. – P. 1631-1643.
4. Rodrigues Zalipynis R.A. Efficient isolines construction method for visualization of gridded georeferenced data / Rodrigues Zalipynis R.A. // Scientific works of Donetsk National Technical University. Series "Problems of Modeling and Design Automation". – Donetsk: DonNTU, 2011. – Vol. 10 (197). – P. 111-123.
5. Rodrigues Zalipynis R.A. ChronosServer: real-time access to "native" multi-terabyte retrospective data warehouse by thousands of concurrent clients / Rodrigues Zalipynis R.A. // Scientific works of Donetsk National Technical University. Series "Informatics, cybernetics and computer engineering". – Donetsk: DonNTU, 2011. – Vol. 14 (188). – P. 151-163.
6. Google Планета Земля: [Электронный ресурс]. – Режим доступа: <http://www.google.com/intl/ru/earth/index.html> (14.05.2012).
7. Unidata IDV: [Электронный ресурс]. – Режим доступа: <http://www.unidata.ucar.edu/downloads/idv/> (14.05.2012).
8. Giovanni — GES DISC: Goddard Earth Sciences, Data & Information Services Center: [Электронный ресурс]. – Режим доступа <http://disc.sci.gsfc.nasa.gov/giovanni/overview/index.html> (14.05.2012).
9. Интернет сервис НАСА по заказу файлов данных продуктов MOIDS: [Электронный ресурс]. – Режим доступа: <http://modwebsrv.modaps.eosdis.nasa.gov/> (14.05.2012).
10. Climate Wikience: [El. resource]. – URL: <http://wikience.donntu.edu.ua> (14.05.2012).

Надійшла до редколегії 14.05.2012

**Р.А. РОДРИГЕС ЗАЛПИНИС**

Донецький національний технічний університет

**R.A. RODRIGES ZALIPYNIS**

Donetsk National Technical University

**НОВІ МЕТОДИ ТА ЗАСОБИ ДОСТУПУ ДО ДАНИХ ДИСТАНЦІЙНОГО ЗОНДУВАННЯ ЗЕМЛІ, ЇХ ОРГАНІЗАЦІЇ ТА ВІЗУАЛІЗАЦІЇ У СИСТЕМАХ ЕКОЛОГІЧНОГО МОНІТОРИНГУ ДОВКІЛЛЯ**

**NEW METHODS AND TECHNIQUES FOR ACCESSING REMOTE SENSING DATA, THEIR ORGANIZATION AND VISUALIZATION IN ECOLOGICAL MONITORING SYSTEMS OF ENVIRONMENT**

У статті описані нові методи та засоби, які дозволяють: а) організувати дані моніторингу різних форматів та відомств у єдину інформаційну систему; б) забезпечити постійний доступ до всього існуючого обсягу даних у реальному часі; в) виконати інтерактивну тривимірну візуалізацію даних. Спеціалісти у галузі природоохоронної діяльності отримують можливість більш оперативно та наочно аналізувати великі обсяги даних моніторингу, що підвищує ефективність підтримки прийняття рішень у галузі екологічної безпеки.

This paper describes new methods and techniques that allow: (i) organize monitoring data in diverse formats and from different agencies into a single information system; (ii) provide permanent access to all existing data volumes in real time; (iii) perform interactive 3D visualization of the data. Experts in ecological security gain the ability to analyze large volumes of monitoring data more operatively and clearly. This increases the efficiency of decision support in ecological security domain.

**Ключові слова:** дані дистанційного зондування Землі, доступ у реальному часі, тривимірна візуалізація, людино-машинна взаємодія, екологічний моніторинг.

**Keywords:** Earth remote sensing data, real time access, 3D visualization, human-computer interaction, environmental monitoring.