

- [Главная](#)
- [Вопросы и ответы](#)

Экспорт данных HDF и HDF-EOS с помощью hdp

Рассмотрены некоторые вопросы, связанные с форматом данных HDF-EOS, а также представлены простейшие методы выгрузки подобного рода данных в двоичный и текстовый ASCII форматы.

[Обсудить в форуме](#) Комментариев — 2

Иерархический формат данных (Hierarchical Data Format, HDF) — стандартный формат данных для всех продуктов системы NASA Earth Observing System (EOS). HDF — довольно сложный формат, развитием и поддержкой которого в настоящее время занимается организация The HDF Group.

Оглавление

- [1. Введение в HDF и HDF-EOS](#)
- [2. Выгрузка данных HDF](#)
 - [2.1 Выгрузка данных HDF в двоичный формат](#)
 - [2.1.1 Unix/Linux](#)
 - [2.1.2 Windows NT/98/2000/XP](#)
 - [2.2 Выгрузка данных HDF в текстовый ASCII формат](#)
- [3. Географическая привязка данных HDF-EOS](#)



Введение в HDF и HDF-EOS

Формат HDF был разработан с целью обеспечения передачи научной информации между разнообразными операционными системами и компьютерными платформами, путем использования интерфейсов языков программирования C и Fortran, а также различных утилит. HDF поддерживает широкий спектр типов данных: n-мерные массивы, таблицы, текстовые описания, всевозможные типы растровых изображений и связанные с ними цветовые палитры и метаданные. Библиотека HDF содержит интерфейсы, позволяющие сохранять и получать данные вышеуказанных типов в их сжатом или несжатом виде.

Для каждого элемента данных в HDF файле существуют специальные теги, определяющие его тип, объём и размерность. Свойство самоописания формата HDF помогает пользователям полностью понять структуру и содержание хранящейся в файле информации. При работе с форматом HDF, пользовательская программа интерпретирует, а затем идентифицирует типы тегов и обрабатывает соответствующие данные. Единственный HDF файл может содержать в себе данные различных типов, например, текстовые, числовые или графические; однако, зачастую растровые изображения и многомерные массивы не содержат в себе информацию о географической привязке. Поскольку большинство наборов данных о Земле нуждаются в подобного рода привязке, The HDF Group разработала формат HDF-EOS, введя дополнительные соглашения и типы данных для HDF файлов.

HDF-EOS поддерживает три типа геоданных: grid, points и swath, обеспечивая стандартный доступ в геопространственном контексте. Библиотека для работы с форматом HDF-EOS позволяет пользователю производить выборку и вырезать фрагменты данных из файла путём указания необходимых координат и времени. Инструменты предназначенные для работы со стандартными файлами HDF, также могут читать и файлы HDF-EOS, однако вызовы стандартной библиотеки HDF не смогут получить доступ к данным о географической привязке, данным времени и метаданным, как в случае использования библиотеки HDF-EOS.



Выгрузка данных HDF

The HDF Group предоставляет различные утилиты для просмотра содержимого файлов HDF и извлечения из них растровых, двоичных или ASCII данных. Одна из них - hdp, обеспечивающая простейший способ выгрузки объектов HDF в двоичный или ASCII формат. Есть несколько команд, используемых в качестве ключей утилиты hdp, которая вызывается следующим образом:

hdp

Например команда `dumpsds` позволяет отобразить данные, содержащиеся в массивах (Scientific Data Set, SDS). На сайте The HDF Group присутствует [документация](#) по использованию `hdp`, но это по сути та же информация, что выводится на экран, если набрать:

```
hdp -h <command>
```

2.1 Выгрузка данных HDF в двоичный формат

2.1.1 Unix/Linux

A. Загрузка библиотек для работы с форматом HDF с сайта The HDF Group

Библиотека сжатия `szip` необходима для запуска `hdp` на платформах Unix/Linux. Следующие инструкции предполагают, что вы используете `sh`, `csch` или `tcsh` на своей Unix/Linux системе.

1. Перейдите в каталог, содержащий прекомпилированные двоичные файлы библиотеки `szip`:

<ftp://ftp.hdfgroup.org/lib-external/szip/2.0/bin/>

2. Библиотека `szip` в скомпилированном виде распространяется двумя способами: в первом случае с поддержкой функции кодирования и декодирования (может потребоваться лицензия), а во втором — только с поддержкой возможности декодирования (без лицензии). В зависимости от того, какой вариант библиотеки используется, соответствующее приложение может создавать, записывать и считывать данные, сжатые с помощью `szip` или только считывать. Скачайте библиотеку `szip` в соответствии с используемой операционной системой в свой домашний каталог.

Например: `szip2-linux2.6-enc.tar.gz`

3. Разархивируйте скачанный файл с помощью следующей команды:

```
gunzip szip2-linux2.6-enc.tar.gz
```

4. Распакуйте полученный на предыдущем этапе `tar` файл:

```
tar xvf szip2-linux2.6-enc.tar
```

Замечание: выполнение этой операции создаст каталог `szip2-linux2.6-enc` в вашей домашней директории.

5. Проверьте, существует ли переменная окружения `LD_LIBRARY_PATH`:

```
echo $LD_LIBRARY_PATH
```

6. Если переменная окружения `LD_LIBRARY_PATH` не определена, создайте ее:

```
export LD_LIBRARY_PATH=$HOME/szip2-linux2.6-enc/lib
```

Если же переменная окружения `LD_LIBRARY_PATH` определена, просто добавьте в нее новый путь:

```
export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:$HOME/szip2-linux2.6-enc/lib
```

B. Загрузка утилиты `hdpc` сайта The HDF Group

1. Перейдите в [каталог скомпилированных бинарных дистрибутивов](#).

2. Выберите директорию, соответствующую вашей платформе.

3. Перейдите в каталог `utilities`.

4. Загрузите утилиту `hdp`.

5. Наберите следующую команду, чтобы сделать файл исполняемым:

```
chmod 755 hdp
```

C. Выбор выгружаемых переменных HDF файла

Пользователи могут конвертировать как все содержимое HDF файла в двоичный формат, так и содержимое отдельных массивов (SDS). Имена массивов представлены в виде списка в заголовке файла HDF.

1. Выгрузите информацию заголовка и укажите, что необходимо выгрузить только имена переменных:

```
hdp dumpsds -h | grep "Variable Name"
```

Например: Имена переменных будут отображены согласно следующему формату:

```
Variable Name = CMG_day  
Variable Name = CMG_night
```

2. Запишите куда-нибудь имена переменных, которые в дальнейшем необходимо будет выгрузить в двоичный формат.

D. Выгрузка данных HDF в двоичный формат

1. Наберите следующую команду:

```
hdp dumpsds -n <variable1 name, ...,variableN name> -o <outputfilename> -b  
<inputfilename.hdf>
```

В результате выполнения вышеприведенной команды необходимые данные (-n <variable1 name,...,variableN name>) будут выгружены в указанный файл (-o <outputfilename>) в двоичном виде (-b).

Замечание 1: Используйте двойные кавычки вокруг имен переменных, если те состоят более чем из одного слова, например "Sea Ice by Reflectance".

Замечание 2: Порядок байтов для платформ Sun/Solaris, SGI, HP и IBM UNIX - от старшего к младшему («big-endian»), о чем необходимо помнить.

2.1.2 Windows NT/98/2000/XP

A. Загрузка утилиты hdp с сайта The HDF Group

1. Перейдите по ссылке: ftp://ftp.hdfgroup.org/HDF/HDF_Current/bin/windows/utilities.
2. Загрузите файл hdp.exe
3. Пользователям Windows XP необходимо загрузить три библиотеки в ту же директорию, где расположен файл hdp.exe: [zlib1.dll](#), [szlibdll.dll](#) и [szlibdll.lib](#):
4. Запустите командную строку.
5. Перейдите в каталог с утилитой hdp.exe. Для удобства можете также поместить свой *.hdf файл в этот же каталог.

B. Выбор выгружаемых переменных HDF файла

1. Выгрузите информацию заголовка в файл в текстовом формате (используется по умолчанию):

```
hdp dumpsds -h -o <outputfilename> <inputfilename.hdf>
```

2. В текстовом редакторе откройте получившийся файл, пролистав его в самый конец, вы увидите имена переменных.
3. Отметьте для себя имена тех переменных, которые необходимо будет выгружать.

C. Выгрузка данных HDF в двоичный формат

1. В командной строке наберите команду:

```
hdp dumphds -n <variable1 name, ...,variableN name> -o <outputfilename> -b
<inputfilename.hdf>
```

Замечание 1: Используйте двойные кавычки вокруг имен переменных, если те состоят более чем из одного слова, например “ Day Tile Snow Cover ”.

Замечание 2: Порядок байтов для платформ Windows - от младшего к старшему («little-endian»), о чём необходимо помнить.

2.2 Выгрузка данных HDF в текстовый ASCII формат

Выгрузка данных в текстовый ASCII формат производится аналогично рассмотренному выше способу выгрузки в двоичный формат, только в этом случае при использовании утилиты hdp вместо ключа -b нужно использовать ключ -x, либо вообще не использовать никакие ключи (ни -b, ни -x), так как по умолчанию выгрузка итак производится в текстовом формате.

Пример:

```
hdp dumphds -n <variable1 name, ...,variableN name> -o <outputfilename>
<inputfilename.hdf>
```

Замечание: В приведенном выше способе данные выгружаются вместе с заголовком HDF файла, чтобы в выходном файле содержались только данные, нужно воспользоваться ключом -d:

```
hdp dumphds -n <variable1 name, ...,variableN name> -o <outputfilename> -d
<inputfilename.hdf>
```

Пример: Рассмотрим выгрузку данных файла [MOD14.A2007201.0735.005.2007202182452.hdf](#) из набора MOD14:

```
hdp dumphds -n FP_longitude,FP_latitude -o output.txt -d MOD14.A20072
01.0735.005.2007202182452.hdf
```

Содержимое выходного файла output.txt будет иметь вид:

```
47.116291 47.114323 47.106380 47.101460 47.099499 47.092617 46.951347
47.659725 47.416451 47.414543 47.093536 46.154545 47.343292 47.342487
47.024929 46.972633 47.022923 46.974819 47.086983 46.443035 47.088673
46.272400 46.828228 46.196171 46.829624 46.827862 46.676620 46.662701
46.643841 46.478455 46.523300 46.478695 46.475666 46.398930 46.344063
46.342270 46.349342 46.335598 43.827866 43.822529 43.706970 43.694920
...
```

```
53.022594 53.035557 53.026371 52.997627 53.010582 52.994797 52.066730
41.673935 44.352989 44.378529 46.810658 53.409946 41.138630 41.130955
42.455647 43.126461 42.577427 43.120789 37.113361 48.199680 37.104752
48.800392 40.626141 48.836739 40.618580 40.658440 41.889912 41.888447
41.714169 41.719471 40.454330 41.438107 41.430126 41.529949 42.034565
...
```



Географическая привязка данных HDF-EOS

Основная функция формата HDF-EOS – хранение географически привязанной информации в одном файле. Каждый объект HDF-EOS (point, swath или grid) поддерживает различные способы привязки. Point и swath используют данные долготы и широты. А для grid-структур библиотека HDF-EOS использует пакет GCTP (U.S. Geological Survey (USGS) General Cartographic Transformation Package), предназначенный для хранения информации о проекции и системе координат. Инструменты, полностью поддерживающие файлы HDF-EOS, способны автоматически считывать массивы широт и долгот, связанных с объектами point и swath и, используя методы библиотеки HDF-EOS, конвертировать значения широт и долгот в столбцы и строки индексов для объектов grid.

Чтобы вручную определить географическое покрытие файла данных, нужно обратиться к файлу метаданных XML (*.xml), который идёт совместно с файлом данных *.hdf. Глобальный текстовый атрибут HDF StructMetadata.0 содержит метаданные о привязке для всех наборов данных содержащихся в HDF. Данные swath второго уровня включают наборы Geolocation Fields, содержащие массивы широт и долгот в каждом

VGroup. Структурные метаданные указывают как эти массивы соотносятся с массивами данных. Подробнее о Geolocation Fields можно прочитать в Разделе 5.1.2 Geolocation Fields, стр. 27 [HDF-EOS Library User's Guide, Volume 1](#). Для HDF-EOS данных третьего уровня, проекция и ее параметры хранятся в структурных метаданных, в формате GCTR.

Свободная утилита преобразования [HDF-EOS в GeoTIFF \(HEG\)](#), позволяет привязывать, вырезать фрагменты и сшивать наборы данных HDF-EOS. На Web сайте утилиты HEG можно найти [список наборов данных](#), поддерживаемых данной программой.

[Обсудить в форуме](#) Комментариев — 2

Ссылки по теме

- [Импорт данных в формате HDF](#)
- [Импорт и обработка данных в формате HDF - пример](#)
- [HDF-EOS Tools and Information Center](#)
- [The HDF Group: Software using HDF \(4\)](#)
- [Langely Atmospheric Science Data Center \(ASDC\): Data Manipulation Software](#)
- [MODIS Atmosphere: Graphical User Interface Tools](#)

Последнее обновление: March 02 2011

Последнее обновление: March 02 2011

Дата создания: 19.01.2009

Автор(ы): [Денис Рыков](#)