Time Series data model
in the Virtual Observatory

F.Bonnarel (CDS)

on behalf of Mark Cresitello, Mireille Louys, Laurent Michel and Ada Nebot

TimeSeries VO interoperability

- Gaia
- LSST
- Catalogues in VizieR
- GAPS ? → yes , if TS VO effort successful, it should also tackle GAPS (attempts Marco M./François B.)
- Tasks : ---->
  - Discovery
  - Access
  - Representation and serialization (datamodel and mapping)

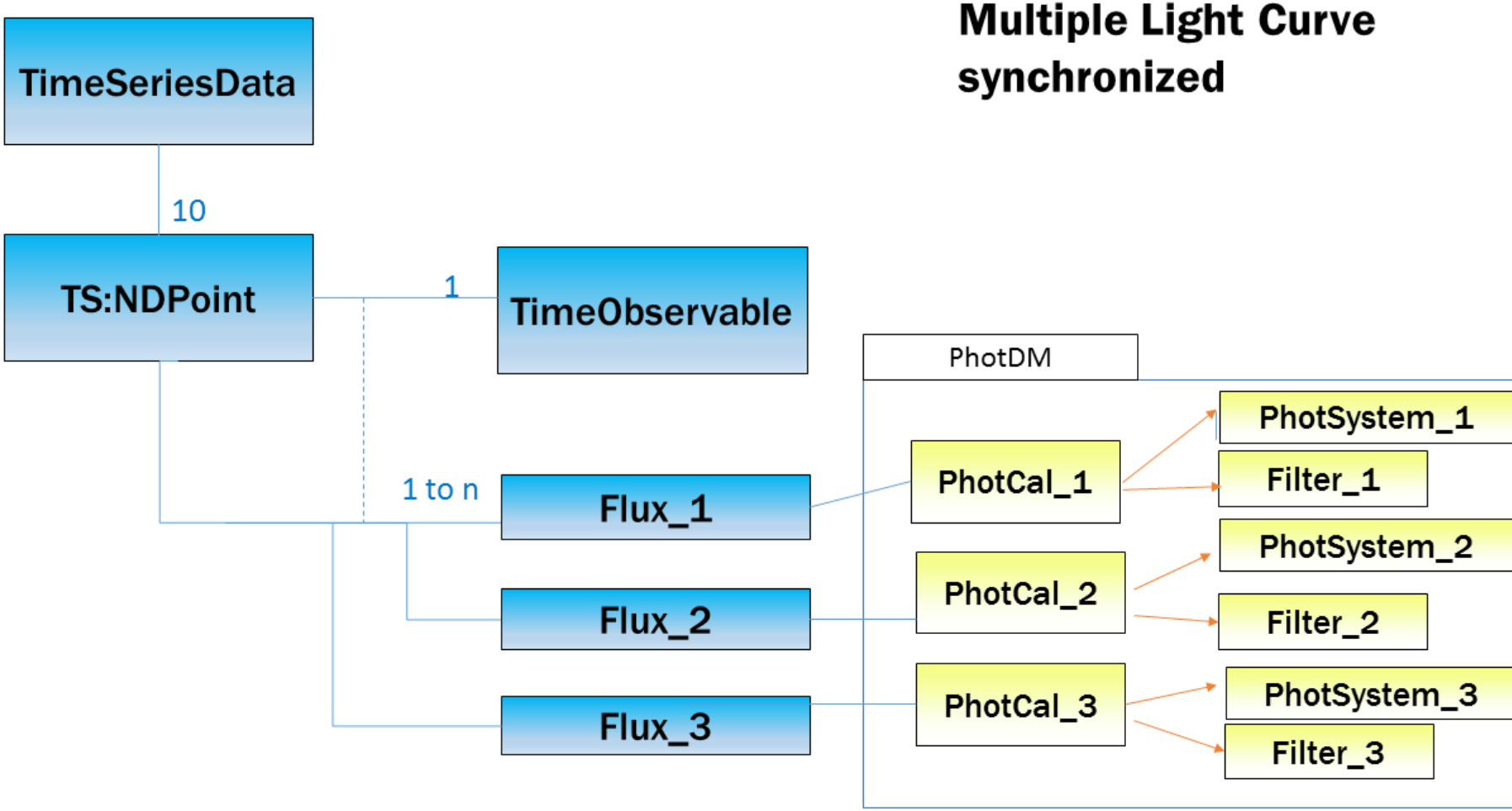
Data Representation: data model

- DataModel has to represent structures and relationship for all data and metadata
- Extension of Cube DataModel with specialization of TimeAxis
- Has to tackle scalar observables (mag, flux, radial velocity, etc..) but also variable data products

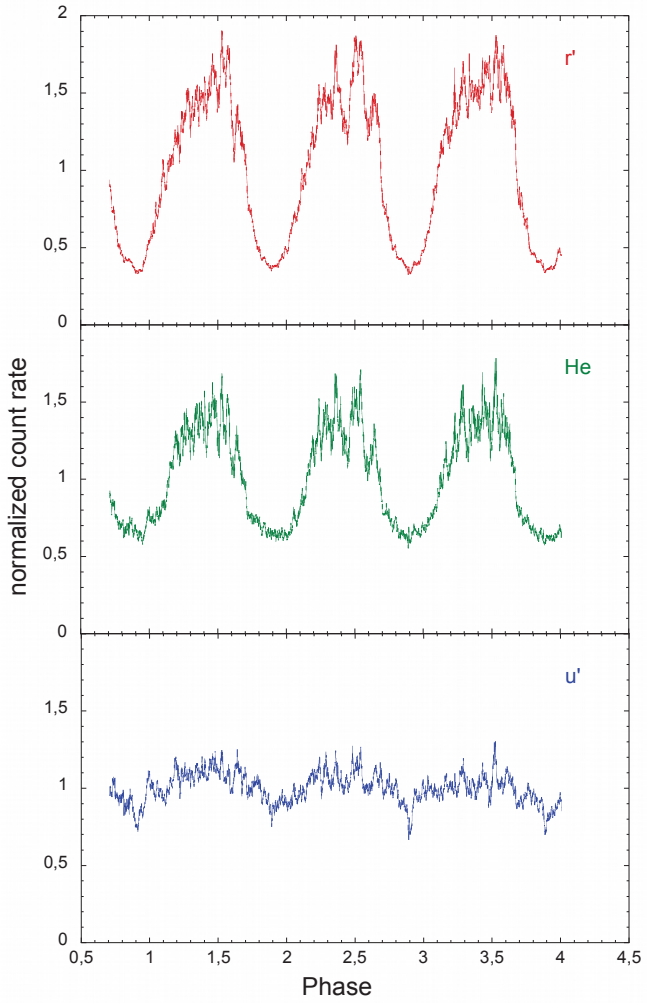
DataRepresentation : serialisation

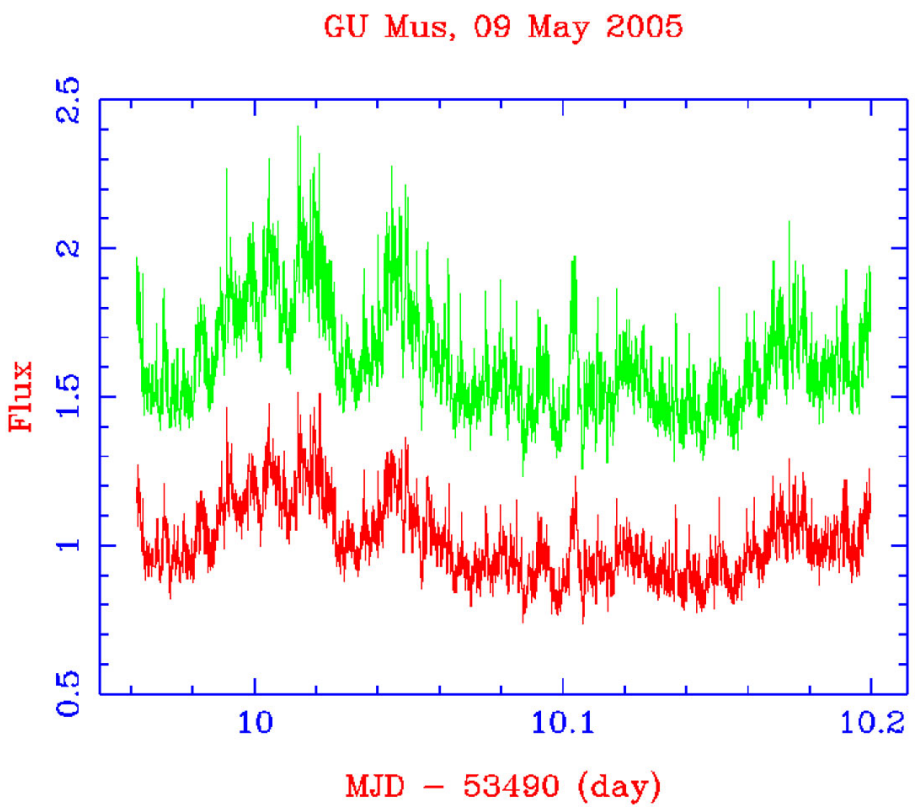
- The Model has a formal xml representation (vo-dml-xml) :
  - Direct translation of the UML diagram with special IVOA rules
  - Useful for exchanging and importing models but not usable in serialisations
  - → Need for a mapping
- Mapping of « VO-DML » structure into VOTable :
  - Full model structure on top of the table with pointers to columns (so called « vo-dml mapping »)
  - Pointers from columns to the model (so called « utypes »)
- Need test implementation and consuming for decision (see VizieR example). Feedback from data providers.

Various combinations use cases

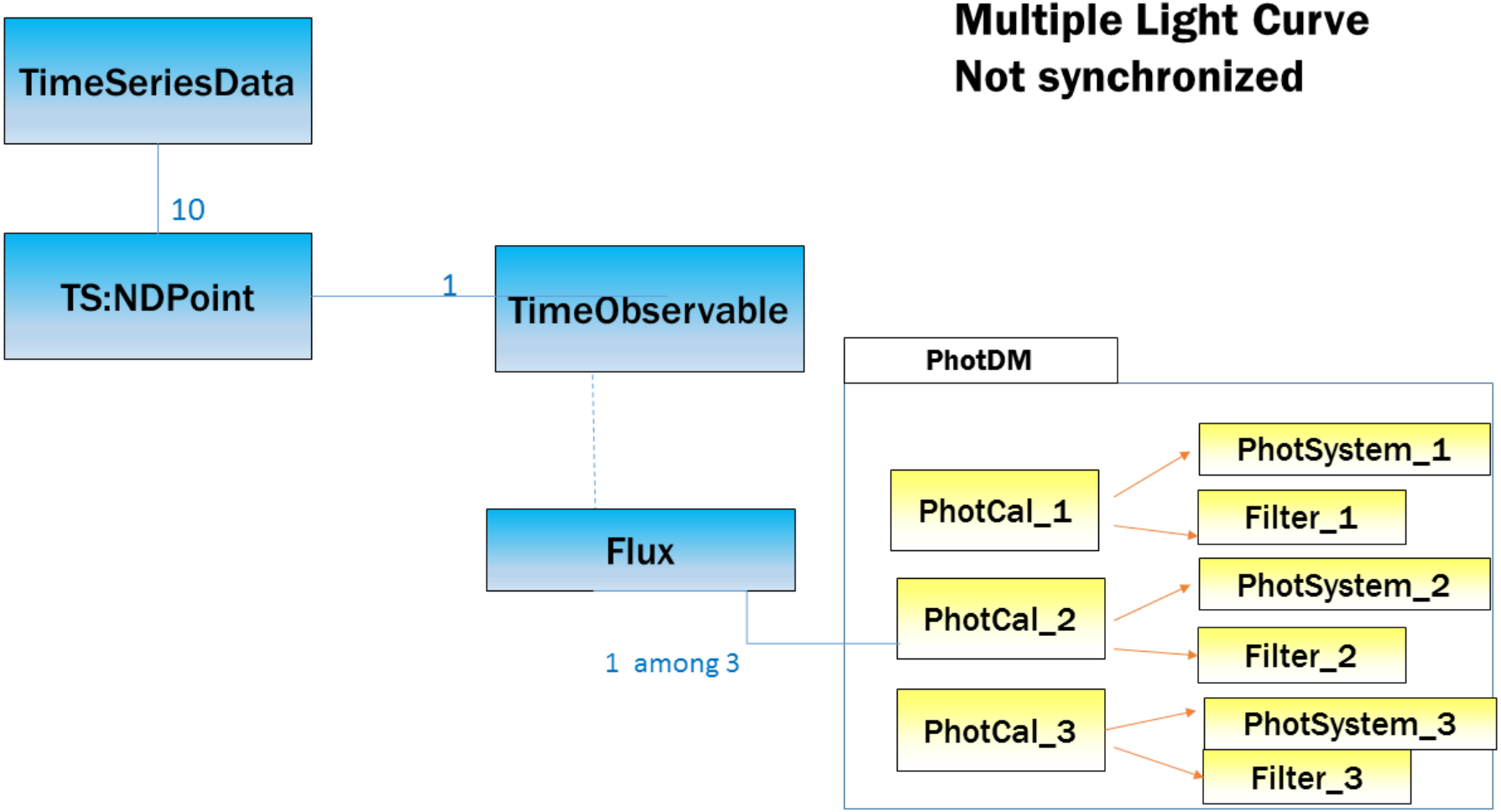


Ultracam time series





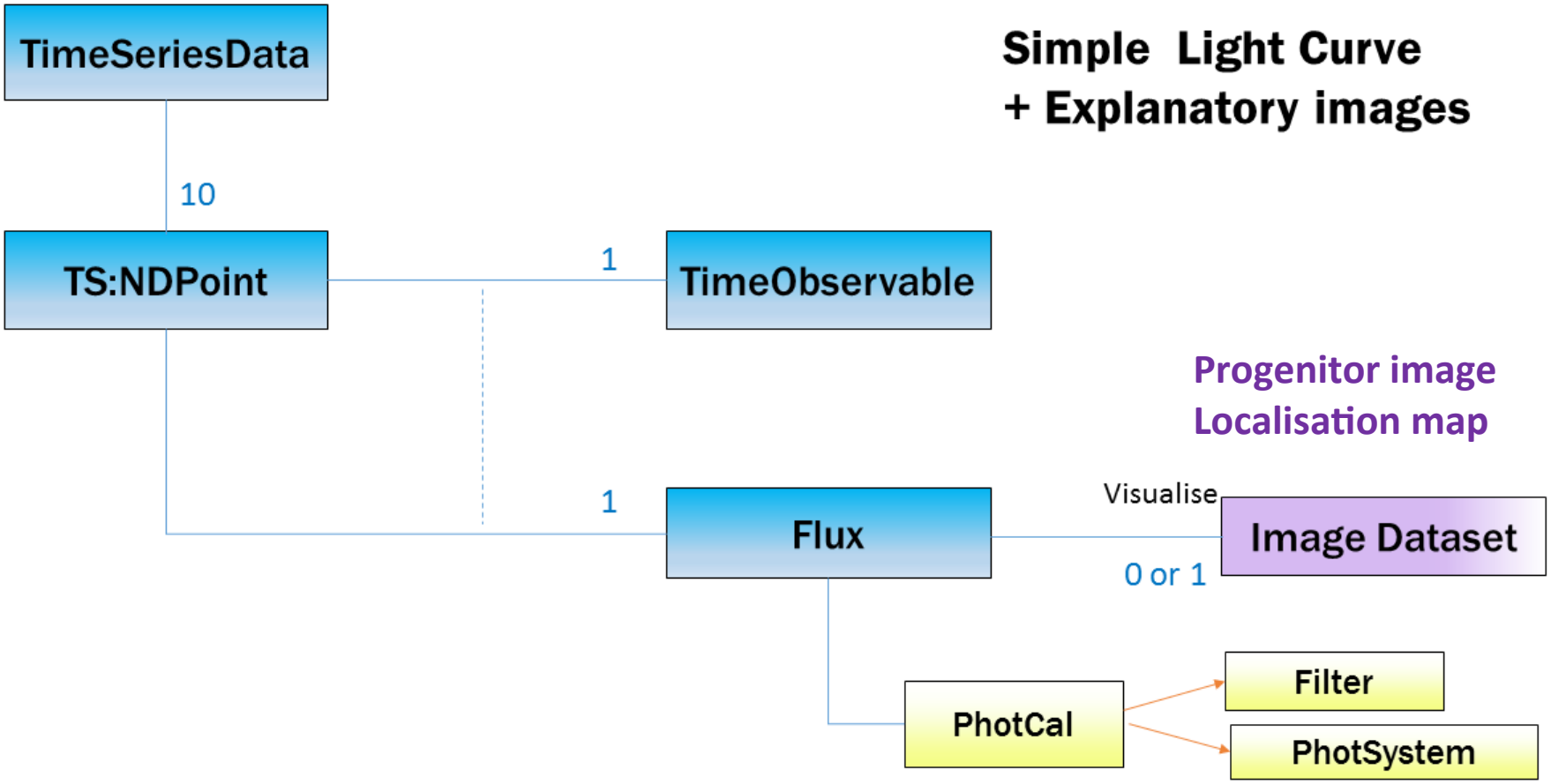
Various combinations use cases



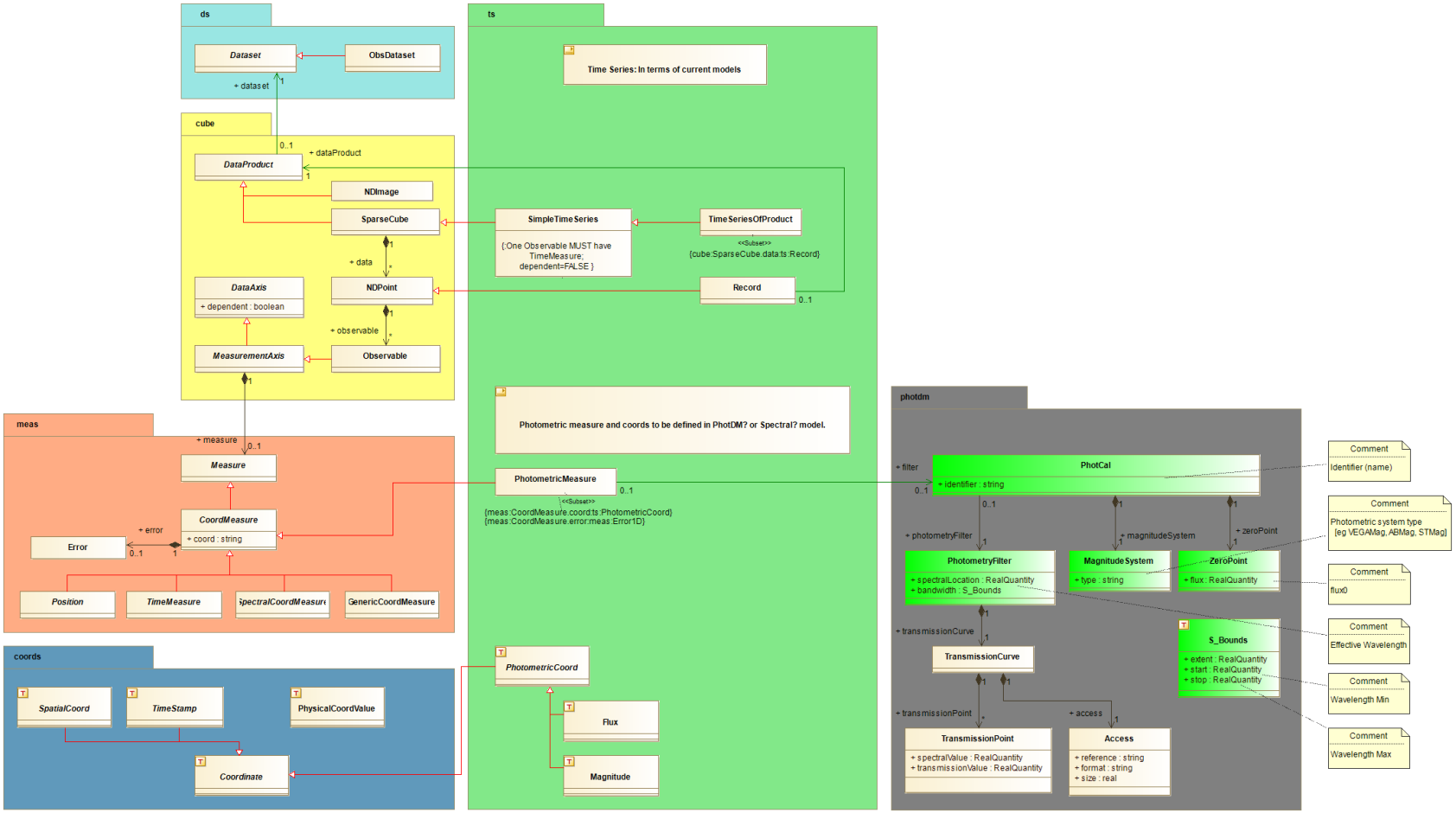
Multiband Flux measures

Coord/ Measure	T1	T2	T3	T4	T5	T6	T7	T8	Time range	Min time period	T-xel
magB	+		+	+							3
Err_magB	+		+	+							3
magV		*		*		*		*	T8-T2	Min (tj+1-tj)	4
Err_magV		*		*		*		*			4
magU				^	^	^	^	^	T7 –T4		5
Err_magU				^	^	^	^	^			5

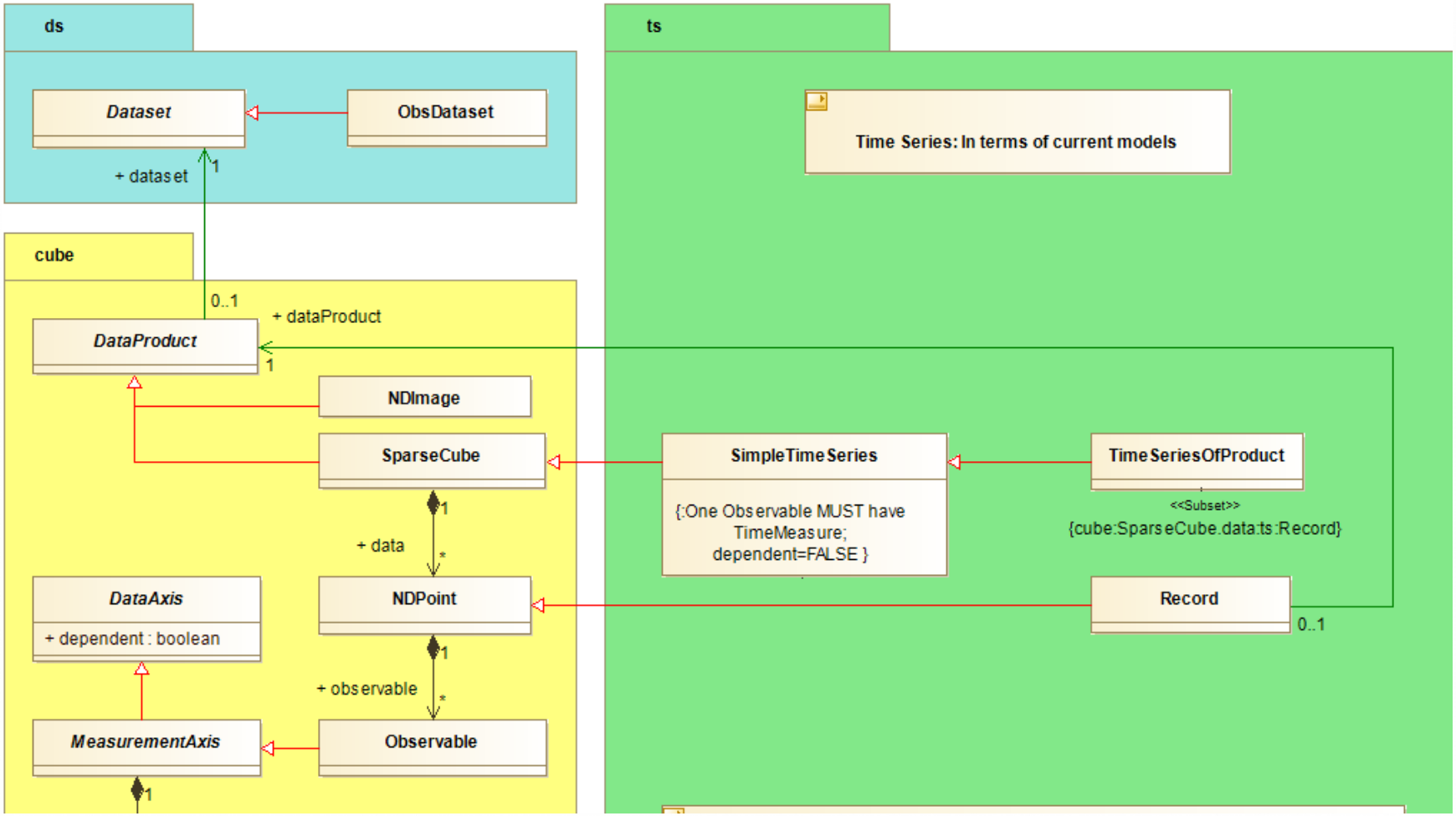
Measures + datasets



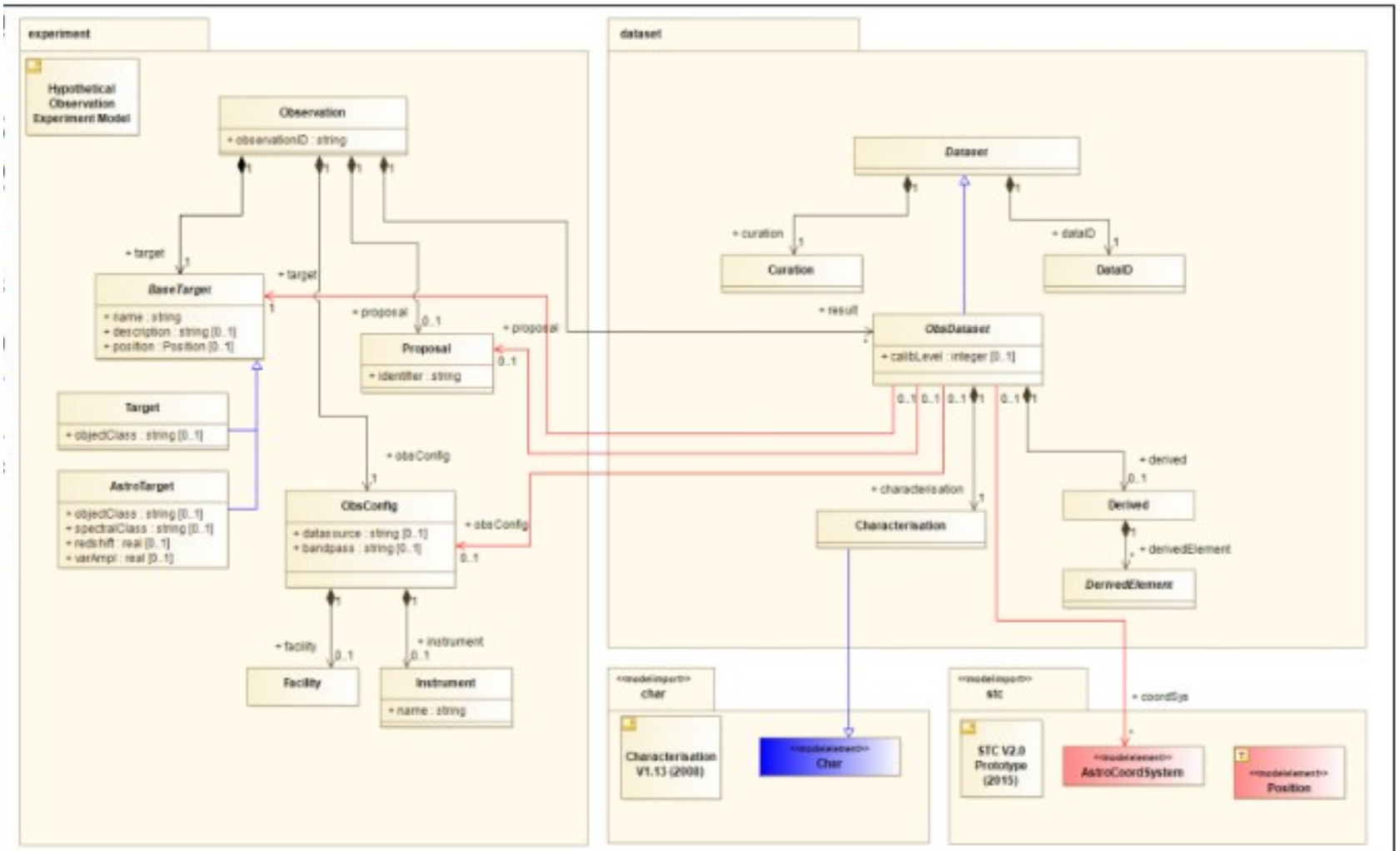
TimeSeries datamodel: full UML diagram



TimeSeries datamodel: DataSet and data structures



Focus on DataSet datamodel



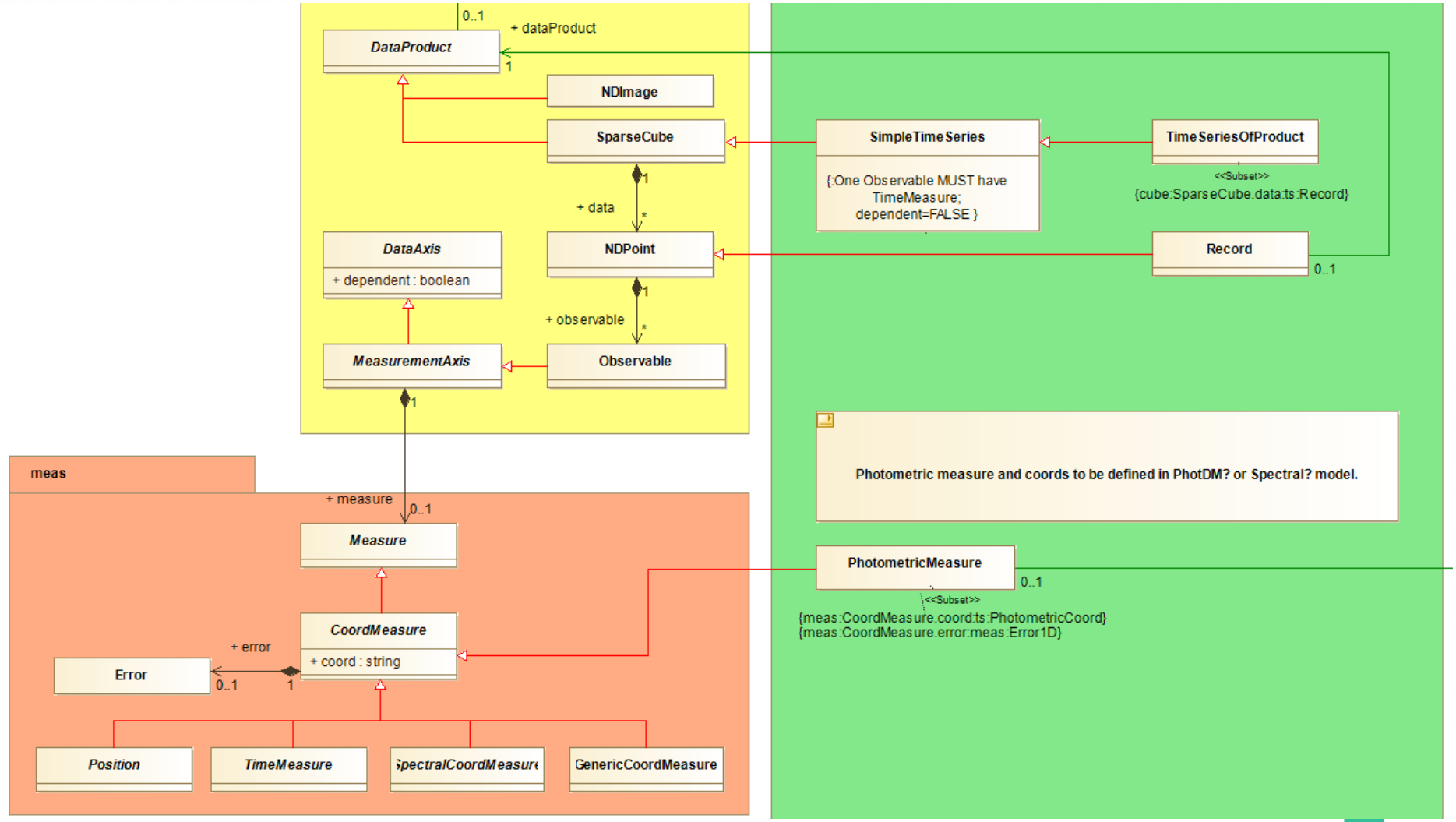
List of useful dataset and characterization : 1)

Metadata for Timeseries / Extension for ObscoreTable for Time series						
Obscore and T	**Definition TD**	**Utype**	**ucd**	**rec. units**	**Mandatory**	**default**
extension keywords		datamodelpath			/optional	
% position on sky						
s_ra	Position (within a certain area)	Char.SpatialAxis.Coverage.RefVal.	pos.eq.ra	deg	man	ICRS
s_dec	Position (within a certain area.	Char.SpatialAxis.Coverage.RefVal	pos.eq.dec	deg	man	
s_resolution	Angular resolution interval	Char.SpatialAxis.Resolution.RefVal	pos.AngResol	arcsec	man	
%target						
target_name	Name of Target	Target.name	meta.id;src	null	opt	
%Observable						
% nb of observables per point						
o_xel	Nb of observables per time point	TSNDpoint.nbMeas	meta.number	null	man	1
%observable types						
%Type of data : one value among (Events, photometry, radial velocities, spectra, images, polarisation, other)						
o_type	List of types of the Observable quantities	Char.ObservableAxis.observableTypeList ??	meta.class	null	opt	scalar
%One value in [scalar, image , spectrum, cube,...] as dataproduct_type in the Obscore vocabulary.						
% Physical nature of observable						
o_ucd	Physical nature attached to observable	Char.ObservableAxis.ucd	meta.class	null	man	
%Limits along observable axis						
% ex: Magnitudes / Fluxes/ counts, etc interval (min)						
o_min	Minimum value for Observable	Char.ObservableAxis.Coverage.BoundsLimits.LoLim	$(o_ucd); stat.min	o_units'	opt	
o_max	Maximum value for Observable (ex. Mag max)	Char.ObservableAxis.Coverage.BoundsLimits.hiLim	$(o_ucd);stat.max	o_units'	opt	
o_unit	Unit of the dependent observable	Char.ObservableAxis.unit	meta.unit	null	opt	
o_complextype	specifies if complex data are compiled value or observed with the first instruments in [false,true]	Char.ObservableAxis.status?	null	opt		
% sensitivity , max detection limit. TBC						
%o_upperlimit	upperlimit is a limiting value for the estimated faintest object in the observation (LSST, ZTF)					
o_upperlimit	flag in the data indicating that some values are upperlimits and not detections measurements. not queryable	Char.ObservableAxis.Coverage.Sensitivity.Quality???	meta.code.qual	null	opt	no

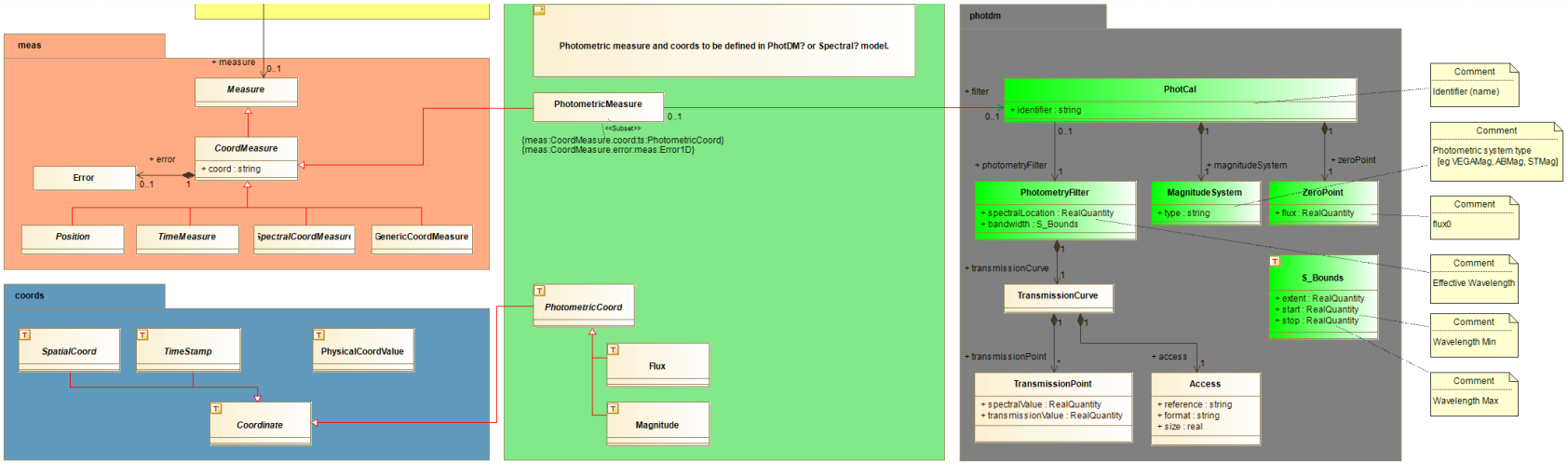
List of useful dataset and characterization : 2)

% spectral coverage						
em_min	spectral interval (min)	Char.SpectralAxis.Coverage.Bounds.Limits.LoLim	em.interval;stat.min	'em_unit'	man	nm
em_max	spectral interval	Char.SpectralAxis.Coverage.Bounds.Limits.HiLim	em.interval;stat.max	'em_unit'	man	nm
% Must be qualified by a ucd em.freq if spectral axis is in Frequency						
em_ucd	Wavelength/ Frequency/ Energy	Char.SpectralAxis.ucd	meta.ucd	null	opt	
em_unit	Unit along the spectral axis	Char.SpectralAxis.unit	meta.unit	null	opt	
% Polarisation states						
pol_states	Polarization state list	Char.Polarization.List	meta.class	null	opt	
%time features						
t_min	Time start of the sequence(min)	Char.TimeAxis.Coverage.Bounds.Limits.LoLim	time.start;obs.sequence	s	man	
t_max	Time end of the sequence	Char.TimeAxis.Coverage.Bounds.Limits.HiLim	time.end;obs.sequence	s	man	
% NB: the time span, or elapsed time for the sequence is then t_max -t_min						
t_exposure	Exposure time (sum of multiple exposures)	Char.TimeAxis.Support.Extent	time.duration;obs.exposure	s	man	
t_exp_min	Exposure time of samples (min)	Char.TimeAxis.Sampling.Extent.loLim	time.duration;obs.exposure;stat.min	s	man	
t_exp_max	Exposure time of samples (max)	Char.TimeAxis.Sampling.Extent.hiLim	time.duration;obs.exposure;stat.max	s	man	
%time space between 2 time samples / cadence						
t_sampling_step_min	minimal length of time interval between 2 observations / cadence (min)	Char.TimeAxis.Sampling.Period.loLim	time.interval;obs.sequence;stat.min	s	opt	
t_sampling_step_max	maximal length of time interval between 2 observations / cadence (min)	Char.TimeAxis.Sampling.Period.hiLim	time.interval;obs.sequence;stat.max	s	opt	
%NB : the UCD time.period is rather dedicated to a physical event. Not appropriate here						
%nb of sample along the time axis						
t_xel	nb of time stamps in the series	Char.TimeAxis.numBins	meta.number	null	man	
%Time CoosSystem						*imposed for discovery*
t_origin	Time(frame origin)	stc:TimeFrame.timeOrigin	time.epoch	?	opt	
t_scale	Time frame scale	stc:TimeFrame.timeScale	time.scale	?	opt	
t_refPosition (barycenter, heliocenter, ...)	Time reference position	stc:TimeFrame.refPosition	?	?	opt	
t_refDirection (for solar observations)	Time reference direction	stc:TimeFrame.refDirection	?	?	opt	
%Time representation ISOtime , MJD, JD , Time offset a la STC ?						
t_format	Time representation	?	?	null	man	MJD ?

TimeSeries Points as ND points made of Observable and/or DataProducts



CoordMeasure made of
Photometric Measure (PhotCal model)
or space/time/spec/pol measurements



Data Model Reuse from the IVOA

- A Time series is
  - a dataset → reuse ObsDataset from **DatasetMetadata DM**
  - A *multi axis dataset* → reuse *SparseCube* **Cube DM**
  - A collection of points of multiple dimensions Cube NDPoint
  - The principal Cube *DataAxis* is TimeAxis

Its properties can be summarized with **Characterization DM (as in ObsCore and DatasetMetadata DM)**

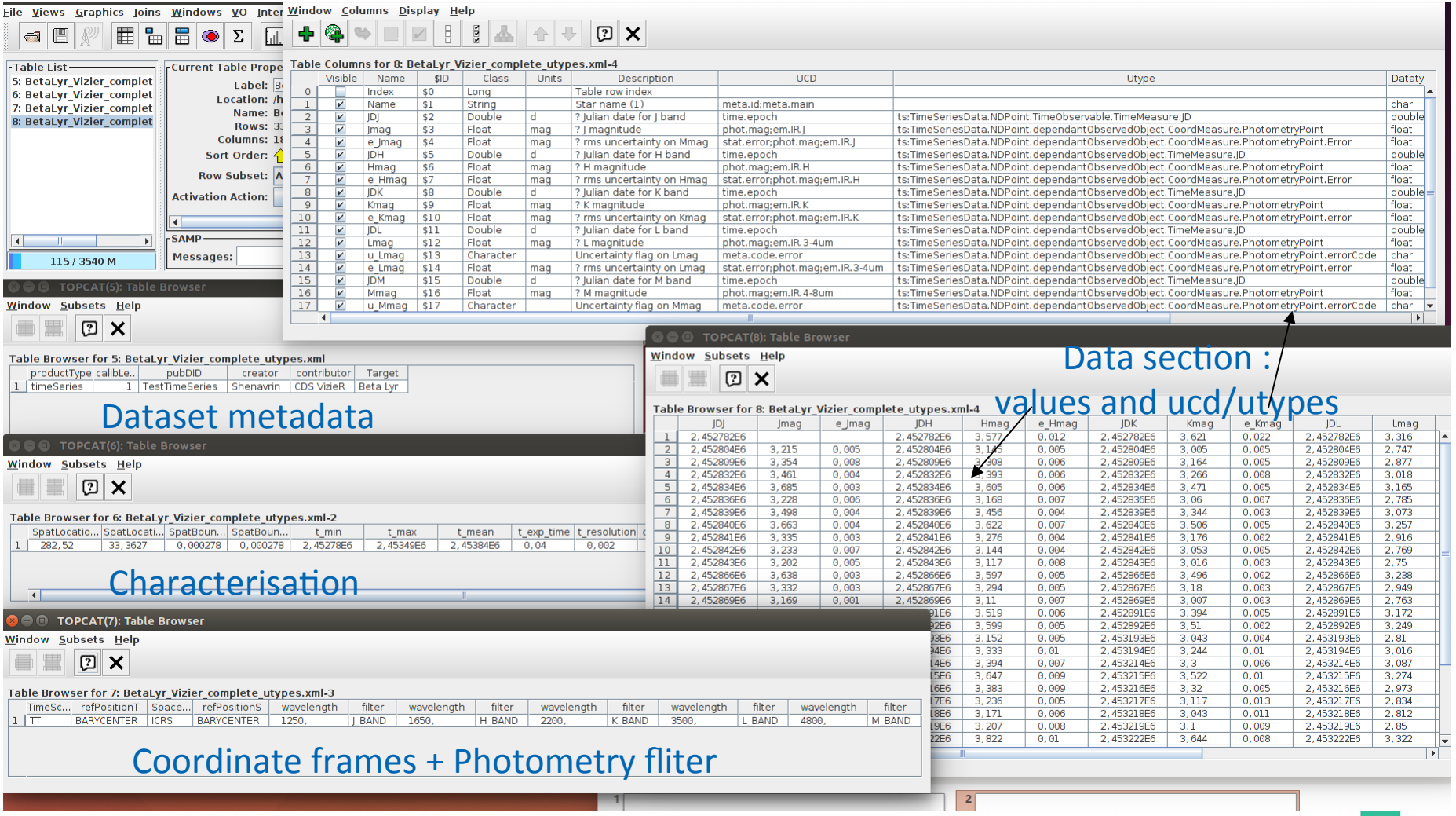
  - Measures/Observations depend on time samples
  - Simple measurement → reuse CoordMeasure as in **STCv2.0 DM**
  - Structured measures as data products → DataProducts element from CubeDM.

Beta Lyrae case

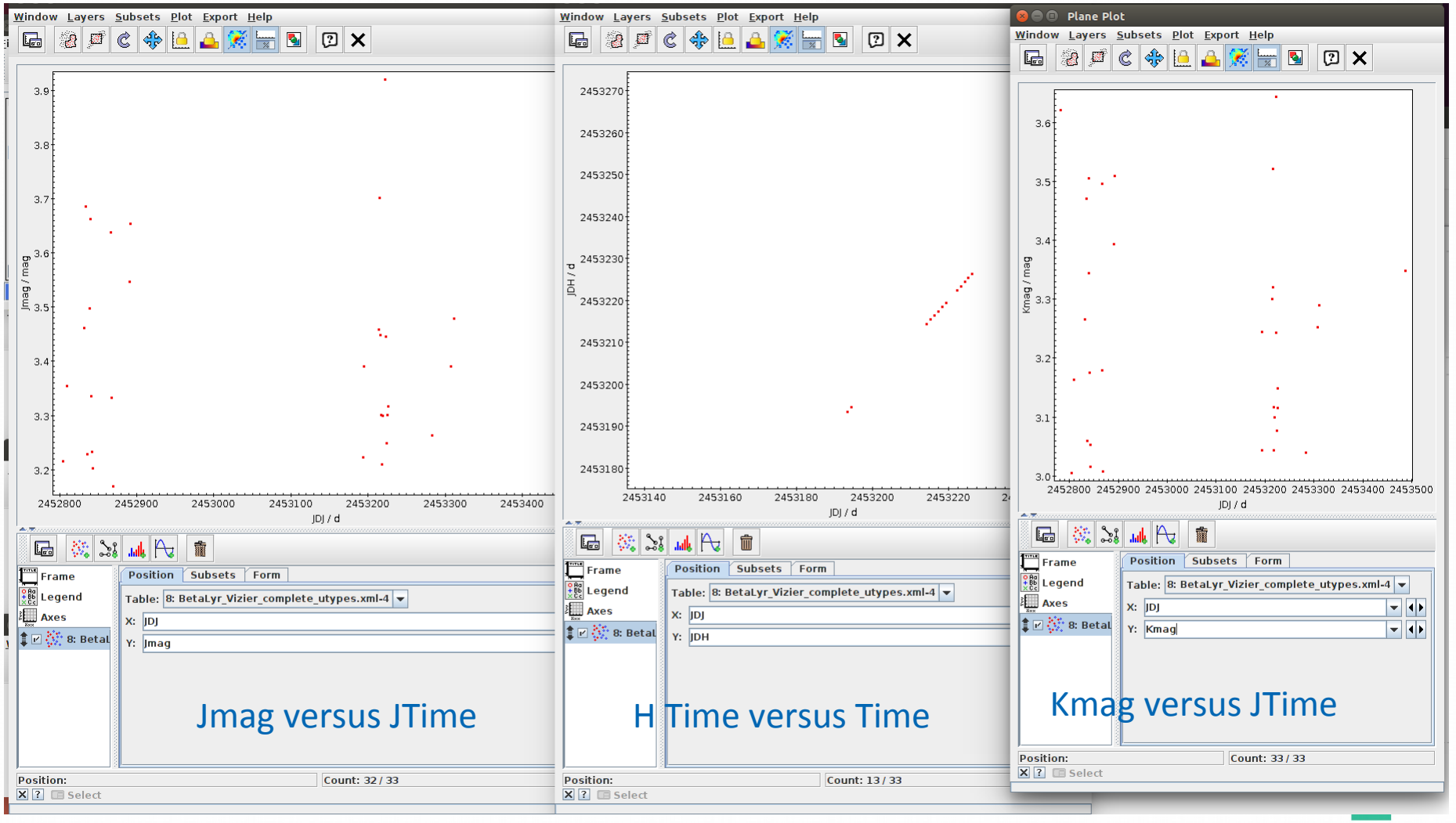
- Characterization and TimeFrame Metadata [Time Frame ,

  → TimeScale, ReferencePosition, TimeRepresentation, (TimeOrigin), Photometric filters]
- Several colors : relationShips to « Frames » (or Photometric filters) managed by reference.

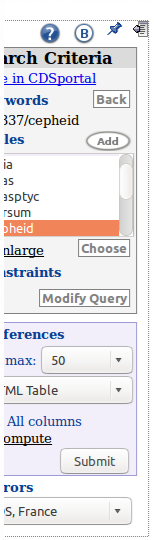
Beta Lyrae case (in TOPCAT tool)



Beta Lyrae case (in TOPCAT tool)



VizieR prototye



VizieR

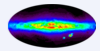
Show the target form
Show constraint information

The 4 columns in ***color*** are computed by VizieR, and are ***not part of the original data***.

I/337/cepheid *Post-annotation*

Gaia DR1 (Gaia Collaboration, 2016)
Cepheid stars identified in table VariableSummary as classification="CEP" (original column names in green) *(599 rows)* [METAtab] [METAcola] [stats]

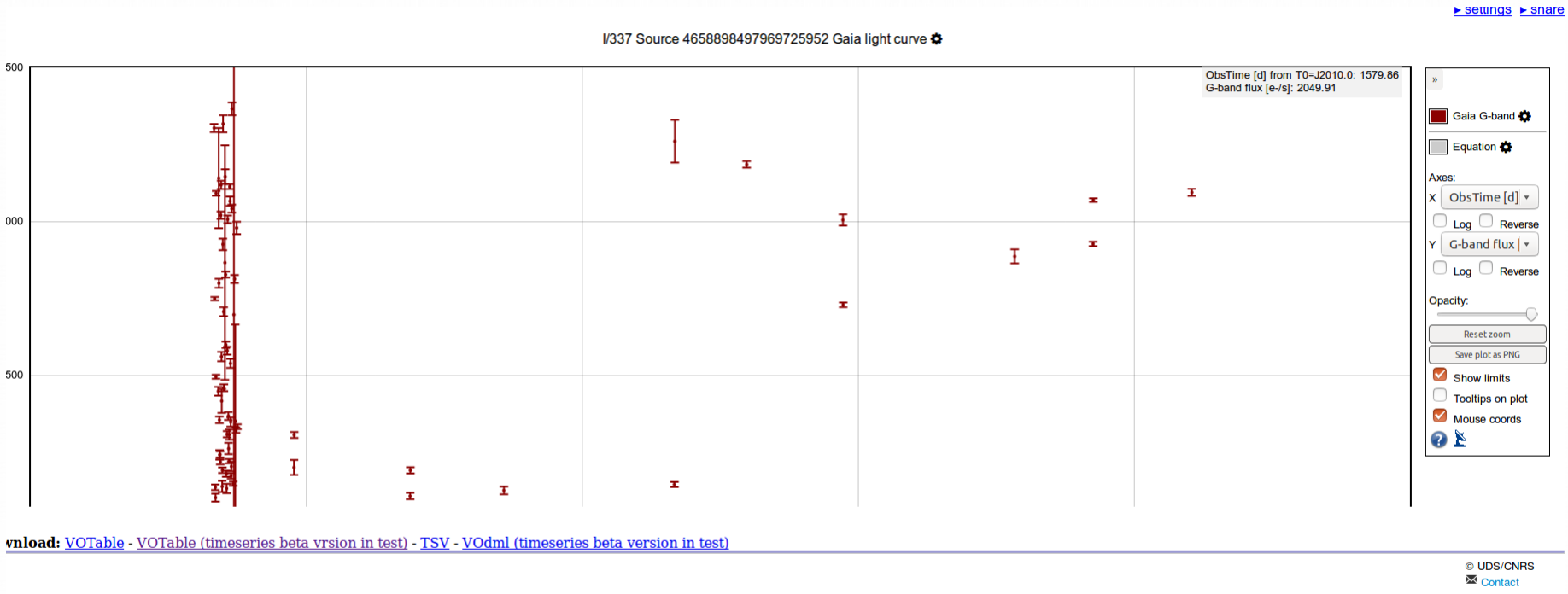
2016A&A...595A...1G ReadMe+ftp timeSerie



start AladinLite plot the output query using TAP/SQL

Full	LC	fov	TBest	TBest2	Mbest	Source	P1 d	EpG d	<Gmag> mag	AmpG mag	NHP1	R21G	phi21G	RA_ICRS deg	DE_ICRS deg	_RA.icrs deg	_DE.icrs deg
1	LC	fov	DCEP	—	UNDEFINED	4658898497969725952	0.81104349	1664.04407304	17.0100	0.419	3	0.193	4.139	80.4417418279	-66.9861900876	80.4417418279	-66.9861900
2	LC	fov	DCEP	—	FIRST_OVERTONE	4658898738488020864	3.38448730	1658.89869278	15.0480	0.338	4	0.182	3.695	80.4115015243	-66.9476771952	80.4115015243	-66.9476771
3	LC	fov	DCEP	—	UNDEFINED	4658925092406745984	2.69331244	1659.84418704	17.3720	0.112	1			78.8871350309	-67.1440231713	78.8871350309	-67.1440231
4	LC	fov	DCEP	—	FUNDAMENTAL	4658939214286774400	3.56278072	1658.55295617	15.6500	0.344	2	0.306	4.344	79.8382641810	-67.1136249309	79.8382641810	-67.1136249
5	LC	fov	DCEP	—	FIRST_OVERTONE	4658950381175117824	2.79569245	1660.51381113	15.1510	0.174	2	0.041	4.049	79.9507863890	-66.8238448295	79.9507863890	-66.8238448
6	LC	fov	DCEP	—	FUNDAMENTAL	4658956119278242688	5.22238334	1655.10218660	15.0250	0.862	4	0.477	4.468	79.4267360254	-66.6480295442	79.4267360254	-66.6480295
7	LC	fov	DCEP	—	FIRST_OVERTONE	4658960276779885056	1.39962121	1662.74142587	16.1550	0.362	3	0.203	4.138	79.0774160047	-66.7777281836	79.0774160047	-66.7777281
8	LC	fov	DCEP	—	FIRST_OVERTONE	4658968110800455040	2.18564218	1663.21217901	15.7010	0.289	2	0.115	4.385	79.2276629254	-66.6274827232	79.2276629254	-66.6274827
9	LC	fov	DCEP	—	FIRST_OVERTONE	4658969072873169536	2.44906328	1663.77790963	15.2650	0.337	2	0.060	4.418	79.1437679628	-66.5865626749	79.1437679628	-66.5865626
10	LC	fov	DCEP	—	FUNDAMENTAL	4658970241104217472	2.90669246	1661.81108296	16.0130	0.759	4	0.432	4.283	78.8939951447	-66.6467466	78.8939951447	-66.6467466
11	LC	fov	DCEP	—	FUNDAMENTAL	4659456740670442752	3.57527302	1658.73958036	15.4900	0.753	3	0.476	4.268	85.7639993508	-67.0764661314	85.7639993508	-67.0764661
12	LC	fov	DCEP	—	FIRST_OVERTONE	4659458527346797696	1.32562578	1663.71261361	16.1450	0.320	2	0.241	4.345	86.2315630903	-67.0800549791	86.2315630903	-67.0800549
13	LC	fov	DCEP	—	FUNDAMENTAL	4659460623290935168	5.42560958	1654.50555644	15.2020	0.413	3	0.337	4.668	86.2892914828	-67.0158001931	86.2892914828	-67.0158001
14	LC	fov	DCEP	—	FUNDAMENTAL	4659461241765373184	2.29321460	1660.71604944	16.0490	0.683	5	0.437	4.205	86.0456401016	-66.9992666985	86.0456401016	-66.9992666
15	LC	fov	DCEP	—	FIRST_OVERTONE	4659464024903476352	3.45100693	1657.81942313	15.0230	0.297	3	0.121	3.290	85.7003223028	-66.9427449711	85.7003223028	-66.9427449
16	LC	fov	DCEP	—	FIRST_OVERTONE	4659464883897843200	1.94143160	1660.88201578	15.7390	0.296	2	0.130	4.825	86.0436726396	-66.9232307845	86.0436726396	-66.9232307
17	LC	fov	DCEP	—	FIRST_OVERTONE	4659465227502800640	1.81105612	1661.77500792	15.7990	0.295	3	0.091	4.735	85.8820525670	-66.8796533252	85.8820525670	-66.8796533
18	LC	fov	DCEP	—	FIRST_OVERTONE	4659483339391441408	2.02794197	1661.46977585	15.5300	0.364	2	0.140	4.321	85.0131886781	-67.0716616361	85.0131886781	-67.0716616
19	LC	fov	DCEP	—	FUNDAMENTAL	4659494124040684032	7.47743782	1650.38164083	14.8800	0.180	2	0.143	5.660	84.6932441338	-67.0852792344	84.6932441338	-67.0852792
20	LC	fov	DCEP	—	FUNDAMENTAL	4659494879948476032	2.92545152	1660.07294073	15.7020	0.765	5	0.428	4.242	84.6697967355	-67.0349677043	84.6697967355	-67.0349677
21	LC	fov	DCEP	—	FIRST_OVERTONE	4659495154825994880	3.61472915	1657.36716036	14.9540	0.307	3	0.129	3.309	84.9383688599	-67.0564302421	84.9383688599	-67.0564302
22	LC	fov	T2CEP	W_VIR	NOT_APPLICABLE	4659497285129779584	12.34489304	1632.09072772	17.2710	0.128	1			84.4995522692	-67.0530549618	84.4995522692	-67.0530549
23	LC	fov	DCEP	—	FIRST_OVERTONE	4659499759031442432	3.83869397	1655.77401254	14.7890	0.277	3	0.149	3.529	84.4736181954	-66.9468487937	84.4736181954	-66.9468487
24	LC	fov	DCEP	—	FIRST_OVERTONE	4659502061133359232	3.65107544	1656.67707148	14.8270	0.295	3	0.101	3.186	85.5126413247	-67.1186067944	85.5126413247	-67.1186067
25	LC	fov	DCEP	—	FIRST_OVERTONE	4659510170032228736	2.56262634	1661.70271521	15.1910	0.374	2	0.114	4.396	85.0264708788	-66.8730142130	85.0264708788	-66.8730142
26	LC	fov	DCEP	—	FIRST_OVERTONE	5289779853168752384	2.57602789	1666.00543643	15.4970	0.110	1			119.2865231322	-62.3245006863	119.2865231322	-62.3245006
27	LC	fov	DCEP	—	FUNDAMENTAL	4659510307501954176	2.01301613	1660.46386339	16.2580	0.687	6	0.499	4.073	84.9958673647	-66.8370151586	84.9958673647	-66.8370151
28	LC	fov	DCEP	—	FUNDAMENTAL	4659512437774965120	4.96058526	1656.23198078	14.9100	0.760	5	0.464	4.437	85.4302310976	-66.8499643471	85.4302310976	-66.8499643
29	LC	fov	DCEP	—	FUNDAMENTAL	4659518588169281920	4.01889675	1656.19077016	15.5420	0.631	4	0.476	4.323	85.0478121914	-66.6840281518	85.0478121914	-66.6840281
30	LC	fov	DCEP	—	FIRST_OVERTONE	4659523845208113280	1.81235621	1661.43937585	16.0590	0.350	2	0.127	4.546	84.5374159062	-66.8811016974	84.5374159062	-66.8811016
31	LC	fov	DCEP	—	FUNDAMENTAL	4659525597553927680	8.72307463	1647.03095256	14.4750	0.364	3	0.289	5.651	84.3223320993	-66.7533061145	84.3223320993	-66.7533061
32	LC	fov	T2CEP	W_VIR	NOT_APPLICABLE	4659525872450052480	16.20356951	1621.96307480	16.8770	0.157	2	0.047	5.612	84.5891433732	-66.7890195521	84.5891433732	-66.7890195
33	LC	fov	DCEP	—	FIRST_OVERTONE	4659526044231432832	1.10378136	1663.85552967	16.5870	0.318	3	0.226	3.949	84.5292603186	-66.7851409564	84.5292603186	-66.7851409

VizieR prototype



TimeSeries discovery

- 3 discovery modes
  - Source driven (direct or via DataLink)
  - ObsCore/SIAV2-like driven (extensions are needed ?)
  - Physical Content driven (project specific?)

- One more in a near future --→ TMOC ?

VO Solutions

- Source driven discovery :
  - VO catalog services (« SCS » or « TAP » for VO friends)
  - Possibility to discover a TimeSeries in the response (« LINK », or « DataLink » access reference in the service response)

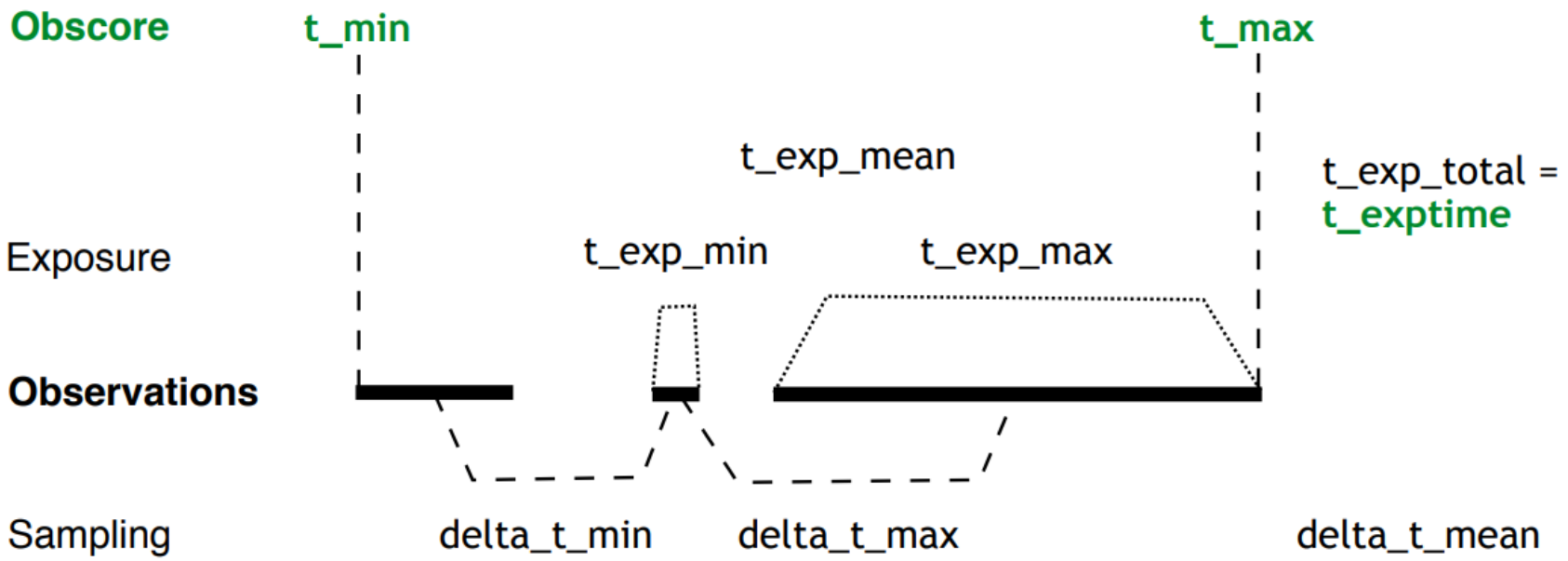
VO Solution :
TimeSeries via TAP or « simple access » protocols

- TimeSeries discovery requires additional parameters/attributes (see above)
- Extension of DAL VO (ObsTAP and « SIAV-2.0 ») protocols for these parameters/attributes

VO solutions : mixed discovery

- Extended ObsCore (TSCore ?) and a catalog describing sources managed together in the same « TAP » service

What do we have/need in Obscore for discovering time series data?



Number of observations **t_xel**

Time resolution (precision of the measurement) **t_resolution**

What kind of observations do I have? **photometry (light curves)**, spectra, images,..

ObsCore extension

- What should be added ?
  - Time Support (when do we have significant observation)
  - Time Support summury (min/max of « parts »)
  - Time sampling frequency, or frequency bounds
  - Time sample width bounds
  - Extend o_* domain : what is varying with time and how much ?

Data Access

- Assuming a representation/serialisation solution is completed
- 1) Full retrieval of archived TimeSeries
- 2) Interface (= « SODA ») to retrieve a « built on the fly » « Time Series » with a given :
  - Time range
  - Time resolution, sampling frequency
  - Representation, format