

# **ALMA Memo No. xxx**

## **ALMA FITS header keywords: a study from the archive User perspective.**

**Authors: Elisabetta Liuzzo, Marcella Massardi, Kazi L. J. Rygl, F. Stoehr, A. Giannetti, S. Burkutean, A. Richards, M. Bonato, M. Lacy, J. Brand**

### **Summary**

ALMA products are stored in the Science Archive in the form of FITS images. It is a common idea that the FITS image headers should collect in their keywords all the information that an archive User might want to search for in order to quickly select a dataset, compare several datasets, or discard useless datasets.

With this perspective in mind, we first present a short description of the current! status of the ALMA FITS archive and images. We realized that at the moment most of the parameters that could be useful for a general User are still missing. We then provide a python code generating the image header keywords that we suggest to be relevant for the scientific exploitation of the ALMA archival data. The proposed tool could be applied to several types of interferometer data. An example of the scientific application of the keywords is also discussed.

In the following, the organization of the Memo is described:

### **Contents**

- 1. Scope**
- 2. Current status of ALMA Science Archive**
  - 2.1 Current status of ALMA FITS image header
- 3. FITS keywords for ALMA Archive purposes**
  - 3.1 Data acquisition and reduction keywords
  - 3.2 Image description keywords
- 4. The ALMA Keywords Filler project**
- 5. The AKF web interface**
- 6. Science User cases for archive FITS keywords**

### **Appendix**

1. Recommended list of keywords
2. Example of the FITS header

## 1. SCOPE

Images (maps and cubes) in the ALMA Science Archive (ASA) are stored in the Flexible Image Transport System (FITS) format which is the standard archival data format for astronomical data sets providing a general way to encode both a definition of the data in the header and the data themselves in an operating system independent format.

In this document, we focus on the FITS headers. It is commonly accepted that a general image header should fully describe the image itself summarizing its content, the production process (i.e., data acquisition and calibration), and additional information to make it accessible to the common image analysis tools. For archived images, the FITS header should then be a repository of this information. For each ASA image, all the necessary details of the observation and reduction process that generated it (including the raw data, the calibration tables and the script that generated it) are accessible through the ASA itself. However, to date, to infer them it is typically necessary to download all the raw data, run the calibration scripts with the proper Common Astronomy Software Applications (CASA<sup>1</sup>) version and consult the calibrated ms data.

Moreover to extract other quantities, it is necessary to open the image and measure them. All these processes could be tedious and the results could be inhomogeneous if they have to be done on a large number of images.

We foresee an ASA that collects images with informative FITS headers that (even if not due to be science-ready by themselves) are fully representative of the content and the processes applied to the data and provide a satisfactory indication of the quality of the achievable results.

In this perspective *the FITS image header should collect all the information that an archive User might want to search for in order to quickly select a dataset, compare several datasets, or discard useless datasets.*

Hence, in addition to the established keywords with a well defined meaning identified by the FITS Standard 3.0 documents (Pence et al. 2010), and soon to be updated to the 4.0 version (see [https://fits.gsfc.nasa.gov/fits\\_standard.html](https://fits.gsfc.nasa.gov/fits_standard.html)), we suggest a list of new keywords that, on the basis of our experience, could be useful to the general ASA User for purposes like those described above. In order to drive our suggestions for a comprehensive FITS header we have investigated some scientific applications that could benefit from an easily accessible, fully documented keyword set.

We thus present in Sect. 3 the “ALMA Keywords Filler” (AKF), tool that generates the image header keywords presented in the current document. Their values could be either obtained from the image itself or from the calibrated ms according to the given keyword definitions. Finally, we provide the updated list (Appendix I) of recommended keywords for the ASA FITS image header, including those identified by the FITS standard 4.0 and those suggested in the present document, and an example of the resulting FITS header (Appendix II). We caveat the reader that other keywords could be present in the FITS images in addition to such a list as result of the process of data reduction for data handling purposes and we do not at any stage recommend to remove or overwrite them.

1 <https://casa.nrao.edu>

## **2. Current status of ALMA Science archive**

The structure of the data trees stored in the archive reflects the project processing structure. An ALMA project is split into Science Goals (SG, the minimum observation settings and targets to reach a scientific purpose), each of which is translated at the observing stage into a Group Units and split into Member Observing Units (MOUs) separating the different settings of the array, each of which is translated into code instruction to the array called Scheduling Blocks (SBs). In order to maximize the efficiency of the telescope dynamical schedule and, as a consequence of the observations, SB are limited in time and repeated as many times as needed to reach the PI requested sensitivity and resolution: each SB repetition is called Execution Block (EBs) and it constitute an independent observing run enclosing its own calibrating source observations. Hence, each analyst should calibrate each EB of a given MOU and combine them all to produce the product images to estimate the reached sensitivity and angular resolution for a given observational setting in a SG: the quality assessment definition works at this level.

In ALMA, in fact, a layered quality assessment (QA) process is applied to all the datasets: after checking the optimal telescope conditions for the data to be taken and stored on the telescope site (QA0 and QA1), the data are fully calibrated and a minimal imaging is performed to verify that the ! resolution and sensitivity requested by the PI are reached (QA2): in case of negative response additional EB are observed (if possible), otherwise, they are delivered to the PI either as “QA2-pass” or “QA2-semipass”. Please notice that for manual assessors imaging of at least for one source in continuum and in one spectral line is requested, so that any additional produced image is done on the assessor goodwill/time/capability. Product images are not intended to be science ready (as calibrated data are), and it is expected that PIs or archive miners use them only as indicators of data quality. In the archive, raw data for each of the >6000 EBs observed so far are stored and properly linked to the project tree they belong to. Scripts to calibrate each EB resultant of the (manual- or, more recently pipeline-based) QA, as well as preliminary scripts for imaging of the whole MOU are stored as well, together with the images produced during the QA and a set of calibration diagnostic plots. Cycle 0 data constitute an exception, as only the raw data are stored and they are not organized according to the above described data tree.

Currently, an archive miner can choose for each MOU whether to download only the products (scripts, images and diagnostic plots), with a typical size of a few 100MB and/or the raw data that could reach size of several 100GB depending on the observing settings (number of antennas, frequency channels, Ebs,...). The data download might thus take several hours to days, which for some users might be difficult because of server timeouts at their institutes. In order to complement or re-produce the product FITS images, once all the data are download, calibration scripts should be run with the same CASA version that the analyst used to generate them (also this process might require hours to days). Only then, the User should use (and frequently improve or adjust) the distributed imaging scripts to produce the FITS images they need, sometimes to discover that the target they were looking for is undetected or for any reason, not observed.

## 2.1 Current status of ALMA FITS image header

In the following, we show one example of the FITS header of an archival ALMA FITS image.

### *ALMA FITS header*

```
SIMPLE =          T /Standard FITS
BITPIX =         -32 /Floating point (32 bit)
NAXIS =           4
NAXIS1 =          84
NAXIS2 =          80
NAXIS3 =         3772
NAXIS4 =           1
EXTEND =          T
BSCALE = 1.000000000000E+00 /PHYSICAL = PIXEL*BSCALE + BZERO
BZERO = 0.000000000000E+00
BMAJ = 1.206100316999E-03
BMIN = 7.315292685941E-04
BPA = -7.971664204539E+01
BTYPE = 'Intensity'
OBJECT = 'pi_Gru '
BUNIT = 'Jy/beam ' /Brightness (pixel) unit
RADESYS = 'ICRS '
LONPOLE = 1.800000000000E+02
LATPOLE = -4.594799027778E+01
PC01_01 = 1.000000000000E+00
PC02_01 = 0.000000000000E+00
PC03_01 = 0.000000000000E+00
PC04_01 = 0.000000000000E+00
PC01_02 = 0.000000000000E+00
PC02_02 = 1.000000000000E+00
PC03_02 = 0.000000000000E+00
PC04_02 = 0.000000000000E+00
PC01_03 = 0.000000000000E+00
PC02_03 = 0.000000000000E+00
PC03_03 = 1.000000000000E+00
PC04_03 = 0.000000000000E+00
PC01_04 = 0.000000000000E+00
PC02_04 = 0.000000000000E+00
PC03_04 = 0.000000000000E+00
PC04_04 = 1.000000000000E+00
CTYPE1 = 'RA---SIN'
CRVAL1 = 3.356843454167E+02
CDELTA1 = -2.166666666667E-04
CRPIX1 = 4.300000000000E+01
CUNIT1 = 'deg '
CTYPE2 = 'DEC--SIN'
CRVAL2 = -4.594799027778E+01
CDELTA2 = 2.166666666667E-04
CRPIX2 = 4.100000000000E+01
CUNIT2 = 'deg '
CTYPE3 = 'FREQ '
CRVAL3 = 3.423942288270E+11
CDELTA3 = 4.883342805786E+05
CRPIX3 = 1.000000000000E+00
CUNIT3 = 'Hz '
CTYPE4 = 'STOKES '
CRVAL4 = 1.000000000000E+00
CDELTA4 = 1.000000000000E+00
CRPIX4 = 1.000000000000E+00
CUNIT4 = ' '
PV2_1 = 0.000000000000E+00
PV2_2 = 0.000000000000E+00
RESTFRQ = 3.433000000000E+11 /Rest Frequency (Hz)
SPECSYS = 'LSRK ' /Spectral reference frame
ALTRVAL = 7.909798029106E+05 /Alternate frequency reference value
```

```

ALTRPIX = 1.000000000000E+00 /Alternate frequency reference pixel
VELREF = 257 /1 LSR, 2 HEL, 3 OBS, +256 Radio
COMMENT casacore non-standard usage: 4 LSD, 5 GEO, 6 SOU, 7 GAL
TELESCOP='ALMA '
OBSERVER='srams '
DATE-OBS='2013-10-08T02:46:00.960000'
TIMESYS='UTC '
OBSRA = 3.356843454167E+02
OBSDEC = -4.594799027778E+01
OBSGEO-X= 2.225142180269E+06
OBSGEO-Y= -5.440307370349E+06
OBSGEO-Z= -2.481029851874E+06
INSTRUME='ALMA '
OBJECT='pi_Gru '
TELESCOP='ALMA '
DISTANCE= 0.000000000000E+00
FIELD='pi_Gru '
ITER = 1
SPW = '18 '
TYPE = 'pbcorimage'
DATE = '2016-09-16T12:02:40.934000' /Date FITS file was written
ORIGIN = 'CASA 4.7.38256-REL (r38256)'
HISTORY CASA START LOGTABLE
HISTORY 2016-09-16T10:57:12 INFO SRCCODE='::setmiscinfo'
HISTORY Ran ia.setmiscinfo
HISTORY 2016-09-16T10:57:12 INFO SRCCODE='::setmiscinfo'
HISTORY ia.setmiscinfo(info={...})
HISTORY CASA END LOGTABLE
END

```

Investigating different ALMA FITS images, we noticed that their FITS headers that are currently stored in the science archive contain keywords that can change on the basis of the CASA version or the procedures that produced them.

Moreover, a lot of keywords lack a clear definition in the FITS header documentation (see orange lines). Other keywords appear to be redundant (see yellow lines).

Some additional keywords might be specifically requested by external analysis tools or to deal with data in other spectral bands. In this document, we focus mostly on those that, according to our experience, could be of use for general radio to sub-mm interferometric scientific purposes.

### 3. FITS Keywords

In the light of the current status of the ALMA archive, we describe in the present section the keywords that we suggest to be useful to the general ASA User in addition to those identified by the FITS Standard v. 4.0 document providing for all of them a appropriate definition.

We classified two main categories of such keywords:

- *data acquisition and reduction keywords*: these refer to the properties of the telescope during the observation and to the calibration process. Their values have to be extracted from the calibrated measurement sets (typically browsing the data tables), the calibration products or the proposal documents
- *image description keywords*: these refer to the properties of the FITS image that hosts them. Their values could be extracted and calculated according to the definitions given below.

#### 3.1 Data acquisition and reduction keywords

##### RA\_TARG and DEC\_TARG

Description: RA and DEC of target as listed in the ms.

Units: degrees

Type: float

Notes: in the presence of multiple ms with different target coordinates or mosaic, a list of RA and DEC is given

##### UVRANGE

Description: Median, first, and third quartile of the UV length distribution

Units: float

Type: kilowavelengths

Notes: in the presence of multiple ms, the total UV length distribution is the sum of each ms UV length distribution

##### BAND

Description: The receiver band used for the image

Values: 'BX', where X= integer 3-9

Type: character

##### MINPRBL and MAXPRBL

Description: Minimum and maximum projected baseline

Units: m

Type: float

Notes: in the case of multiple ms, it calculates the minimum and maximum projected baseline of each ms and takes the minimum and maximum value among them. In the case of multiple arrays, the minimum and maximum among the array is reported.

##### MAXANGSC

Description: The maximum angular scale resolved by the 12m, and 7m array

Units: arcseconds

Type: float

Notes: it is calculated as  $37100/((\text{BNDCTR} \cdot 10^{-9}) \cdot \text{MINPRBL})$ . In the case of multiple arrays, the MAXANGSC of each array is listed.

### PADLIST

Description: list of ALMA pads contributing to the data

Type: character

Notes: Antennas with all data flagged are discarded. In the case of multiple ms, the final list contains all the pads, even if some are present only in one of the ms. In the case of multiple arrays, all the pads of each array is listed.

### NANT12M, NANT7M, NANTTP

Description: Number of ALMA 12m, 7m, and ACA total power antennas contributing to the data.

Type: integer

Notes: In the case of multiple ms, the maximum number of antennas among each single ms is taken. Antennas with all data flagged are not considered.

### MINELT12, MINEL7, MINELTP

Description: Minimum Elevation range achieved during observations of target 12m, 7m and TP data.

Units: degrees

Type: float

Notes: in the case of multiple ms, it calculates the minimum elevation of each ms and takes the minimum value among them. Only data completely not flagged are taken.

### DATE-OBS, DATE-END

Description: Observation start and end time

Units: hh:mm:ss

Type: character

Notes: it calculates minimum and maximum observing time (min\_obst and max\_obst)

(DATE-OBS = min\_obst - 0.5\*int, DATE-END = max\_obst + 0.5\* int where int is the integration time)

### MJD-OBS, MJD-AVG

Description: Modified Julian Date (JD - 2,400,000.5) of start and mid-point of the observations

Type: float with a format F5.5

## 3.2 Image description keywords

### BNDRES

Description: Frequency resolution of the image defined as increment of frequency axis

Units: Hz

Type: float

Notes: overlap with CDEL (Ctype = FREQ)

### CHNRMS

Description: the inter-quartile range of the pixel values in the image

Units: Jy/beam

Type: float

Notes: in the case of spectral line data, the inter-quartile range is taken considering the pixel values of all the channels together. In the case of polarization data, one value for each stokes is provided.

## SPATRES

Description: geometric average of the min and the max beam axes.

Units: arcsecond

Type: float

Notes: For spectral line and polarization data, it loops over the channel and stokes axes, takes the mean of bmin and bmax among all channel and stokes values and then it calculates the geometric average of bmin and bmax.

## STOKES

Description: list of image stokes parameters

Value: either "I" or "I, Q, U, V"

Type: character

Notes: it does not cope with the case of shape (nstokes) = 2 or 3.

## DATAMIN and DATAMAX

Description: minimum and maximum values in the image

Units: Jy/beam

Type: float

Notes: in the case of spectral line data, the minimum and the maximum among all the channel is taken. In the case of polarization data, one value is given for each stokes.

## DYNRANGE

Description: Estimation of image dynamic range for each stokes defined as  
DATA MAX /CHANRMS

Type: float

Notes: For polarization data, one value is provided for each stokes. As one estimation of CHANRMS (see its definition above) is given for all channel, one value of DYNRANGE for each channel is given.

## BNDCTR

Description: the center frequency of the image calculated as  $\frac{1}{2}$  (the reference value for the frequency axis \* number of channels) .

Units: Hz

Type: float

Notes: none

## BNDWID

Description: the effective bandwidth of the image calculated as the increment in the frequency axis \* number of channels

Units: Hz

Type: float

Notes: none

## FLUXTOT

Description: integrated flux of the source obtained masking the image below 3 \* CHANRMS

Units: Jy

Type: float

Notes: For spectral line data, the sum of pixels with flux above 3\*CHANRMS of each channel is taken. In the case of polarization data, one value is given for each stokes.



#### 4. The ALMA Keywords Filler

The ALMA Keywords Filler (AKF) is a Python-based script that exploits existing CASA tasks and toolkit<sup>2</sup>.

It consists of two codes (that could eventually be called in sequence), each of which addresses one of the keyword categories described in Sect. 3 and ingests the resulting keywords in the header of an input image.

1. **AKF\_v1.0\_ms.py**: it generates the keywords listed in Sect. 3.1. It requires that (only) the ms from which the input FITS image was produced is in the same folder as the image. Since the program browses the ms tables, it can take some minutes to run (and up to ~30 min for the biggest datasets as large as 100 Gb). However, it is possible to request the production of only specific keywords.
2. **AKF\_v1.0\_fits.py**: it generates the keywords listed in Sect. 3.2. It calculates the values from the image itself and it is usually faster than the ms code. Also in this case it is possible to produce all or a selection of the keywords.

The result of the codes is a Python dictionary as reported in the Fig. 1.

The code could be recovered from the Italian ALMA Regional Center web page ([http://www.alma.inaf.it/index.php/ALMA\\_FITS\\_Keywords](http://www.alma.inaf.it/index.php/ALMA_FITS_Keywords))

In the following, the instructions to launch the codes are described:

- enter the folder (yourDir) where your fits files and ms are stored
- launch CASA (any version > 5.0)
- define the image to process  
imageName = 'yourDir/yourImage.fits'
- define the list of keywords you want calculate  
kwlist = ['KWDi', 'KWDj',....., 'KWDk'] where KWDi,j,k are one or more keywords belong to the same category, either data acquisition or image description keywords
- launch the correct code according to the requested category of keywords with  
execfile('AKF\_v1.0\_ms.py') and/or execfile('AKF\_v1.0\_fits.py')

```
{'BNDRES': 244201.59716800001, 'BNDWID': 348231477.56156802, 'BNDCTR': 3130392757  
53.78076, 'STOKES': 'I', 'SPATRES': '0.24', 'CHANRMS': {'YY': None, 'I': 0.009886  
0654979944229, 'Q': None, 'XX': None, 'U': None, 'V': None}, 'DYNRANGE': {'YY': N  
one, 'I': 46.003972084247678, 'Q': None, 'XX': None, 'U': None, 'V': None}, 'DATA  
MAX': {'YY': None, 'I': 0.45479828119277954, 'Q': None, 'XX': None, 'U': None, 'V  
' : None}, 'FLUXTOT': {'YY': None, 'I': 4283.909631695642, 'Q': None, 'XX': None,  
'U': None, 'V': None}, 'DATAMIN': {'YY': None, 'I': -0.31712988018989563, 'Q': No  
ne, 'XX': None, 'U': None, 'V': None}}
```

Fig. 1: Example of the dictionary provided by the AKF\_v1.0\_fits.py where kwlist contains all the implemented keywords up to now.

<sup>2</sup> <https://casa.nrao.edu/docs/CasaRef/CasaRef.html>

## 5. The AKF web interface for image-description keywords

The AKF code is implemented in the Keywords of Astronomical FITS-Images Explorer (KAFE - Burkutean et al. in prep), a Python tool suite that exploits CASA tasks and toolkit.

Even if originally conceived in the framework of our ALMA activities, KAFE is well suited to analyze FITS images produced by most of the major radio to submm interferometric facilities.

In particular, KAFE offers a web interface (see Fig. 2) that allows:

- to fill the header of the input FITS images with products of the image post-processing (e.g. total flux, dynamic range, rms) applying the image-description keywords as a part of the AKF code. This offers quick-look catalogues, and a fast, homogeneous and coherent comparison tool for image sets (see Sect. 6 for scientific applications);
- to provide advanced diagnostic plots (e.g. moments maps, spectrum, polarization vectorial images, SEDs, light curves ...);
- to cross-match the image with the most widely used astronomical catalogues and databases;
- to exploit some visualization tools (e.g. 3-colour images, Mollweide representation).

Allowing a quick comparison of the data contents in the form of advanced image metadata as well as diagnostic plots, KAFE is then suitable for all the major data mining purposes.

The screenshot displays the KAFE web interface with a teal background. At the top, there is a file upload section with the text "send this file:" followed by a "Browse..." button and "No files selected." Below this is a "filters" section containing six input fields for range settings: POSRANGE, CNTRFREQ RANGE, FREQRANGE, ANGRESRANGE, CHANRMS RANGE, and FLUX TOTAL RANGE, each with a "0" value. The bottom section is divided into three columns: "requested keywords", "spectrum analysis options", and "further analysis options". Each column contains a list of options with checkboxes, all of which are currently unchecked.

filters	
POSRANGE	0
CNTRFREQ RANGE	0
FREQRANGE	0
ANGRESRANGE	0
CHANRMS RANGE	0
FLUX TOTAL RANGE	0

requested keywords	spectrum analysis options	further analysis options
ALL	ALL	ALL
RA_centre	3D view	Source detection
DEC_centre	continuum subtraction	Source detection SNR layer
SPATRES	Channel gallery	radial average
BNDCTR	moments	Image cuts
BNDRES	Spectrum inner quarter	power spectrum
BNDWID	Spectrum around max	Light curve
CHANRMS	Spectrum slabs	Polarization maps
DYNRANGE	Spectral fit	
FLUXTOT	Cube morph	
DATAMAX		
DATAMIN		
STOKES		

Fig. 2: a screenshot of the KAFE web interface is presented.

The first release of the KAFE implements only the AKF code to derive keywords related to the images (i.e. AKF\_v1.0\_fits.py) . In later versions, also the keywords related to the ms properties will be provided.

## **6. Science Cases for use of keywords in archived FITS images**

As astronomy continues to move towards multi-wavelength, data-driven science, issues of data provenance become of vital concern. Many future User cases for ALMA data will involve the download of FITS images from the archive through protocols such as the Virtual Observatory, where the User will receive the data file with essentially no other piece of information. To make the best possible use of such data in a publication, the metadata in the file must contain sufficient provenance information to permit the replication of the published results, sufficient characterization to allow meaningful statements about the nature of the observations, and sufficient attribution to guarantee that organizations and individuals are given due credit.

FITS is the only data standard used commonly in all fields of astronomy. Thus, the most convenient way to supply these data is via keyword-value pairs in the FITS header. The use of this technology ensures that such data can be easily machine readable, for example into databases and/or Python dictionaries. In addition to that, keywords values could be used for image selection, comparison! and statistical analysis with direct scientific exploitation.

Here we summarize some of the most relevant examples of archived FITS image keywords exploitation, and how they could benefit of the AKF tool.

The AKF codes are particularly useful to compare image products, for example to compare image quality or identify the images to be selected for User scientific purposes. It is important to stress that to exploit AKF for publications we recommend to apply the codes to calibrated data (tagged as science-ready) or re-imaged according to the User's needs.

Some science application examples for the image-description part of the code are presented in Burkutean et al. (in prep.). Here we present a science case that exploits also the data acquisition keywords: the case of light curves and source variability analysis.

Light curves show the brightness of an object over a period of time and they are a simple, but useful tool to understand processes at work within the objects, such as novae, supernovae, variable stars but also blazars, or extragalactic objects in general.

To construct them, it is necessary to retrieve an adequate number of images, typically through archive mining processes, to derive the brightness or flux over time. The AKF code is able to do this in an automatic way for a huge amount of observations in a short time through the calculation of the FLUXTOT and MJD-AVG keywords.

For example, for a proper characterization of blazar population properties, the analysis of blazars light curves on statistical significant samples are mandatory.

Blazars, as a class of Active Galactic Nuclei (Urry & Padovanni 1995), are characterized by strong variability in all wave-bands. Additionally, a broad range of variability time scales is observed, ranging from minutes, as in the cases of PKS 2155-304 (Aharonian et al. 2007) and PKS 1222+216 (Tavecchio et al. 2011), to months (e.g. Abdo et al. 2010). In particular, the very short variability time scales are puzzling, since their emission should be generated in emitting regions much smaller than the event horizon of the AGN black hole, which, instead, should be the lower limit on the jet width size.

Blazars emitting at high energy are peculiar sources for which the emission mechanism and site of the

gamma-ray signal are not fully understood. In this case, the cross-correlation of the radio/millimeter emission with the gamma-ray one is crucial to test the particle acceleration models and the origin of the high energy emission.

The high energy emission and the erratic, rapid and large-amplitude variability observed in all accessible spectral regimes (radio-to-gamma-ray) are two of the main defining properties of blazars (e.g. Ulrich et al. 1997; Webb 2006). The entire non-thermal continuum is believed to originate mainly in a relativistic jet, pointing close to our line of sight. Studies of variability in different spectral bands and correlations of multi-waveband variability patterns allow us to shed light on the physical processes in action in blazars, such as particle acceleration and emission mechanisms, relativistic beaming, origin of flares and size, structure and location of the emitting regions that could be complemented with the high resolution- sensitivity ALMA maps. In Fig. 3 we report an example of light curves derived using the AKF code for the blazar source J0635-7516.

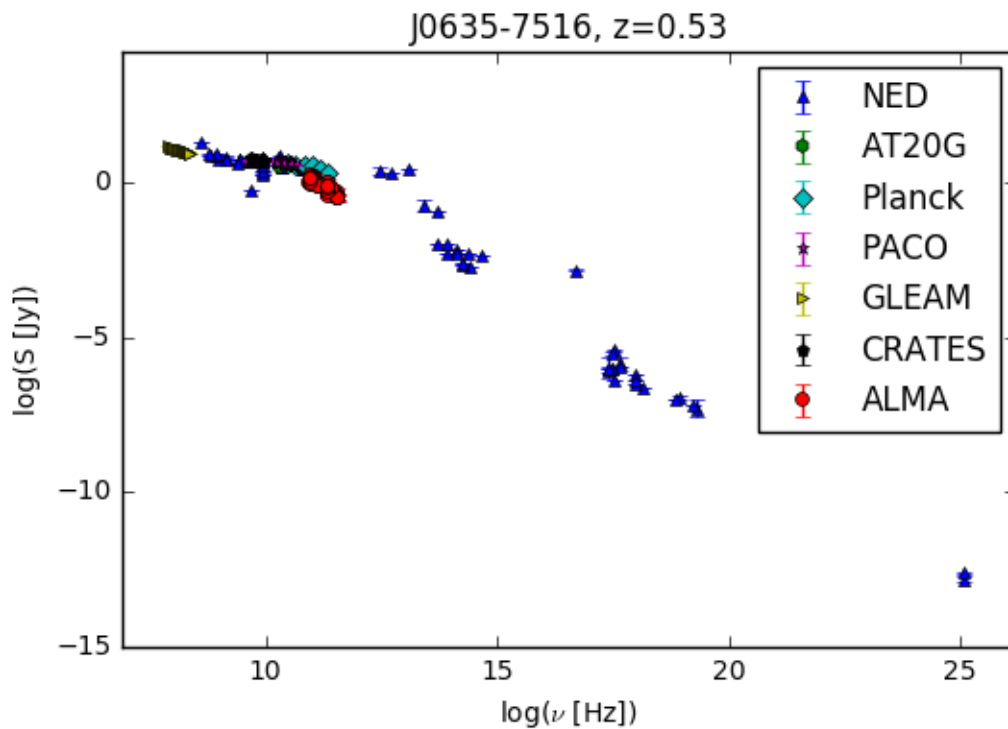


Fig. 3: Example of blazar light curve obtained using the AKF codes: the case of J0635-7516 (Bonato et al. in prep.)

## Appendix 1 - Recommended list of keywords

Here, we list first the list of the existing FITS Standard 4.0 keywords and then the keywords we suggest will be useful to include in the metadata of images to be suitable for a general archive User.

In blue are the keywords already provided by the AKF codes or by the FITS Standard 4.0 or already present in ALMA FITS headers with clear definition.

In red we mark those that are under construction by the AKF project (and will be distributed in future versions of the code) and those for which the information is currently not derivable directly by the FITS or ms files (and we expect might be included in future images by future versions of AKF or of the ALMA imaging pipeline). For the latter we suggest a possible definition, that might be improved during their development.

### \* Existing FITS Standard 4.0

SIMPLE  
BITPIX  
NAXIS  
NAXISn  
PC##i###j  
Pvi\_m  
CTYPEn, CRVALn, CDELTn, CRPIXn, CUNITn  
RADESYS  
SPECSYS  
BSCALE, BZERO, BTYPE  
CHECKSUM  
OBSGEO-X, OBSGEO-Y, OBSGEO-Z  
LONPOLE, LATPOLE  
TELESCOP  
OBJECT  
EQUINOX  
END (no associated value)

### \* Data acquisition and reduction keywords

AKF:

RA\_TARG  
DEC\_TARG  
MOS\_TARG  
UVRANGE  
BAND  
MINPRBL, MAXPRBL  
MAXANGSCALE  
PADLIST  
NANT12M, NANT7M, NANTTP  
MINELT12, MINEL7, MINELTP  
DATE-OBS, DATE-END  
MJD-OBS, MJD-AVG  
**EXPTIM12, EXPTIM7, EXPTIMTP**: exposure time spent on source for each of 12m, 7m and TP array (currently under construction in AKF)  
**FOV**: field of view size for each array  
**PROJCODE**: identifier code for the project the image belongs to  
**CASAVR**: CASA versions used for calibration and imaging  
**GROUP**: id of the group observing unit set for which the image is one of the products  
**MEMBER**: id of the member observing unit set for which the image is one of the products

**SGOAL:** id of the science goal for which the image is one of the products  
**SBNAMES:** id of the scheduling blocks for which the image is one of the products  
**QA2FLAG:** flag given to the member observing unit set during the quality assessment 2 stage

\* Image description keywords

AKF  
BNDRES  
BNDCTR  
BNDWID  
CHNRMS  
FLXTOT  
SPATRES  
DYNRANGE  
STOKES  
DATAMAX, DATAMIN  
BMAJ,BMIN,BPA

**Appendix 2 – Example of suitable FITS header**

**EXISTING FITS STANDARD 4.0 KEYWORDS**

SIMPLE = T	/ Standard FITS
BITPIX = -32	/ Floating point (32 bit)
EXTEND =	F
NAXIS = 4	/ Number of axes in the associated data array.
NAXIS1 = 240	/ NAXIS 1 dimension
NAXIS2 = 240	/ NAXIS 2 dimension
NAXIS3 = 1	/ NAXIS 3 dimension
NAXIS4 = 1	/ NAXIS 4 dimension
BSCALE = 1.000000000000E+00	/ PHYSICAL = PIXEL*BSCALE + BZERO
BZERO = 0.000000000000E+00	/ PHYSICAL = PIXEL*BSCALE + BZERO
BTYPE = 'Intensity'	/ Brightness (pixel) unit
BUNIT = 'JY/BEAM'	/ Physical units in which the quantities in array
EQUINOX = 2.000000000000E+03	/ Equinox of source coordinates
PC001001 = 1.000000000000E+00	/ Transformation matrix terms
PC002001 = 0.000000000000E+00	/ Transformation matrix terms
PC003001 = 0.000000000000E+00	/ Transformation matrix terms
PC004001 = 0.000000000000E+00	/ Transformation matrix terms
PC001002 = 0.000000000000E+00	/ Transformation matrix terms
PC002002 = 1.000000000000E+00	/ Transformation matrix terms
PC003002 = 0.000000000000E+00	/ Transformation matrix terms
PC004002 = 0.000000000000E+00	/ Transformation matrix terms
PC001003 = 0.000000000000E+00	/ Transformation matrix terms
PC002003 = 0.000000000000E+00	/ Transformation matrix terms
PC003003 = 1.000000000000E+00	/ Transformation matrix terms
PC004003 = 0.000000000000E+00	/ Transformation matrix terms
PC001004 = 0.000000000000E+00	/ Transformation matrix terms
PC002004 = 0.000000000000E+00	/ Transformation matrix terms
PC003004 = 0.000000000000E+00	/ Transformation matrix terms
PC004004 = 1.000000000000E+00	/ Transformation matrix terms
PV2_1 = 0.000000000000E+00	/ Parameter value #1 for world coordinate axis #2,
PV2_2 = 0.000000000000E+00	/ Parameter value #2 for world coordinate axis #2,
CTYPE1 = 'RA---SIN'	/ WCS term: type of Axis 1
CRVAL1 = 2.853708750000E+02	/ WCS term: Reference pixel value, axis 1
CDELTA1 = -4.444444444444E-05	/ WCS term: Increment per pixel, axis1
CRPIX1 = 1.210000000000E+02	/ WCS term: Reference pixel number, axis 1
CUNIT1 = 'deg'	/ WCS term: Unit of axis 1

CTYPE2 = 'DEC—SIN' / WCS term: type of Axis 2  
 CRVAL2 = -3.703011111111E+01 / WCS term: Reference pixel value, axis 2  
 CDELTA2 = 4.444444444444E-05 / WCS term: Increment per pixel, axis 2  
 CRPIX2 = 1.210000000000E+02 / WCS term: Reference pixel number, axis 2  
 CUNIT2 = 'deg ' / WCS term: Unit of axis 2  
 CTYPE3 = 'FREQ ' / WCS term: type of Axis 3  
 CRVAL3 = 2.315424966698E+11 / WCS term: Reference pixel value, axis 3  
 CDELTA3 = 3.870856771975E+09 / WCS term: Increment per pixel, axis 3  
 CRPIX3 = 1.000000000000E+00 / WCS term: Reference pixel number, axis 3  
 CUNIT3 = 'Hz ' / WCS term: Unit of axis 3  
 CTYPE4 = 'STOKES ' / WCS term: type of Axis 4  
 CRVAL4 = 1.000000000000E+00 / WCS term: Reference pixel value, axis 4  
 CDELTA4 = 1.000000000000E+00 / WCS term: Increment per pixel, axis 4  
 CRPIX4 = 1.000000000000E+00 / WCS term: Reference pixel number, axis 4  
 CUNIT4 = '' / WCS term: Unit of axis 4  
 RADESYS = 'ICRS ' / Reference system for equatorial coordinates  
 RESTFRQ = 2.315424966698E+11 / Rest Frequency (Hz)  
 SPECSYS = 'LSRK' / Spectral reference frame  
 OBSGEO-X= 2.225142180269E+06 / [m] X-coordinate of observation position wrt Geocentric reference  
 OBSGEO-Y= -5.440307370349E+06 / [m] Y-coordinate of observation position wrt Geocentric reference  
 OBSGEO-Z= -2.481029851874E+06 / [m] Z-coordinate of observation position wrt Geocentric reference  
 LONPOLE = 1.800000000000E+02 / [deg] Long. in native coordinate system of celestial system's north pole  
 LATPOLE = -7.667444444444E-01 / [deg] Lat in native coordinate system of celestial system's north pole  
 END / End of HDU

#### AKF IMAGE KEYWORDS

RA = 2.853708750000E+02 / [deg] Image centre RA  
 DEC = -3.703011111111E+01 / [deg] Image centre Dec  
 OBJECT = 'Mrk590' / target name  
 BMAJ = 2.228875623809E-04 / [arcsec] Restoring beam major axis  
 BMIN = 1.697528362274E-04 / [arcsec] Restoring beam minor axis  
 BPA = 5.713778686523E+01 / [deg] Restoring beam position angle  
 DATAMAX = 1.5 / [Jy/beam] Maximum value in the FITS image!  
 DATAMIN = -.05 / [Jy/beam] Minimum value in the FITS image  
 DYNRANGE = 5.0 / Dynamic range in the image  
 CHANRMS = 0.0003 / [Jy/beam] RMS per channel of FITS image  
 SPATRES = 0.7 / [arcsec] Spatial resolution of the FITS image  
 BNDCTR = 2.315424966698E+11 / [Hz] Center frequency of data in the FITS array!  
 BNDWID = 1.875E+9 / [Hz] Bandwidth of the FITS image  
 BNDRES = 0.488281E+6 / [Hz] Frequency resolution in the FITS image  
 STOKES = 'I' / List of data Stokes parameters!

#### AKF DATA ACQUISITION AND REDUCTION KEYWORDS

TIMESYS = 'UTC ' / Time system for time-related kwds in the HDU!  
 DATE = '2012-10-11T09:27:32.760000' / Date FITS HDU file was written!  
 DATE-OBS = '2012-06-17T05:56:15.792000' / Date and time of start of observations !  
 MJD-OBS = 55927.50000 / Modified Julian Date of start of the observation,!  
 MJD-AVG = 55928.54321 / Modified JD of the mid-point of observation. !  
 DATE-END = '2012-06-18T05:56:15.792000' / Date and time of last observations !  
 RA\_TARG = 2.853708750000E+02 / [deg] PI-defined target RA  
 DEC\_TARG = -3.703011111111E+01 / [deg] PI-defined target Dec  
 BAND='3' / observing band of the observation  
 MAXANGSC = 2000.0 / [arcsec] Maximum ang. scale of the FITS image  
 UVRANGE = [202345, 123456, 345677] / [klambda] Median, 1st and 3rd quartile of the UV length distribution in klambda  
 MINPRBL / [m] Minimum baseline  
 MAXPRBL / [m] Maximum baseline  
 PADLIST / List of ALMA pad names contributing to data  
 NANT12M / Number of ALMA 12 m antennas used in data  
 NANT7M / Number of ALMA ACA 7 m antennas used in data  
 NANTTP / Number of ALMA ACA total power antennas used in data

MINEL12M	/ [deg] Minimum elevation of ALMA 12 m antennas
MINEL7M	/ [deg] Minimum elevation of ALMA ACA 7 m antennas
MINELTP	/ [deg] Min. elevation of ALMA ACA total power antennas
FOV = 0.1234	/ [deg^2] Total field of view of the image!
SPW = 23	/ Spectral window identification number
EXPTIM12	/[s] On-source obs time of ALMA 12 m antennas
EXPTIM7	/[s] On-source obs time of ALMA ACA 7 m antennas
EXPTIMTP	/[s] On-source obs time of ALMA total power antennas

#### ADDITIONAL KEYWORDS

ZSOURCE = 0.000485787	/ Redshift of Source
TELESCOP= 'ALMA '	/ Telescope name
OBSERVER= 'koayjy '	/Alma id of the PI
ORIGIN = 'JAO'	/ Organisation responsible for producing dataset.
PROPCODE = "2011.0.00101.S"	/ ALMA proposal ID
CASAVR = 'CASA 3.4.0 (release r19988)!'	/ Version of CASA used by pipeline to produce FITS data
GROUP = 'uid://A005/X006/X007'	/ Level at which data is combined in this FITS
MEMBER = 'uid://A008/X009/X010'	/ Member observing unit set status ID
SGOAL = 'uid://A009/X010/X011'	/ Science Goal Observing unit set ID
SBNAMES = 'SBNAMES = 'exampleSB1 '	/ Names of scheduling blocks contributing to data in array
QA2FLAG = 'PASS'	/ QA2 flag description

#### References:

Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, ApJ, 722, 520 ;  
 Aharonian, F., Akhperjanian, A. G., Bazer-Bachi, A. R., et al. 2007, ApJL, 664, L71;  
 Pence et al 2010, A&A, 524,  
 Tavecchio, F., Becerra-Gonzalez, J., Ghisellini, G., et al. 2011, A&A, 534, A86;  
 Ulrich, M.-H., Maraschi, L., & Urry, C. M. 1997, ARA&A, 35, 445;  
 Urry, C. M., & Padovani, P. 1995, PASP, 107, 803:  
 Webb, J. R. 2006, Blazar Variability Workshop II: Entering the GLAST Era, 350, 3