

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

# Импорт и обработка данных в формате HDF - пример

Пример импорта растровых данных HDF и их простейшей обработки с помощью ArcGIS.

[Обсудить в форуме](#) Комментариев — 1

Данный пример иллюстрирует импорт и простейшую обработку hdf-файла с помощью ArcGIS в формат, пригодный для любого рода дальнейших вычислений.

В качестве примера будут использоваться данные по цветности вод. Эти данные имеют 8-дневную периодичность и разрешение 4 или 9 км. Такие данные общедоступны и доступны в HDF-формате по адресу: <http://oceancolor.gsfc.nasa.gov/cgi/level3.pl>. Особенностью этих данных является наличие внедренной в них самих документации (в файлы HDF) их описывающей, которую можно также легко извлечь и использовать для более эффективного импорта данных. Если Вы собираетесь использовать другие данные, распространяемые в формате HDF, например данные MODIS, данный подход может не подойти, нужно использовать специальное программное обеспечение, [предназначенное специально для этих данных](#).

Для импорта и обработки нам понадобятся:

- [HDF2Bin](#) – утилита, предназначенная для конвертирования растров, содержащихся в HDF файле (если они там есть) в простой бинарный растр;
- [HDF Explorer](#), либо любая иная программа (NeoSYS и т.д.), позволяющая просматривать структуру и содержимое hdf-файлов, экспортировать их в другие форматы не понадобится, поэтому достаточно простого вьюера, такого как HDF Explorer ([подробнее](#));
- ArcGIS любой версии.



## Получение данных

Скачиваем архив и распаковываем файл (для распаковки можно использовать WinRAR), который будет служить нам в качестве образца, например:

[http://oceancolor.gsfc.nasa.gov/cgi/getfile/A20061452006152.L3m\\_8D\\_CHLO\\_4.bz2](http://oceancolor.gsfc.nasa.gov/cgi/getfile/A20061452006152.L3m_8D_CHLO_4.bz2)

Данный файл представляет образец данных с разрешением 4 км. Дальнейшие расчеты будут основываться именно на этой цифре, если вы используете данные имеющие другое разрешение (например, 9 км), параметры импорта будут другими. При импорте данных, абсолютно необходимо четко представлять, какое разрешение имеют ваши данные. Часто, разрешение данных нигде не "прописано" в самих файлах.



## Получение растрового слоя

Запускаем утилиту [HDF2Bin](#) (она должна находиться в одной папке с импортируемым файлом), используя следующую команду:

```
имя_утилиты  название_импортируемого_файла > param.txt.
```

В нашем случае запись будет выглядеть следующим образом:

```
hdf2bin-win2000.exe A20061452006152.L3m_8D_CHLO_4 >param.txt
```

В результате работы утилиты HDF2Bin будет создан бинарный dat-файл: SDS\_l3m\_data.dat, соответствующий единственной растровой матрице содержащейся в исходном HDF файле. Данный файл является самым простым типом растрового файла и не содержит информации о привязке, размерах растра, разрешение и т.д. Для нормальной работы, понадобится описать структуру файла с помощью внешних файлов описания.

Далее, из полученного файла param.txt необходимо получить информацию о размерах растра по оси X

(количество колонок - NCOLS) и Y (количество рядов - NROWS), а так же числовом формате (разрядность - NBITS). Пример файла param.txt (жирным выделены искомые параметры):

```
uint8 SDS_l3m_data.dat (fakeDim0=4320, fakeDim1=8640)
Data type: unsigned 8-bit integer
Number of Dimensions: 2
Dim 1: "fakeDim0" = 4320
Dim 2: "fakeDim1" = 8640
```



### Создание сопутствующих файлов

Создаем 2 пустых файла (их можно создать в любом текстовом редакторе, например, Блокноте), имеющие расширения \*.blw и \*.hdr, названия файлов должны быть точно равны названию файла HDF. Файл BLW отвечает за географическую привязку изображения ([подробнее о таких файлах](#)), а файл HDR описывается структуру данных, количество рядов, колонок, разрядность и т.д. Изменим расширение dat на bil, чтобы он был "узнан" ArcGIS.

Заполняем файлы \*.blw и \*.hdr данными. Для этого запускаем HDF Explorer (или иную программу для просмотра структуры hdf-файла) и в иерархии находим следующие ключи: latitude step (шаг по широте), longitude step (шаг по долготе), SW point latitude (широта левого нижнего пиксела) и SW point longitude (долгота левого нижнего пиксела). Заносим в blw-файл эти данные:

Параметр	Описание
0.0416667	longitude step
0	0
0	0
-0.0416667	latitude step (пишется с минусом)
-179.9792	SW point longitude
89.9792	SW point latitude (пишется без минуса, т.к. в ArcGIS – это левый верхний угол (NW), а в HDF – левый нижний угол(SW))

Таким образом содержание blw-файла должно быть следующим:

```
0.0416667
0
0
-0.0416667
-179.9792
89.9792
```

Также из hdf-файла берем значения количества строк, колонок и разрядности данных, и записываем в hdr-файл:

### Содержание hdr-файла

### Описание

BYTEORDER I

—

NROWS 4320	Number of lines (Количество строк)
NCOLS 8640	Number of columns (Количество полей (колонок))
NBANDS 1	Количество каналов
NBITS 16	Битность (разрядность) данных

Таким образом содержание hdr-файла должно быть следующим (в зависимости от того, какие у вас данные, эти параметры могут меняться):

```
BYTEORDER I
LAYOUT BIL
NROWS 4320
NCOLS 8640
NBANDS 1
NBITS 16
```



### Загрузка данных в ArcGIS

Запускаем ArcGIS и загружаем в него bil-файл.

При загрузке появляется окно:



Выбираем No.



### Исключение значений NODATA

Импортируемые данные содержат значения NODATA, т.е. участки, которые не несут информацию (на рисунке они белые). В нашем случае эти значения равны 65535, для других данных они могут быть другими (например встречались значения 255, -9999 и другие). Для того, чтобы определить, какое значение является

значением NODATA, необходимо загрузить данные в Arcview/ArcGIS и используя инструмент Identify определить текущее значение в точках, где значения в принципе должны отсутствовать (в нашем случае это например вся суша, где оценки хлорофилла не производилось). После этого, определенное значение и будет нашим NODATA. До проведения корректных расчетов необходимо эти значения исключить из массива данных. Для этого в ArcToolBox запускаем Spatial analyst tools\conditional\set null. В появившемся диалоговом окне указываем:

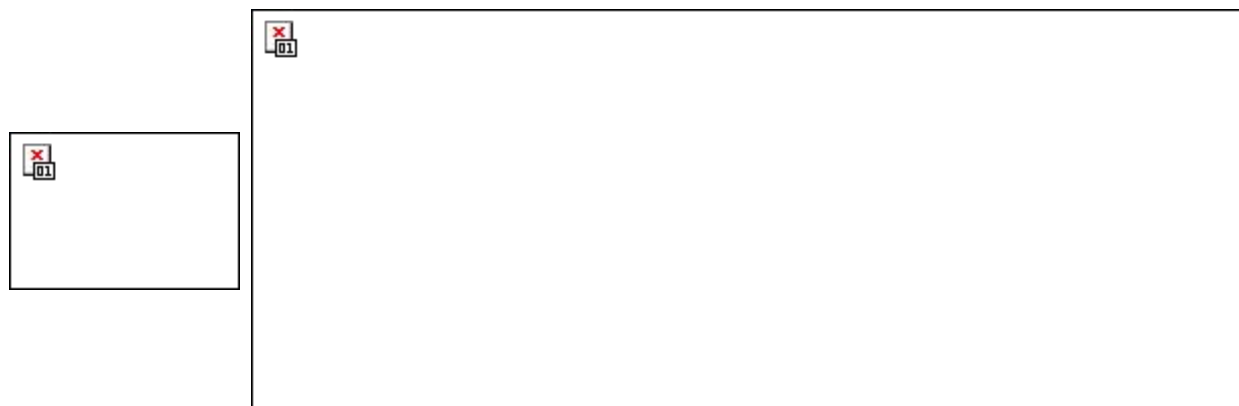
Input conditional raster: SDS\_I3m\_data.bil

Input false raster or constant value: SDS\_I3m\_data.bil

Output raster: имя выходного растра, не длиннее 13 символов

Expression (optional): Value = 65535

Нажимаем ОК. После обработки создается новый слой, который добавляется автоматически во фрейм данных и отличается от предыдущего отсутствием в легенде (и, соответственно, в самих данных) значения 65535.



#### Перерасчет данных

Для того, чтобы получить значения, соответствующие концентрации хлорофилла в воде в мг/м<sup>3</sup>, необходимо, используя формулу, содержащуюся в ключе **Scaling Equation**, и параметры из одноименных ключей из исходного hdf-файла, пересчитать все значения только что полученного растра (без значения 255). Формула имеет следующий вид:

$$\text{Base}^{((\text{Slope} * \text{I3m\_data}) + \text{Intercept})} = \text{Parameter value}$$

где в нашем случае

base – 10

slope – 0.000058138

intercept – -2

setnull\_sds\_3 – данные исходного слоя.

Для перерасчета удобно использовать модуль Raster Calculator из расширения Spatial Analyst. Для его использования необходимо подключить в ArcGIS модуль Spatial Analyst (Tools\Extensions...) и на его панели (View\Toolbars\Spatial Analyst) выбрать Raster Calculator.

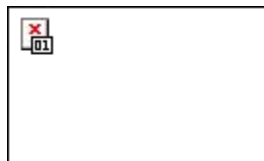
В появившемся окне записываем формулу со всеми коэффициентами. В нашем случае она будет иметь вид:

**$$\text{Pow}(10, (0.000058138 * \text{setnull\_sds\_3} + (-2)))$$**

где setnull\_sds\_3 – название слоя, из которого удалены значения NODATA. Разделитель целой и дробной части может быть как точкой, так и запятой. После ввода формулы нажимаем Evaluate.

Результат – очередной растровый слой, содержащий данные по концентрации пигмента в воде, и его можно (и нужно) раскрасить, щелкнув мышкой по серой полосе под Value в слое, в появившемся диалоговом окне выбрав цвет, при необходимости установив галочку для Invert. Здесь мы можем оценить правильность расчетов. Дело в том, что на всех созданных нами слоях есть шкала, имеющая надписи Low (минимум) и High (максимум) с определенными значениями. В нашем случае в рассчитанном слое, имеющем по умолчанию название Calculation, Low равно 0,01, а High – 64,5664. Эти значения соответствуют значениям, приведенным в

ключах **Scaled data minimum** и **Scaled data maximum** исходного hdf-файла, значит все расчеты произведены верно.



### Сохранение данных

Всё, данные импортированы. Теперь надо их сохранить. Для этого правой клавишей щелкаем по слою Calculation, в контекстном меню выбираем Data\Export data..., а в диалоговом окне указываем местоположение (Location), название (Name) и формат (Format) экспортируемого файла.

Описание вырезки фрагмента данных из полученных растровых слоев можно прочитать в статье "[Вырезка фрагмента данных в формате GRID](#)".

[Обсудить в форуме](#) Комментариев — 1

## Ссылки по теме

- [Импорт данных в формате HDF](#)
- [Вырезка фрагмента данных в формате GRID](#)
- [Документация по формату HDF, отличия форматов и др.](#)

Последнее обновление: March 01 2011

Дата создания: 19.02.2007

Автор(ы): [Максим Дубинин](#)