Subject: Re: Getting mean from HDF SD files - stack in envi or read into IDL array?
Posted by kathryn.davies1 on Sun, 06 Jul 2008 22:38:17 GMT
View Forum Message <> Reply to Message

On Jul 6, 2:25 pm, bulrushmo...@gmail.com wrote:
> On Jul 6, 4:10 am, kathryn.davi...@googlemail.com wrote:
>
>
>
>
>
>> On Jul 5, 10:32 pm, bulrushmo...@gmail.com wrote:
>
>>> On Jul 5, 2:45 pm, bulrushmo...@gmail.com wrote:
>
>>>> On Jul 5, 1:58 pm, bulrushmo...@gmail.com wrote:
>
>>>> > On Jul 5, 10:46 am, kathryn.davi...@googlemail.com wrote:
>
>>>> > > On Jul 5, 4:33 pm, bulrushmo...@gmail.com wrote:
>
>>>> > > > On Jul 4, 8:41 am, kathryn.davi...@googlemail.com wrote:
>
>>>> > > > > Hi
>
>>>> > > > > I am extremely new to IDL (2 weeks!) and have previously only used
>>>> > > > > envi on a small scale.
>
>>>> > > > > I want to read one SD dataset from from a each of a huge number of
>>>> > > > > MODIS files and having looked at IDL and envi batch routines can't
>>>> > > > > decide which is the best way. Bear in mind my limited knowledge and a
>>>> > > > > very short timeframe.... Should I write an envi batch programme and
>>>> > > > > create a big (3000bands +) envi file or should I put straight into the
>>>> > > > > an IDL array. I need to get a mean value (one image or array) and
>>>> > > > > even if it is easier in envi batch mode, would the routine
>>>> > > > > ENVI_SUM_DATA_DOIT with the Mean option deal with the missing
>>>> > > > > values???
>
>>>> > > > > Looking at IDL I have managed to open HDF file from command line, read
>>>> > > > > in appropriate data set to an array but how then could I build 3D
>>>> > > > > array from absolutely loads of 2D arrays.
>
>>>> > > > > Big questions I know - I am desperate to do this in a short time.
>
>>>> > > > > Any help on any aspect much appreciated.
>
>>>> > > > > K
>

>>>> > > > Tell me more about how many bands you have in HDF file and how many
>>>> > > > bands you want to read into IDL?- Hide quoted text -
>
>>>> > > > - Show quoted text -
>
>>>> > > Well I am going to be using around 3-4000 MODIS HDF files but I only
>>>> > > want one band (the first) from each i.e. the Land Surface
>>>> > > Temperature. Since my last post I have thought about creating a huge
>>>> > > multiband file in ENVI and then exporting as a variable to IDL (if the
>>>> > > ENVI_SUM_DOIT doesn't work for the mean, as it may not deal with
>>>> > > missing values very well, I need them to not be counted as opposed to
>>>> > > counting as zero). However that means extracting the SD dataset from
>>>> > > all of the HDF files, converting them to ENVI standard files to build
>>>> > > multi-band image. I hope the data values are not corrupted by being
>>>> > > converted to ENVI standard. Also I could create an image stack in
>>>> > > iIMAGE or mess about with iDataManager in some way but they do not
>>>> > > seem to like reading ENVI standard files and keep asking me to fill in
>>>> > > binary information - will the data values still be OK?
>
>>>> > > Thanks
>
>>>> > > Kathryn- Hide quoted text -
>
>>>> > > - Show quoted text -
>
>>>> > The simplest way to do it:
>>>> > I am assuming you have IDL and ENVI, initiate batch mode by doing the
>>>> > following
>
>>>> > 1. define the file directory
>>>> > 2. read them into IDL using envi_open_data_file
>>>> > 3. get their mean by
>>>> > 4. print them into a txt file
>
>>>> > Try this code
>
>>>> > Pro Mean_HDF
>>>> > envi, /restore_base_save_files
>>>> > envi_batch_init, log_file='batch.txt'
>
>>>> > ; Open the file directory and searh for HDF files to read, then
>>>> > select the directory manually
>>>> > files=file_search(dialog_pickfile(/dir),'*.HDF', count=numFiles);
>>>> > or you can use files=file_search('D:\MODIS*.hdf', count=numFiles)
>
>>>> > ; loop for the whole data set in the directory
>>>> > FOR K = 0, numFiles-1 do begin
>>>> > ; get the file name only without file directory for final

>>>> > output filename
>>>> > fname = file_basename(files[K])
>>>> > ;select input file directory to subset
>>>> > hdf_bands = 1   ; determines the HDF dataset bands to read
>
>>>> > ;start looping through opening bands from HDF
>>>> > for i = 0, hdf_bands -1 do begin
>>>> > envi_open_data_file, files[K], r_fid=fid, /hdf_sd,
>>>> > hdfsd_dataset=i, hdfsd_interleave=0
>
>>>> > ;query new file for ns, nl, dims;
>>>> > envi_file_query, fid, dims=dims, bnames=bnames, ns=ns,
>>>> > nl=nl, nb=nb
>>>> > pos=0
>
>>>> > ;get the mean of the data
>>>> > result = MEAN(fid)
>>>> > endfor
>>>> > ;if you want to export the results in screen do as
>>>> > print, results
>>>> > ;if you want to export them into a txt file
>
>>>> > OpenW, Lun, 'D:\test.txt', /get_lun
>>>> > str= fname
>>>> > printf,lun,str
>>>> > endFOR
>
>>>> > End- Hide quoted text -
>
>>>> > - Show quoted text -
>
>>>> I wonder if you are trying to get mean of each band you read or the
>>>> mean of thousands of bands over each pixel.
>>>> If you want to read just band you can get rid of the inside loop from
>>>> above code. Let me know I will help you figure out.- Hide quoted text -
>
>>>> - Show quoted text -
>
>>> If you are looking making a mean of all of the data bands you read,
>>> try the following
>
>>> Pro Mean_HDF
>>> envi, /restore_base_save_files
>>> envi_batch_init, log_file='batch.txt'
>
>>> ; Open the file directory and searh for HDF files to read, then
>>> select the directory manually
>>> files=file_search(dialog_pickfile(/dir),'*.HDF', count=numFiles);

```
>
>>>     out_fid = lonarr(numFiles)
>>>     ; loop for the whole data set in the directory
>>>     FOR i = 0, numFiles-1 do begin
>>>        ; get the file name only without file directory for final
>>>  output filename
>>>        fname  = file_basename(files[i])
>>>        ;select input file directory to subset
>>>        hdf_bands = 1   ; determines the HDF dataset bands to read
>
>>>        ;start looping through opening bands from HDF
>>>        envi_open_data_file, files[i], r_fid=fid, /hdf_sd,
>>>  hdfsd_dataset=1, hdfsd_interleave=0
>>>        ;query new file for ns, nl, dims;
>>>        envi_file_query, fid, dims=dims, bnames=bnames, ns=ns, nl=nl,
>>>  nb=nb
>>>           pos=0
>
>>>       out_fid[i]=fid
>>>     endFOR
>>>     ; Set the keywords to process all the
>>>     ; spectral data.
>>>     ; Set the keyword COMPUTE_FLAG to
>>>     ; compute the sum of the bands, the
>>>     ; sum squared of the bands, the mean
>>>     ; of the bands, and the standard
>>>     ; deviation of the bands.
>>>    out_pos = lonarr(numFiles)
>>>   envi_file_query, fid, dims=dims, nb=nb
>>>   out_name = 'Mean.img'
>>>   compute_flag = [1,1,1,1,0,0,0,0]
>>>   ;
>>>   ; Call the processing routine to
>>>   ; sum the data together.
>>>   ;
>>>   envi_doit, 'envi_sum_data_doit', $
>>>     fid=out_fid, pos=out_pos, dims=dims, $
>>>     out_name=out_name, compute_flag=compute_flag
>>>   ;
>>>   ; Exit ENVI
>>>   ;
>>>   envi_batch_exit
>
>>>  End- Hide quoted text -
>
>>>  - Show quoted text -
>
>>  Hi
```

>
>> I am trying to get the mean of thousands of bands over each pixel.
>> Many thanks for the code I will give it a go tonight when I get some
>> time to play with it. My only worry is whether it will deal with the
>> missing values - I will let you know.
>
>> Many, many thanks
>
>> Kathryn- Hide quoted text -
>
>> - Show quoted text -
>
> Missing values? you mean the background values as Zero or not? If yes,
> no problem, but if there is missing balues even in the data, it would
> be a little tricky.- Hide quoted text -
>
> - Show quoted text -

The metafiles says the fill value is zero where there is no value
(i.e. a cloudy pixel and therefore no temperature reading) so I hope
that is OK - obviously I want the mean to not count in the zero value
so the number the total value for the pixels needs to be divided by to
obtain the mean needs to be adjusted when there is a non-valid value
i.e. zero - do you think it will work? I can test it on a few files.

I am using MODIS daily files and I do have another complication
however as I have realised I need to use two datasets from each HDF
file not one i.e. the first dataset (daytime temperature) and the
fifth (nighttime temperature)to get the daily mean temperature but
again if one of the two values is zero I don't want the value to be
calculated, the pixel in the resulting daily image needs to be zero so
it wont be counted as if only one temperature (day or night) is there
the daily value can't be calculated. Does that make sense....

Cheers

Kathryn