astrolibR: Astronomy Users Library for R

Arnab Chakraborty[12], Eric D. Feigelson[234], G. Jogesh Babu[243]

1 Introduction

While astronomical research involves a wide variety of observational studies of celestial objects, certain aspects of data analysis arise very frequently. Observers must link objects with fixed locations on the celestial sphere to changing locations in the local sky at the current time. Knowledge of the location of Earth with respect to the Sun and other bodies of the Solar System is often needed. Positions and brightnesses are affected by several subtle local effects such as precession and nutation or the Earth's rotation axis, aberration and refraction by the Earth's atmosphere, and absorption by the Galaxy's interstellar medium.

Software algorithms and computer codes have long been developed to treat these problems, and many are embedded in large data analysis software packages such as AIPS (Astronomical Image Processing System) for radio astronomy, IRAF (Image Reduction and Analysis Facility), and MIDAS (Munich Image Data Analysis System) which date to the 1970-80s. The Interactive Data Language (IDL) emerged in the 1980s as a flexible environment for software development, and numerous utilities and codes were written in IDL to serve the community. The Interactive Data Language[5] is proprietary software system with a C-like grammar quite similar to the R public domain software system. While IDL specializes in image analysis and R specializes in statistical analysis of tabular data, both are general purposes languages for data analysis. It is therefore not difficult to adapt IDL codes to R.

Over 25 years, observational research astronomers have developed 'The IDL Astronomy Users Library (*astrolib*), an extensive collection of $\sim$ 500 low-level utilities and codes for data analysis implemented in the Interactive

[1]Applied Statistics Unit, Indian Statistical Institute, 203 Barrackpore Trunk Road, Kolkata 700108, India, arnabc@isical.ac.in

[2]Center for Astrostatistics, 417 Thomas Building, Pennsylvania State University, University Park PA 16802

[3]Department of Astronomy & Astrophysics, 525 Davey Laboratory, Pennsylvania State University, University Park, PA 16802

[4]Department of Statistics, 325 Thomas Building, Pennsylvania State University, University Park PA 16802

[5]http://www.exelisvis.com/ProductsServices/IDL.aspx

Data Language. This influential library is curated by Wayne Landsman at NASA's Goddard Space Flight Center[6].

2 Scope of astrolibR

This present CRAN package *astrolibR* adapts 64 astrolib utilities to the R language. They treat various calculations for: time, coordinate and proper motion transformations; terrestrial precession and nutation, atmospheric refraction and aberration, barycentric corrections, and related effects; utilities for astrometry, photometry, and spectroscopy; and utilities for planetary, stellar, Galactic, and extragalactic science. These programs are listed in Table 1, where the first column gives the IDL program name and the second column gives the IDL one-line description[7].

Each IDL utility is adapted into an R function with the same name, but in lower-case letters (e.g., *ADSTRING.pro* in IDL is *adstring.R* in R). The operation of each function is generally simple with scalar/vector/string quantities as input arguments and scalar/vector/string quantities as output values. Outputs quantities are returned directly, or within an R *list* structure, or (occasionally) within an R *data.frame* structure. Each utility can be used in isolation, although combinations are often convenient. The R help files for *astrolibR* functions are modeled closely on the internal documentation of the IDL utilities, and have examples that illustrate their use. None of the *astrolibR* functions create a special R *class* of output objects. Standard R functions such as *summary*, *str*, *plot*, and *write.table* can be applied to show and utilize *astrolibR* outputs.

3 IDL *astrolib* procedures not included in *astrolibR*

While the 64 *astrolib* procedures included in *astrolibR* constitute only a small fraction of the complete IDL *astrolib* library with ∼500 functions, they represent an important class of utilities with broad applications that are mostly absent from R and CRAN codes. Table 2 gives IDL *astrolib* procedures that are not included in *astrolibR*. Some are applicable to specific astronomical software environments such as the DAOPHOT software suite for stellar photometry, the IRAF code systems developed by the U.S. National Optical

[6]http://idlastro.gsfc.nasa.gov and https://github.com/wlandsman/IDLAstro

[7]http://idlastro.gsfc.nasa.gov/contents.html. Note that an R function POLYIDL was created to substitute for IDL's POLY.pro procedure to avoid duplication with an existing R function.

Astronomical Observatory serving ground-based telescopes, and the STSDAS software system developed by the Space Telescope Science Institute serving the orbiting Hubble Space Telescope. Others refer to the Flexible Image Transport System (FITS)[8] which defines universal standards for astronomical image and table formatting. FITS input/output is treated by other CRAN packages. Yet other IDL *astrolib* procedures implement primitive operations that do not involve astronomical data.

However, several dozen IDL *astrolib* procedures, mostly in the categories of "Math and statistics", "Plotting procedures", and "Robust statistics", have similar or identical functionalities in R or CRAN. Table 2 lists these ~ 65 IDL procedures (in capital letters) and the corresponding R function or CRAN package.

Thus, between the astronomical utilities adapted in *astrolibR* (Table 1) and the mathematical, statistical and plotting functionalities already treated in R and CRAN (Table 2), about 130 procedures in the IDL *astrolib* library are available in the R software system. Astronomers with codes dependent on IDL *astrolib* procedures who need the advanced statistical capabilities of R/CRAN can convert their codes to R with small or moderate effort.

4 Related CRAN packages and R resources for astronomy

The CRAN *Task View on Chemometrics and Computational Physics*[9] lists the rapidly growing CRAN packages associated with astronomy and astrophysics. Some of these CRAN packages − notably *astro*, *astroFns*, *celestial*, *cosmoFns*, and *moonsun* − have functions that overlap those of *astrolibR*, although sometimes with simplified calculations. The *FITSio* and *fitsR* CRAN packages provide access to astronomical data in FITS format.

Some ancillary services may be useful to the astronomer involved in data analysis using R. The *Astrostatistics and Astroinformatics Portal* (ASAIP)[10] provides a variety of resources such as recent papers, lists of jobs and meetings, links to blogs, brief articles and discussion forums related to advanced statistical analysis in astronomy. The Facebook group *Astronomy with R*[11] presents informal discussion for using R in astronomical research. The textbook *Modern Statistical Methods for Astronomy with R Applications* (Feigel-

[8] https://fits.gsfc.nasa.gov
[9] http://cran.r-project.org/web/views/ChemPhys.html
[10] https://asaip.psu.edu
[11] http://www.facebook.com/groups/astro.r

son & Babu, 2012) gives many examples of R/CRAN usage for astronomy. The R scripts and astronomical datasets for the text are available from Penn State's Center for Astrostatistics[12].

Acknowledgements: This work was supported by NSF's Scientific Software Infrastructure program, SI2-SSE grant AST-1047586 (P.I. G. J. Babu). AC performed most of the coding, and EDF wrote most of the documentation for *astrolibR*.

Bibliography

Breen, J. 2011, Accessing databases from R, http://www.slideshare.net/jeffreybreen/accessing-databases-from-r

Feigelson, E. D. and Babu, G. J. 2012, *Modern Statistical Methods for Astronomy with R Applications*, Cambridge Univ. Press

Ripley, B. D. 2001, Using databases with R, *R News*, 1(1):18-20

[12] http://astrostatistics.psu.edu/MSMA

Table 1: IDL *astrolib* procedures adapted to R in *astrolibR*

Name	Purpose
ADSTRING	Format RA and DEC as a character string
AIRTOVAC	Convert air wavelengths to vacuum wavelengths
AITOFF	Convert longitude,latitude to X,Y using Aitoff equal-area projection
ALTAZ2HADEC	Convert Horizon (Alt-Az) coordinates to Hour Angle and Declination
BARYVEL	Compute components of barycentric Earth velocity, given Julian date
BPRECESS	Precess coordinates, proper motion from J2000 to B1950
CALZ_UNRED	Deredden a galaxy spectrum using the Calzetti et al. (2000) formula
CCM_UNRED	Deredden a spectrum using the Cardelli et al. (1989) parameterization
CIRRANGE	Force an angle into the range 0 <= ang < 360
CO_ABERRATION	Calculate changes to Ra and Dec due to aberration effects
CO_NUTATE	Calculate changes in RA and Dec due to nutation of the Earth's rotation
CO_REFRACT	Calculate correction to altitude due to atmospheric refraction (with CO_REFRACT_FO
COSMO_PARAM	Derive a full set of cosmological parameters given a subset
CT2LST	Convert from civil time to local sidereal time
DAYCNV	Convert from Julian Date to calendar date
DEREDD	Deredden Stromgren indices (called by UVBYBETA)
ECI2GEO	Convert Earth-centered inertial coordinates to geographic coords
EQ2HOR	Convert celestial (ra-dec) coords to local horizon coords (alt-az)
EQPOLE	Convert longitude,latitude to X,Y using polar equal-area projection
EULER	Astronomical coordinate system conversions
FLUX2MAG	Convert from flux units to magnitudes
FM_UNRED	Deredden a spectrum using the Fitzpatrick & Massa (1998) parameterization
GAL_UVW	Calculate the Galactic space velocity (U,V,W) of a star
GALAGE	Derive a galaxy age as a function of redshift for a cosmological model (with DTDZ)
GCIRC	Compute rigorous great circle distance
GEO2ECI	Convert geographic coordinates to Earth-centered inertial coords
GEO2GEODETIC	Convert from geographic to geodetic coordinates
GEODETIC2GEO	Convert from geodetic to geographic coordinates
GLACTC_PM	Convert between celestial and Galactic (or Supergalactic) proper motion
GLACTC	Convert between Galactic and equatorial coordinates at any equinox
HADEC2ALTAZ	Converts Hour Angle and Declination to Horizon (alt-az) coordinates
HELIO_JD	Convert geocentric (reduced) Julian date to heliocentric Julian date
HELIO_RV	Compute radial velocity given binary star orbit parameters
HELIO	Compute (low-precision) heliocentric coordinates of the planets
HOR2EQ	Convert local horizon coords (alt-az) to equatorial (ra-dec)
IMF	Return values for a multi-component power law initial mass function
ISMEUV	Compute EUV optical depth due to photoionization of HI, HeI and HeII

JDCNV	Convert from calendar date to Julian date (with INTDIV)
JPRECESS	Precess positions & proper motions from B1950 to J2000
JULDATE	Convert from calendar date to reduced Julian date
LSF_ROTATE	Create a 1-d convolution kernel to broaden a spectrum from a rotating star
LUMDIST	Return luminosity distance for a given redshift & cosmological model
MAG2FLUX	Convert from magnitudes to flux units
MONTH_CNV	Convert a month name to the equivalent number or vice-versa
MOONPOS	Compute the RA and Dec (and distance) of the Moon at a given date
MPHASE	Compute illuminated fraction of the Moon's disk for given Julian dates
NUTATE	Compute the nutation in longitude and latitude for given Julian date(s)
PLANCK	Returns a blackbody flux for a given effective temperature
PLANET_COORDS	Return low-precision RA and Dec of planets give a date(s)
POSANG	Compute the position angle between sources of specified RA and Dec
PRECESS	Precess RA and Dec to a new equinox
PREMAT	Returns precession matrix from equinox 1 to equinox 2
RADEC	Format RA, Dec as Hours, Min, Sec, Deg, Min, Sec
RHOTHETA	Compute separation and position angle of a binary star
SIXTY	Convert decimal number to sexigesimal
SPHDIST	Return angular distance between two points on a sphere
SUNPOS	Compute the RA and Dec of the Sun at a given date
TEN	Convert sexigesimal number to decimal
UBVYBETA	Use Stromgren indices to derive dereddened colors, metallicity, and T_{eff}
VACTOAIR	Convert vacuum wavelengths to air wavelengths
XYZ	Compute heliocentric rectangular coordinates at given Julian date (with PRECESS_XY
YMD2DN	Convert year,month,day to day number of the year
YDN2MD	Convert day number of the year to year, month,day
ZANG	Compute angular size as a function of redshift in a Friedman cosmology

Table 2: Other *astrolib* procedures with counterparts in R

Category	**IDL**	**R/CRAN**
Utilities	AITOFF_GRID	CRAN *mapproj*
	EQPOLE_GRID	CRAN *mapproj*
	IMCONTOUR	R *contour*, lattice *contourplot*
	TIC*[13]	R *par, axTicks*
Photometry	(23 procedures)	not available
Databases	(28 procedures)	CRAN *DBI*, *RSQLite*[9]
IRAF I/O	(5 procedures)	not available
FITS I/O	(122 procedures)	note 10
STSDAS I/O	(12 procedures)	not available
Image manipulation	(19 procedures)	CRAN *adimpro*, Bioconductor *EBImage*
Math/stat	AVG	R *mean*
	AVSINH	R *asinh*
	CIC	R *loess*, CRAN *gstat, gsoR, RandomFields*11
	CSPLINE	CRAN *splines*[11]
	FACTOR	CRAN *gmp*
	FITEXY	not available (but see CRAN *simex*)
	FLEGENDRE	CRAN *gaussquad, orthopolynom*
	GAUSSIAN	R *rnorm*
	HERMITE	R *splinefun*
	KS*[12]	R *ks.test*
	KUIPER*[12]	CRAN *CircStats*
	(M)LINMIX_ERR	not available
	LINTERP	R *approx*[11]
	MEANCLIP	note 13
	MINF*[12]	R *optim, constrOptim*
	MRANDOM	CRAN *mnormt, mvtnorm*
	MULTINOM	R *rmultinom*
	NGP	CRAN *RSAGA*
	PCA	R *princomp*
	PENT	not available
	PERMUTE	R *sample*
	POIDEV	R *rpois*
	POLINT	R *loess*[11]
	POLYLEG	CRAN *gaussquad, orthopolynom*
	POLY_SMOOTH	CRAN *RTisean, signal*[11]

	PROB_KS	CRAN *ks, kolmin*
	PROB_KUIPER	CRAN *CircStats*
	QSIMP	R *integrate*, CRAN *Bolstad*
	QTRAP	R *integrate*
	QUADTERP	R *loess*, CRAN *gstat, gsoR, RandomFields*[11]
	RANDOMCHI	R *rchisq*
	RANDOMDIR	CRAN *hyperdirichlet, MCMCpack*
	RANDOMGAM	R *rgamma*
	RANDOMP	CRAN *VGAM*
	RANDOMWISH	R *rWishart*
	SIXLIN	not available
	TABINV	R *which*
	TRANSFORM_COEFF	not available
	TRAPZD	not available
	TSC	R *loess*, CRAN *gstat, gsoR, RandomFields*[11]
	TSUM	R *integrate*
	ZBRENT	R *uniroot, optimize*
Plotting	AL_LEGEND	R *legend*, CRAN*ggplot2*
	MULTIPLOT	R *plot*, CRAN *ggplot2*
	PLOTERROR	CRAN *gplots, ggplot2, Hmisc, psych*
	PLOTHIST	R *hist*, CRAN *ggplot2, gplots*
	PARTVELVEC	R *arrows*, CRAN *fields*
	SUNSYMBOL	note 14
	VSYM	R *points*
IDL structure	(7 procedures)	not available
Robust stat	AUTOHIST	R *hist*, CRAN *ggplot2*
	BIWEIGHT_MEAN	CRAN *biwt*
	HISTOGAUSS	R *hist*, MASS*fitdistr*
	MEDSMOOTH	R *smooth*[11]
	RESISTANT_MEAN	R *mean*, CRAN *robustbase, robust*
	ROBUST_LINEFIT	R *line, rlm*, CRAN *robustbase, robust*
	ROBUST_POLYFIT	R *rlm*, CRAN *robustbase, robust*
	ROBUST_SIGMA	R *mad*
Web sockets	(5 procedures)	R *make.socket*, CRAN *Rserve, svSocket, websockets*
TV display	(12 procedures)	not available
Miscellaneous	(66 procedures)	note 15

[9] See Ripley (2001) and Breen (2011)

[10] Basic input and output functionalities for FITS (Flexible Image Transport System) files are provided by the CRAN package *fitsR* which is based on the *CFITSIO* codes endorsed by the International Astronomical Union. It places FITS headers into R scalar and vector variables, binary tables into R data frames, images into R arrays. Many of the other functionalities of these IDL procedures can be reproduced using standard R manipulation functions for data frames and arrays.

[11] R has many low-dimensional smoothers and interpolators, a few of which are indicated here. See Chpts. 6, 11 and 12 in Feigelson & Babu (2012).
[12] An asterisk represents a wildcard for several closely related IDL procedures.
[13] A brief R script for 'sigma clipping' is given by Alastair Sanderson at http://www.sr.bham.ac.uk/∼ajrs/R/r-getting_started.html.
[14] Several options for producing the solar symbol (circle with central dot) in R graphics so they appear correctly in PDF and EPS output formats are described at the *Astronomy with R* Facebook group (https://www.facebook.com/groups/astro.r).
[15] These are mostly generic low level manipulations of files, string manipulations, and interactions with the host computer that are not considered here, as they are not specifically astronomical in nature. Most of these capabilities are directly available, or readily coded, in R.