Introduction

Into the **MIA-EWS** reduction package I want to introduce some sort of database of calibrators. The motivation is that there are only a fairly limited number of good MIDI calibrators, and accurate information on them, particularly fluxes and spectra, is hard to get. So it would save a lot of time and effort if the reduction package recognized the name or alias of a calibrator and had at least some information ready to present to the poor gradstudent astronomer who heard from her/his supervisor that she/he should use some particular calibrator.

The danger of this approach is of course that if you compile a list of calibrators you are *de facto* certifying the quality of the list, which can get both you and the user into a lot of trouble. For the time being I am ignoring this issue. I am here only trying to specify data formats for storing the calibrator data, so that lists can be exchanged between users, and somewhat later, tools to maintain the lists. My concept is that some heavyweight institutions (e.g. ESO, Keck) might maintain master lists of calibrators, but that individual users should be able to edit their own private copies according to their own taste and experience.

While my main motivation for this exercise was to create a format for calibrator lists for MIDI, it is obvious that the same exercise could be useful for AMBER or other interferometers and indeed for single telescopes too. So the descriptions of formats below can, with small omissions or inclusions, be used in any of these contexts.

I have used a design based on a relational model: the information is spread out over many tables, each containing a very specific type of information, and there are keys or relations used to bind entries in the tables together. This has proven in many contexts to be the best way to ensure consistency and maintainability between entries. This does not mean that the implementation in any particular package would have to use a relational database. The amount of data will probably be relatively small, so that any reasonable method of accessing the data will work. The description of the tables below is in terms of FITS binary tables. This would be the transport format of the information. The local representation of the data on any particular package would be up to the package designer, but there must be tools to create the FITS tables from data in the package, and conversely ways to update the local representation from the FITS tables.

I describe below ten tables:

(1) **Sources:** Gives a standard name for a source, and position information.

(2) **Aliases** Gives a list of alternative names for sources.

(3) **Bands** Describes the meaning of the wavelength bands, e.g. 'K-band'

(4) **Bandpasses** Spectrally dispersed photometry data

(5) **Diameters** Diameters of sources in various bands

(6) **Photometry** Broad band (polarized) photometry data

(7) **Visibility** Broad band visibility data

(8) **Spectrophotometry** Spectrally dispersed photometry data

(9) **Spectrovisibility** Spectrally dispersed visibility data

(10) **References** Where information came from

The original definitions of each table in terms of keyword and column definitions are specified in XML files. For this document I have translated them automatically into .tex files with a self-crafted tool. The .tex files generate the descriptions given below. In principle similar tools can generate from the .xml files .c, .C, .h, .pro or other such files, which will allow automatic incorporation of the appropriate structures into most existing packages. Jeroen de Jong (jdejong@eso.org) has developed a number of tools to facilitate this process.

The current version is only a draft/suggestion version. I know that other people have been working in this area and may by now have better (developed) models for calibration data. I am certainly eager to see suggestions for improvements, or even to transfer the whole responsibility for this development to someone else who feels the calling. If the politics of agreeing on standard formats becomes too messy, I will simply retreat to the original goal of designing an internal MIDI format.

Table: SOURCES

Describes parameters of calibration sources that are not dependent on band, primarily positional information. This is derived from the Jaffe/Cotton standard which is closely related to the IAU standard, but the actual photometric and/or visibility information has been moved to other tables.

Header Keywords

Keyword | **Type**
REVISION | **INTEGER**

The revision number of table definiton.

ORIGIN | **STRING**

Who created this table

Table Columns

Column | **Format** | **Units**
SOURCE | **16A**

Unique identifier internal to tables. Aliases will be stored in the ALIAS table

RAEPP | **D** | **DEGREE**

Right ascension at mean equinox.

RAEPP_ ERR | **D** | **DEGREE**

Estimated rms uncertainty in RAEPP

DECEPP | **D** | **DEGREE**

Declination at mean equinox.

DECEPP_ ERR | **D** | **DEGREE**

Estimated rms uncertainty in DECEPP

EQUINOX | **A**

Equinox of RAEPP and DECEPP. This should be either B1950.0 or J2000.0

EPOCH | **D**

Epoch of RAEPP and DECEPP. Modified Julian Date.

SYSVEL | **D** | **METER/S**

System velocity of source.

SYSVEL_ ERR | **D** | **METER/S**

Estimated rms uncertainty in SYSVEL.

VELTYP | **A**

Velocity type of SYSVEL, e.g. LSR, BARYCENT, GEOCENTR, TOPOCENT

VELDEF | **A**

Definition of relativistic velocity (OPTICAL or RADIO).

PMRA | **D** | **DEGREE/DAY**

Proper motion in RA. Perhaps DEGREE/YEAR is better than DEGREE/DAY

PMRA_ ERR **D** **DEGREE/DAY**

Estimated rms uncertainty in PMRA

PMDEC **D** **DEGREE/DAY**

Proper motion in DEC. Perhaps DEGREE/YEAR is better than DEGREE/DAY

PMDEC_ ERR **D** **DEGREE/DAY**

Estimated rms uncertainty in PMDEC

PARALLAX **E** **ARCSEC**

Heliocentric Parallax

PARALLAX_ ERR **E** **ARCSEC**

RMS uncertainty in Parallax

Table: ALIASES

Describes alternative names for sources whose **standard** name is contained in the **SOURCES** table. Note that the **standard** name is only standard internal to a particular realization of the table, it need not be an external standard, e.g. an IAU position specification. An alias may be any other name for the source as long as it is unique within the **ALIASES** table. I propose the following syntactic rules for aliases:

(1) They are limited to 32 characters;

(2) they contain only printable ASCII characters; These are ASCII codes 32→126 (decimal).

(3) case of alphabetic characters is not significant, i.e. 'hd102453' is equivalent to 'HD102453';

(4) blanks are permitted but not significant, so 'hd102453' is equivalent to ' hd 102453 '. Positional specifications can be used: '0021+015'. Because of the rule on blanks, it is advisable to always use the plus sign '+' before positive declinations. **Standard** names have the same syntactic rules but are limited to 16 characters (this comes from the previous IAU definition of SOURCE tables). The ALIASES table has two columns in the form of a 1→N relation. That is, the first column has the **standard** name for a source and the second column has an alias for this source. A given **standard** name may appear as many times as necessary to list all of its aliases, which must all be unique. For consistency, the **standard** name should also be present as an alias for itself.

Header Keywords

Keyword	**Type**
REVISION	**INTEGER**

The revision number of table definiton.

ORIGIN	**STRING**

Who created this table

Table Columns

Column	**Format**	**Units**
SOURCE	**16A**	

Unique identifier internal to tables.

ALIAS	**32A**	

Aliases for SOURCE.

Table: BANDS

Describes characteristics of the wavelength bands for which photometry and spectrophotometry can be described. The intention is to make more precise what someone means if he/she says 'K-band' or 'I-band'. We allow modifiers to the band name such as 'MIDI' or 'Gunn'. Thus there may be several definitions of 'N-band' Finally we use a character code to define whether the band is, for example, instrumentally defined, or by precise filter cutoffs or unknown. For the moment I am ambivicating between the choice of wavelength or frequency as the primary coordinate, since I will never get both the optical and radio communities to agree on one system. So the keyword 'VARIABLE' defines whether wavelength or frequency is the variable being used to describe the band. The boolean 'BANDPASS_ FILE' entry indicates whether there is an optional separate table defining the bandpass precisely as a function of frequency/wavelength.

Header Keywords

Keyword	**Type**
REVISION	**INTEGER**

The revision number of table definiton.

ORIGIN	**STRING**

Who created this table

VARIABLE	**STRING**

Acceptable values are 'WAVELENGTH' or 'FREQUENCY'

Table Columns

Column	**Format**	**Units**
BAND	**8A**	

Generic identifier for band. Examples are 'K','L_ RADIO'. 'IRAS12'

MODIFIER	**8A**	

More specific description of band.

EFFECTIVE_ FREQ	**E**	**HZ**

Effective (weighted mean) frequency of band if VARIABLE='FREQUENCY', otherwise not present

EFFECTIVE_ WAVE	**E**	**METERS**

Effective (weighted mean) wavelength of band if VARIABLE='WAVELENGTH', otherwise not present

BANDWIDTH_ FREQ	**E**	**HZ**

Full width of band if VARIABLE='FREQUENCY'

BANDWIDTH_ WAVE	**E**	**METERS**

Full width of band if VARIABLE='WAVELENGTH'

BAND_ CODE	**8A**	

Codes describing definition of band. Examples are 'I'='instrumentally defined'; 'F'='defined by filters','A'=defined by atmosphere. Up to eight codes can be specified. If no code is specified the definition is 'unknown'.

SHORT_ REF **8A**

Short pointer to row in REFERENCE table with literature reference defining band. E.g. 'cotton99'.

BANDPASS_ FILE **L**

Indicates that a file containing precise bandpass information is present.

COMMENT **80A**

Any other useful information about band definition

Table: BANDPASSES

Gives measured or theoretical transmission in specified wavelength bands as a function of wavelength. Because the array of wavelength values will probably be highly varying from band to band, and I don't want to handle variable length arrays in one table, there will only be once row per table. In the FITS transport format there will have to be an additional relational table that specifies the table name for a given source and band. This extra table is not described in this document. For internal formats that easily compass variable length arrays, this problem will be disappear when the FITS files are read into internal format.

Header Keywords

Keyword **Type**

REVISION **INTEGER**

The revision number of table definiton.

ORIGIN **STRING**

Who created this table

VARIABLE **STRING**

'WAVELENGTH' or 'FREQUENCY'

Table Columns

Column **Format** **Units**

BAND **8A**

Generic identifier for band. Serves as reference to 'BAND' file.

MODIFIER **8A**

Reference to band 'MODIFIER' in 'BAND' file.

WAVELENGTH **E** **METERS**

Wavelength at each data point if VARIABLE='WAVELENGTH'.

FREQUENCY **E** **HZ**

Frequency at each data point if VARIABLE='FREQUENCY'.

TRANSMISSION **E**

band transmission at specified wavelength.

Table: DIAMETERS

Gives estimate of stellar diameters in specified wavelength bands, which must be described, for better or worse, in the 'BANDS' table. In addition to the specified diameters and errors and references, I add a 1-byte to describe the method of estimation:

(1) 'L' = lunar occultation

(2) 'P' = photometric

(3) 'I' = direct interferometric measurement and an 8-byte 'TYPE' to indicate what type of model has been fit, e.g.

(1) 'UNIFORM ' = Uniform Disk

(2) ' ' = unknown

(3) 'LIMB1 ' = limb darkened model 1 (?)

Header Keywords

Keyword	Type
REVISION	**INTEGER**

The revision number of table definiton.

ORIGIN	**STRING**

Who created this table

Table Columns

Column	Format	Units
SOURCE	**16A**	

Unique identifier of source.

BAND	**8A**	

Generic identifier for band. Serves as reference to 'BAND' file.

MODIFIER	**8A**	

Reference to band 'MODIFIER' in 'BAND' file.

DIAMETER	**E**	**ARCSEC**

Diameter measured at this band.

DIAMETER_ ERR	**E**	**ARCSEC**

RMS estimated error in Diameter.

DATE	**D**	

MJAD at which diameter was measured.

METHOD	**1A**	

Method used to estimate diameter, e.g. 'L' = lunar occultation.

TYPE **8A**

Model used to fit diameter e.g. 'UNIFORM ' = uniform disk model.

SHORT_ REF **8A**

Short pointer to row in REFERENCE table with literature reference defining source of diameter information. E.g. 'jaffe04a'.

COMMENT **80A**

Any other useful information

Table: PHOTOMETRY

Gives fluxes (in Jy) in specified wavelength bands, which must be described, for better or worse, in the 'BANDS' table. The biggest problem is the 'QUALITY' flag section. As in the 'BAND_ CODE' flags in the 'BAND' table, we need a set of agreed upon codes. As a start I suggest something like:

(1) 'P' = Primary; that is a measurement reducable directly to real physical units like Jy, or a theoretical model in physical units.

(2) 'S' = Secondary; derived by direct instrumental comparison in this band with a Primary calibrator.

(3) 'T' = Tertiary; derived by direct instrumental comparison in this band with a Secondary or other Tertiary calibrator.

(4) 'A' = 'Air mass' or 'Atmospheric'; corrected for atmospheric absorption. This is implicit in Primary calibrators, but not in the others.

(5) 'V' = 'Variable' Known or suspected variable.

(6) 'L' = 'poLarized' Known or suspected to be polarized.

Header Keywords

Keyword **Type**
REVISION **INTEGER**

The revision number of table definiton.
ORIGIN **STRING**

Who created this table
STOKES **STRING**

Describes which stokes parameters are present in the table. Most often this will be just 'I' but sometimes 'IQU' and seldom 'IQUV' or some other combination.

Table Columns

Column **Format** **Units**
SOURCE **16A**

Unique identifier of source.
BAND **8A**

Generic identifier for band. Serves as reference to 'BAND' file.
MODIFIER **8A**

Reference to band 'MODIFIER' in 'BAND' file.
FLUX **E** **JANSKY**

Stokes(I) Flux in band.
FLUX_ ERR **E** **JANSKY**

Estimated rms uncertainty in I Flux.

UFLUX	**E**	**JANSKY**

Stokes(U) Flux in band if specified in keyword 'STOKES'.

UFLUX_ ERR	**E**	**JANSKY**

Estimated rms uncertainty in U Flux.

QFLUX	**E**	**JANSKY**

Stokes(Q) Flux in band if specified in keyword 'STOKES'.

QFLUX_ ERR	**E**	**JANSKY**

Estimated rms uncertainty in Q Flux.

VFLUX	**E**	**JANSKY**

Stokes(V) Flux in band.

VFLUX_ ERR	**E**	**JANSKY**

Estimated rms uncertainty in V Flux if specified in keyword 'STOKES'.

DATE	**D**

MJAD at which flux was measured.

SHORT_ REF	**8A**

Short pointer to row in REFERENCE table with literature reference defining source of flux information. E.g. 'jaffe04a'.

QUALITY	**8A**

Up to 8 agreed upon quality codes. Blank means 'unknown'.

COMMENT	**80A**

Any other useful information

Table: VISIBILITY

Gives broadband visibility in specified wavelength bands, which must be described, for better or worse, in the 'BANDS' table. The biggest problem is the 'QUALITY' flag section. As in the 'BAND_ CODE' flags in the 'BAND' table, we need a set of agreed upon codes. As a start I suggest something like:

(1) 'P' = Primary; that is a measurement reducable directly to real physical units like Jy, or a theoretical model in physical units.

(2) 'S' = Secondary; derived by direct instrumental comparison in this band with a Primary calibrator.

(3) 'T' = Tertiary; derived by direct instrumental comparison in this band with a Secondary or other Tertiary calibrator.

Header Keywords

Keyword **Type**
REVISION **INTEGER**

The revision number of table definiton.

ORIGIN **STRING**

Who created this table

Table Columns

Column **Format** **Units**
SOURCE **16A**

Unique identifier of source.

BAND **8A**

Generic identifier for band. Serves as reference to 'BAND' file.

MODIFIER **8A**

Reference to band 'MODIFIER' in 'BAND' file.

VISIBILITY **E**

Estimated average visibility in band.

VISIBILITY_ ERROR **E**

Estimated rms uncertainty in Visibility.

DATE **D**

MJAD at which visibility was measured.

SHORT_ REF **8A**

Short pointer to row in REFERENCE table with literature reference defining source of visibility information. E.g. 'jaffe04a'.

QUALITY **8A**

Up to 8 agreed upon quality codes. Blank means 'unknown'.

COMMENT **80A**

Any other useful information

Table: SPECTROPHOTOMETRY

Gives fluxes (in Jy) in specified wavelength bands as a function of wavelength. Because the array of wavelength values will probably be highly varying from band to band and source to source, and I don't want to handle variable length arrays in one table, there will only be once source per table. In the FITS transport format there will have to be an additional relational table that specifies the table name for a given source and band. This extra table is not described in this document. For internal formats that easily compass variable length arrays, this problem will be disappear when the FITS files are read into internal format.

Header Keywords

Keyword **Type**

REVISION **INTEGER**

The revision number of table definiton.

ORIGIN **STRING**

Who created this table

Table Columns

Column **Format** **Units**

SOURCE **16A**

Unique identifier of source.

BAND **8A**

Generic identifier for band. Serves as reference to 'BAND' file.

MODIFIER **8A**

Reference to band 'MODIFIER' in 'BAND' file.

WAVELENGTH **E** **METERS**

Wavelength at each data point.

FLUX **E** **JANSKY**

Stokes(I) Flux at specified wavelength.

FLUX_ ERR **E** **JANSKY**

Estimated rms uncertainty in FLUX.

DATE **D**

MJAD at which visibility was measured.

SHORT_ REF **8A**

Short pointer to row in REFERENCE table with literature reference defining source of visibility information. E.g. 'jaffe04a'.

QUALITY **8A**

Up to 8 agreed upon quality codes. Blank means 'unknown'.

COMMENT **80A**

Any other useful information

Table: SPECTROVISIBILITY

Gives measured or computed visibility in specified wavelength bands as a function of wavelength. Because the array of wavelength values will probably be highly varying from band to band and source to source, and I don't want to handle variable length arrays in one table, there will only be once source per table. In the FITS transport format there will have to be an additional relational table that specifies the table name for a given source and band. This extra table is not described in this document. For internal formats that easily compass variable length arrays, this problem will be disappear when the FITS files are read into internal format.

Header Keywords

Keyword **Type**
REVISION **INTEGER**

The revision number of table definiton.

ORIGIN **STRING**

Who created this table

Table Columns

Column **Format** **Units**
SOURCE **16A**

Unique identifier of source.

BAND **8A**

Generic identifier for band. Serves as reference to 'BAND' file.

MODIFIER **8A**

Reference to band 'MODIFIER' in 'BAND' file.

WAVELENGTH **E** **METERS**

Wavelength at each data point.

VISIBILITY **E**

visibility at specified wavelength.

VISIBILITY_ ERR **E**

Estimated rms uncertainty in visibility.

DATE **D**

MJAD at which flux was measured.

SHORT_ REF **8A**

Short pointer to row in REFERENCE table with literature reference defining source of flux information. E.g. 'jaffe04a'.

QUALITY **8A**

Up to 8 agreed upon quality codes. Blank means 'unknown'.

COMMENT **80A**

Any other useful information

Table: REFERENCES

Lists literature references used in other tables. Uses as key an 8 byte character string. The current intention is to provide a simple human-readable reference, so I have left the reference itself as a 132-byte unformatted character string. If we want to stimulate automatic formatting for publications we might adopt Bibtex conventions and have separate columns for 'AUTHOR' 'VOLUME' 'YEAR' etc. For now this seems overly complex.

Header Keywords

Keyword **Type**

REVISION **INTEGER**

The revision number of table definiton.

ORIGIN **STRING**

Who created this table

Table Columns

Column **Format** **Units**

SHORT_ REF **8A**

Abbreviated specification, used in other tables to point to this entry.

REFERENCE **132A**

Complete reference to source of information. This may be the literature, but in many cases information internal to the repository. For example many fluxes may be derived from a specific solution made by someone working at an observatory at a specific date, but unpublished. This solution should be referenced by name, observatory and date.