



IAU Working Group FITS Format Comments

A. Boden & R. Akeson



Outline

- Endorsement (& Gratitude)
- ISC Data Format Architecture
- Status
- Comments on Existing Standards Document
 - Multiplicity
 - ❖ Multiple Arrays
 - ❖ Multiple Wavelengths
 - Calibration/Transfer/Gain Functions
 - Times
 - Metadata
 - ❖ Alternative Formats (e.g. XML per NVO/IVO)



Endorsement

- ISC Wishes to Communicate:
 - Its endorsement of and support for data standards effort
 - Its gratitude to community for advancing the issue
 - ❖ Peter Lawson
 - ❖ Tom Pauls
 - ❖ John Young (for the *excellent* template code!)
 - Its enthusiasm for participating in and contributing to the standards activity
 - ❖ *This will greatly please our ISC oversight committee...*
 - Its commitment to produce data products according to community-adopted standards
 - ❖ Though *not* sole and exclusive mechanism



Data Standards: FITS and Others

- Recommendation: *FITS as Internal Data Standard*
- FITS *not* recommended for internal use by T. Cornwell:
 - “I don't recommend FITS as the internal format for processing. It's very poorly suited for that. The recommendation should be to establish a FITS format for interchange but to use something more capable internally.” (Priv. Comm. 09 April)
- FITS not “suitable” as interchange mechanism to projects (KI, SIM -- Shao):
 - “The ISC would fail to fulfill its prime function during the debugging phase [of KI and SIM] if the FITS format were the sole data format for interchange.” (Priv. Comm. 23 July)



FITS Format

- Agreement With NASA HQ (15 April 2002):
 - “ISC will use FITS as *an* external interface format”
 - ISC will choose internal format during formulation (phase B)
- Existing (high-level) native KI data formats are all ASCII-based
- Developing a generic ASCII—FITS extraction capability
 - Built on top of NASA/HESARC CFITSIO package (& Perl module I/F)
 - Functioning KI V^2 extraction to IAU working group (draft) standard FITS format (particular binary table structure)
 - To:
 - ❖ be packaged with L-2 processing codes for user home environments
 - ❖ become available as a (the default?) format option from data access interface



ISC (KI) Data Format Architecture

- Level-0:
 - Custom binary format to represent structure of instrument telemetry data
- Level-1:
 - Native ASCII format for both data and metadata
 - ❖ *This is the format the project wants as it's primary interface*
 - ❖ *Opinion: This is the format the KI shared-risk users want as primary interface*
 - Extractions (ingestion) to RDBMS storage
 - ❖ On-line storage for archiving/browsing/extraction w/SQL-based search mechanisms
 - Extractions to FITS table format * (Not recommended by IAU)
 - Extractions to XML (eXtensible Markup Language)
 - ❖ Particularly for metadata – to be consistent with emerging I/NVO standards (e.g. Plante June 2002)



ISC Data Format Architecture (2 of 2)

- Level-2:
 - Native ASCII format for both data and metadata
 - ❖ *This is the format the project wants as it's primary interface*
 - ❖ *Opinion: This is the format the KI shared-risk users want as primary interface*
 - Extractions (ingestion) to RDBMS storage
 - ❖ On-line storage for archiving/browsing/extraction w/SQL-based search mechanisms
 - Extractions to FITS table format (IAU standard & extended)
 - Extractions to XML (eXtensible Markup Language)
 - ❖ Particularly for metadata
- Level-3: Not Presently Designed



IAU Format Status at ISC

- Working Prototype for Extraction to FITS Realization for PTI/KI L-1 Data
 - Not particularly in keeping with working group MO, but it makes NASA HQ and our oversight committee feel good
- Working Prototypes for Extractions to Working Standard (and Augmented Standard) for PTI/KI L-2 Data
- Implemented as Custom Perl Programs That Parse Our Native (ASCII) Data Formats and Express in FITS
 - Piece of cake...thanks largely to John!
- (Contemplating Utility of Inverse-Transformation Software)
- FITS-Format Products To Be Available From:
 - ISC data user interface (web tool)
 - Packaged with processing applications (i.e. available in user home environment)



“Arrays”

- Current standard calls for a single instance of an array table (OI_ARRAY), with associated keywords like:
 - Array Name (e.g. PTI, COAST)
 - Array Center Coordinates (cartesian)
- This makes it hard (or at least unnatural) to think of expressing multiple-array data sets
 - e.g. omicron Leo orbit (Hummel et al 2001) using Mark III, NPOI, and PTI data
- No direct expression of “baseline” parameters
 - Remember, a physical interferometric baseline is described by a 3-vector and a (scalar) net bias
 - Need bias representation in order to recreate delay (not currently possible – will address in six slides)



“Arrays” Continued

- Is There a Simple Way to Understand Why the Observation Start Date is a Keyword to OI_ARRAY Table? (Not a normal radio convention)
- We Suggest:
 - Moving some of the keywords to direct table expression (e.g. Name, coordinates)
 - Restructuring table so multiple instances of arrays is more straightforward, or
 - Allow for multiple instances of the OI_TABLE structure entirely
 - Capture baseline parameters directly, or
 - Capture a mid-array propagation delay for each station
 - Think about an alternative storage for date information (see below)



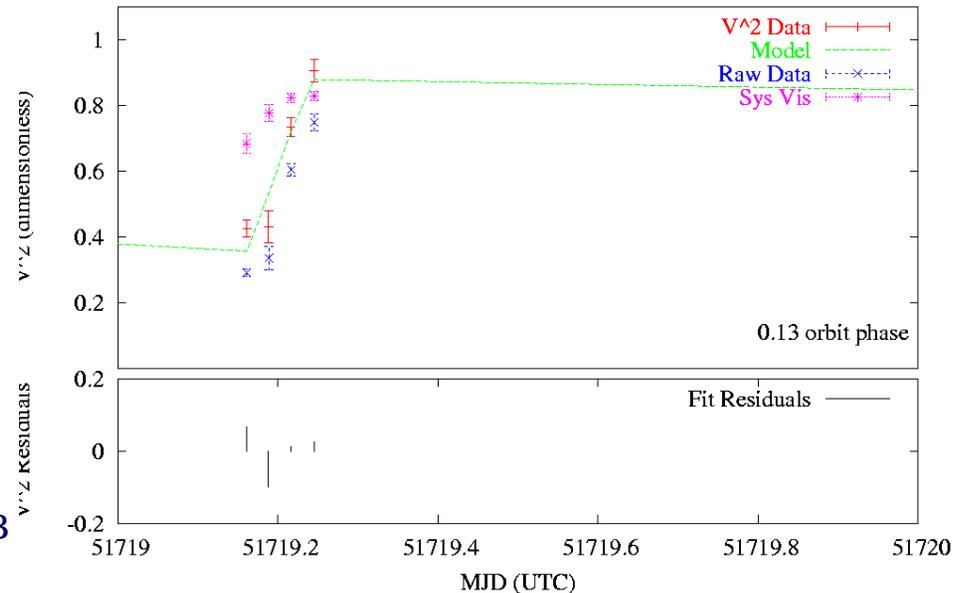
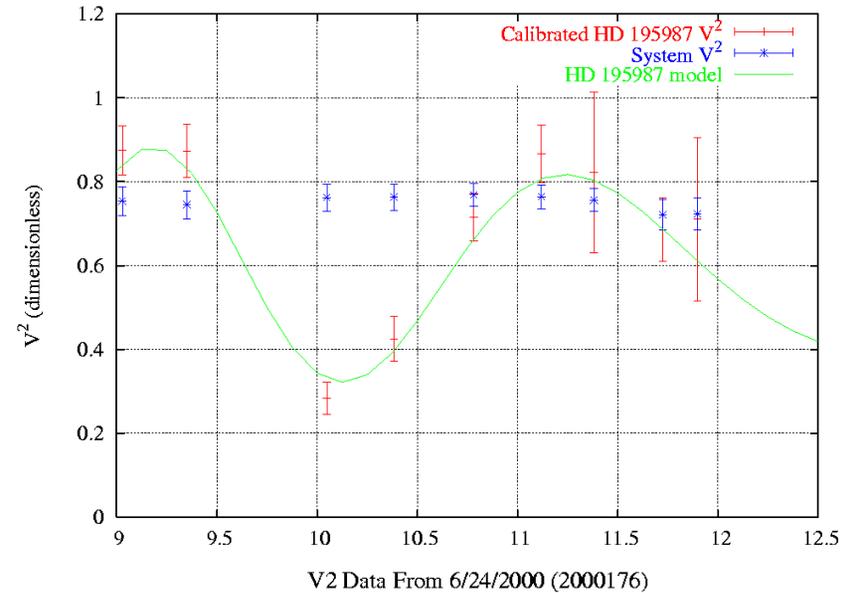
Wavelengths

- Wavelength Table Straightforward Representation of Wavelength Parameters (mid-band, pass-band)
- Is There a Relational Mechanism Between Instances in Wavelength Table and Instances in the Data Tables (e.g. OI_VIS2)?
 - Can data in table represent heterogeneous wavelength set? (Surely this would be the intent!!!)
- Suggest:
 - Create a “Wavelength ID” column in the data tables (like “Target ID”) to provide relational mechanism, or
 - Move wavelength information directly to data tables (wasteful?)



System Visibility/Gain

- No Obvious Provision for Expressing System Visibility Information
- No Obvious Provision for Expressing Calibration Source Information
- Suggest
 - Include raw and system visibility in data table, or
 - Gain tables (AIPS-like)
 - Include calibrator information in metadata, or
 - Include calibrator table to mirror target table



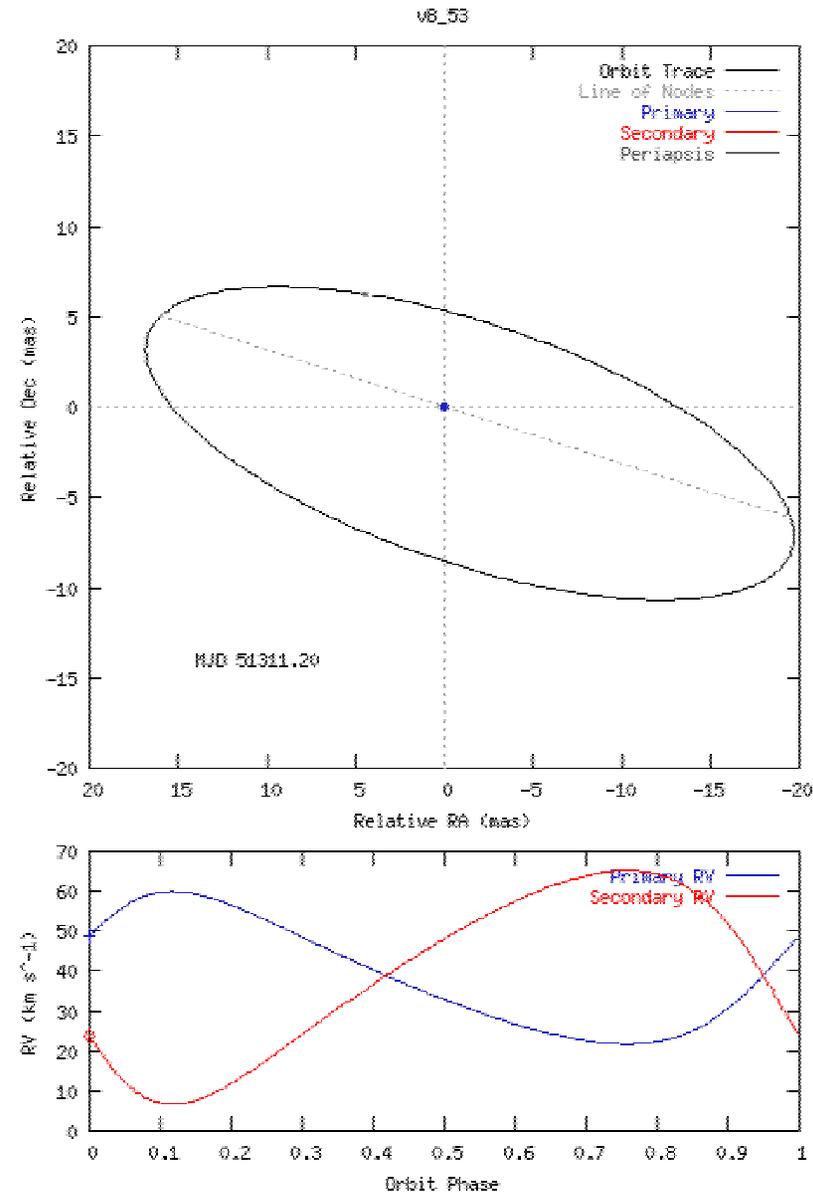
9/13/2002

AFB



Time

- Existing visibility tables use UTC in units of seconds – only
 - (Relative to start date)
 - Asymmetry between DATE-OBS and individual observation times
- For data sets of appreciable duration (multiple yrs) this is like measuring odometer distances in cm





Time (Continued)

- Suggest
 - We expand support for time expression to include other time bases
 - ❖ JD (MJD) at a minimum
 - ❖ Besselian years?
 - Include absolute (rather than purely relative) time information directly in data tables



Delays

- No Direct Realization of Delay in Present Standard
- What If Error in Target Indexing/Naming
 - Delay represents a mechanism to check target astrometry
- What If Instrument Has Delay-Related Systematic Errors?
- Suggest:
 - Delay be directly included in data tables, or
 - Ancillary delay table
 - Array infrastructure be expanded to support this inclusion



Metadata

- Metadata (Contextual Information For Data) is (Mostly) Captured in the FITS Family as Keywords
 - This has apparently worked well over years (?)
 - Will it always be so?
 - ❖ “A foolish consistency is the hobgoblin of little minds.”
- (Inter)National Virtual Observatory Standardizing on XML For Metadata Expression
 - Plante June 2002 (NVO Website)
 - VOTable (ESO Website)
 - FITS/XML Synthesis (I don’t know much about this...)
- Begs Question of Full Expression of OI Data in XML
 - Why not?
 - Should “we” be keeping more of an eye on IT developments?

Current PTI/KI L-2 V² Data Format Contents



- Data
 - Object Designation
 - Level of Averaging
 - JD (MJD)
 - UTC Date/Time
 - UTC Decimal Hours
 - Delay
 - Wavelength Parameters
 - Cal V² & Error
 - Raw V² & Error
 - System V² & Error
 - Number of Calibrators
 - u-v Coordinates
 - Hour Angle
 - Baseline Designation
- Metadata
 - Target Information
 - ❖ Astrometry
 - ❖ Spectral Typing
 - ❖ Comments (“high-proper motion star”)
 - Calibrator Information
 - Calibration Code Information
 - ❖ Designation
 - ❖ Version
 - ❖ Option/Algorithm Settings