



4MOST – 4m Multi-Object Spectroscopic Telescope

What community PIs need to provide: target catalogues, templates, FoM, SSC, etc.

Wing-Fai Thi (MPE) Johan Comparat (MPE),
on behalf of the OpSys team

10 July 2020

www.4MOST.eu



Outline



- What is needed to make a 4FS Exposure Time Calculator (ETC) run and obtain a Scope of Survey fiber-hours?
- Overview of the steps to create a catalogue
- Science requirements:
 - Definition of a subsurvey
 - Scope of Survey Figure of Merit
 - Small Scale Merit
 - Large Scale Merit
 - A_{req} and T_{max}
 - Scope of Survey Figure of Merit Formula
- Required number of targets to be provided: two examples
- Input catalogue description
- Spectral Template
- Spectral Success Criterion: Rules and Rulesets
- An Exposure Time Calculator and Scope of Survey run
 - Assumptions used in the ETC
 - ETC and SoS output

What is a subsurvey and how many subsurveys?



- Targets within a subsurvey share the same Small Scale Merit (SSM), Large Scale Merit (LSM), and resolution mode (LR or HR)
 - If two categories of objects require different completeness level or spatial coverage requirements, each of the category should be assigned a subsurvey
- Within a survey, groups of targets may have several Spectral Success Criteria (SSC) and representative spectral templates
- At minimum one subsurvey or a set of subsurveys constitute a survey

What is needed to make an Exposure Time Calculation and obtain the SoS fiber hours?



- Access to the web interface: see talk by Vincenzo Mainieri
- A tutorial with more technical details will be given on Tuesday July 14th 14h15 “Demonstration session and tutorial on 4FS-Web Interface” by Jake Laas
- What is needed?
 - A catalogue:
 - The catalogue contains the coordinates, flux, spectrum, extend information, subsurvey membership, associated spectral success criterion for all **the** targets, ...
 - Spectral templates
 - **Spectral Success Criteria (SSC)** in form of **Rules** and **Rulesets**
 - Scope of Survey Figure of Merit (FoM) package:
 - **Small Scale Merit (SSM)**: Figure of merit as function of the target completeness
 - **Large Scale Merit (LSM)**: Map of the sky expressing the spatial scientific importance
 - Sky Area required (**A_{req}**): The scientifically required area to be observed
 - The maximum (cap) exposure time **T_{max}**

Survey Scope of Survey Figure of Merit



- We monitor the survey during the simulation/operation via the concept of **Scope of Survey (SoS) Figure of Merit (FoM)**, which is a function of the completeness
 - To compute the SoS FoM, survey P.Is have to provide indications (Small Scale Merit, Large Scale Merit, Areq) how the target completion can be translated into the SoS FoM.
 - The SoS FoM allows progress for vastly different scientific goals to be monitored on a same footage. Surveys have different completeness requirements, which have to be normalized to the same FoM value (for Consortium and Participating surveys).

Small Scale Merit: SSM



- The SSM expresses the relation between object observation completeness (fraction of observations done, i.e. between 0 and 1) and the science achievement level (also between 0 and 1).
- The completeness is defined as the ratio between the number of successfully observed targets and the **total number of targets N_{targets}** provided in the catalogue and **not** the scientifically required number of objects N_{req} with ($N_{\text{targets}} \geq N_{\text{req}}$).
- For Participating surveys, which will share observations and focal plane resources (fibres) with the 4MOST consortium and with other Participating surveys, a SSM=0.5 signals that all the main scientific requirements are met.

Small Scale Merit: SSM

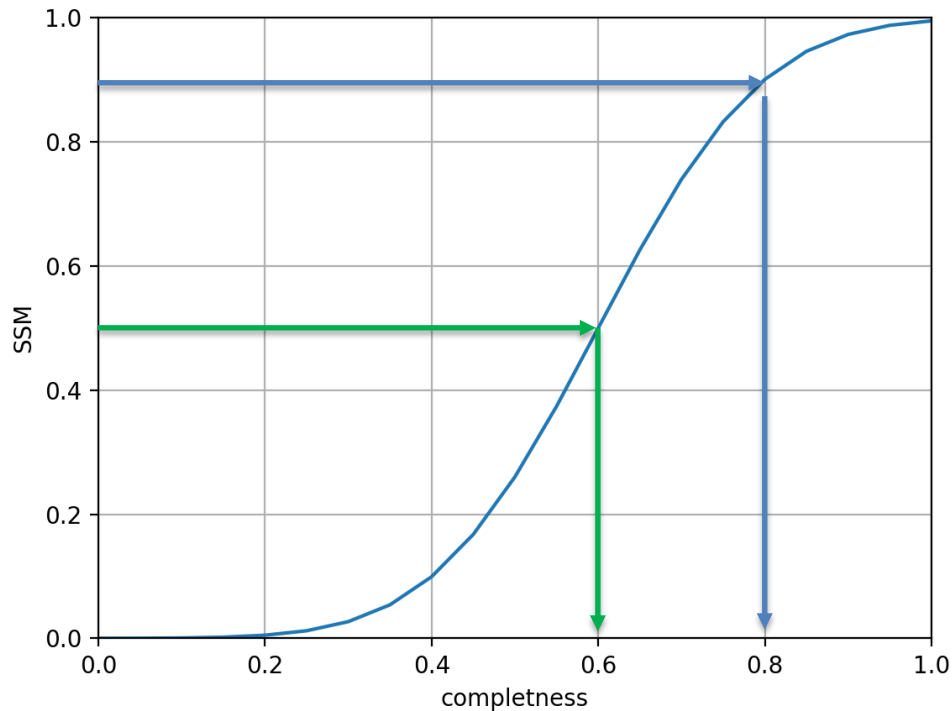


- Since for $SSM=0.5$, the target level completeness is less than 1, surveys provide more targets than needed (see later how to make a catalogue).
- A $SSM=0.9$ means that all the scientific goals are met.
- A SSM is unique for each subsurvey and is applicable to the entire footprint of the subsurvey. If vastly different level of completeness are needed for subset of objects in a survey, it would be better to split into two or more subsurveys.

SSM : example 1 Low completeness

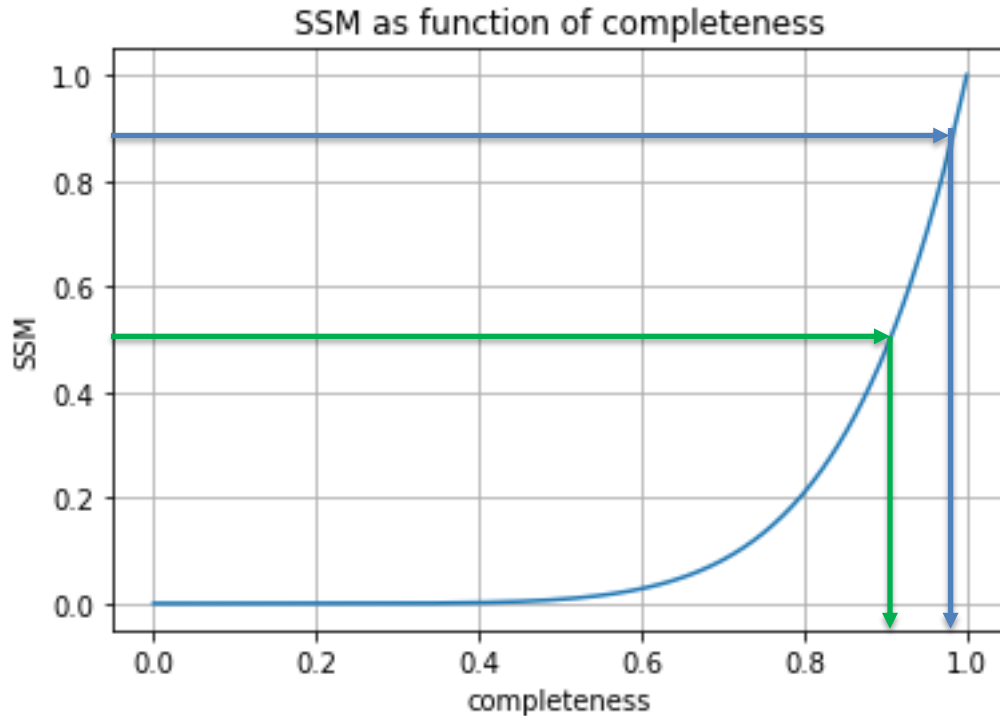


Subsurvey SSM



- SSM requirement (= 0.5) is met when completeness \geq 60%
- SSM goal (= 0.9) is met when completeness \geq 80%

SSM: example 2 High completeness



SSM = 0.5 corresponds to a completeness of ~ 0.905

SSM = 0.9 corresponds to a completeness of ~ 0.985

Large Scale Merit (LSM)



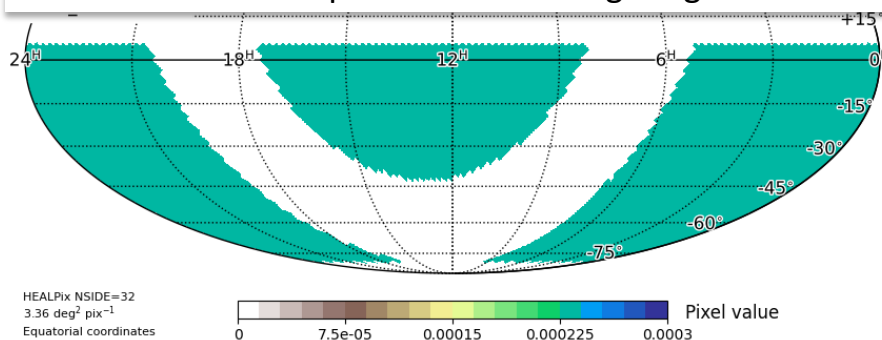
- Each subsurvey requires a Healpix-like map with weights. The weights are chosen to reflect the scientific importance of the given field in the sky relative to the other fields.
- The map is normalized (the sum of all the weights is equal to one)
- LSM is not used in the ETC computation but is part of the SoS FoM
- The Healpix map follows the format:
 - nested=Ring
 - Res = 5
 - nside = 32
 - Npixels = 12288 (full sky)
 - Mean spacing (deg) 1.8323

LSM examples

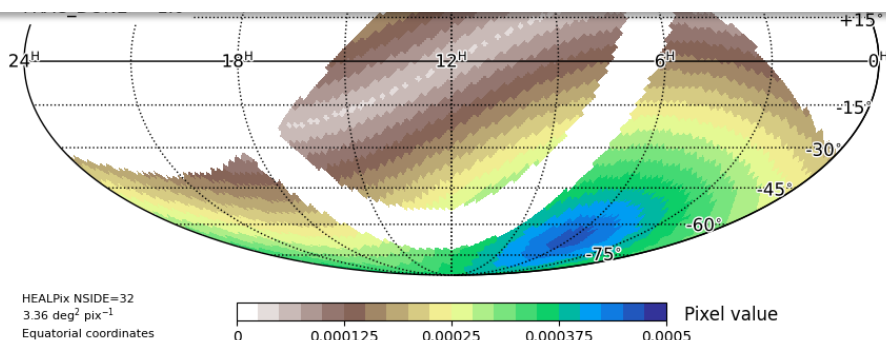


- 1) Spatially unbiased: a survey has 100,000 objects. The science requires that no region is preferred whatever the number of objects in the healpix. The weight for all healpix is $1/12288$
- 2) A survey has 100,000 objects. The science is best done when regions with the high target densities are observed, then the weights can be the object density with a final normalization to one.
- 3) Other examples:

An example of uniform weighting



An example of non-uniform weighting



Areq and Tmax



- Areq = goal area of the survey (degree²)
 - The goal area has to be lower or equal to the total area Atot (footprint) covered by the objects in the catalogues.
 - While the LSM provides a way to prioritize regions in the sky, Areq provides the actual area needed.
 - If the science requires all regions in the catalogue footprint to be observed, then Areq should be closed to the catalogue footprint area ($A_{req} \sim A_{tot}$, $A_{req} < A_{tot}$).
 - If the science requires targets to be observed independently on the regions of the sky where they are located (ie the science can be done with many of the targets observed in an area smaller than the catalogue target area), a smaller Areq can be used.
- Tmax: maximum requested exposure time
 - To avoid extremely long required exposure times, the user has the option to override the exposure times computed by 4FS_ETC on the basis of the Spectral Success Criterion by selecting a maximum exposure time for targets of the Subsurvey.
 - Example: 90% of the objects require 20 min and 10% 3hrs (180 min), the 10% faintest objects (or with the most stringent science success criterion high SNR) will require 40% of fiber hours. If Tmax is set to 2hrs, these 10% will require 30% of the total fiber hours, which also has decreased by 13%

Scope of Survey Figure of Merit



- The LSM, SSM and A_{req} are used in the progress monitoring of the survey during operation (and survey simulation) for each subsurvey:

$$SoS\ FoM = \frac{A_{tot}}{A_{req}} \sum_{all\ healpix} LSM\ weight\ in\ this\ pixel \times SSM\ value\ in\ pixel$$

- A_{tot} is the input catalogue footprint
- A_{req} is the required area
- SSM_val_pix is the SSM value for each pixel, i.e. the number of objects with scientifically valid data according to SSC within that pixel
- LSM provides weights for each of the pixels

How many targets should a participating survey provide?



- To maximise the multiplexing efficiency of the 4MOST global survey, an individual survey should provide more targets N_{targets} than the scientifically required number N_{req} for a SSM of 0.5.
- All progress statistics (SSM value, SoS FoM) are based on N_{targets}
- However the computed fiber hours are for N_{req} and not for N_{targets}
- Two ways to increase N_{req} to have N_{targets}
 - Increase locally (within one 4MOST pointing field of view) the density of targets: this will provide more possibilities for the fiber assignment, thus increasing the fiber usage efficiency
 - Increase the footprint of the input catalogue, i.e. having more objects spread over the reachable sky, independently of A_{req} : this increases the potential RA range to observe the targets, increasing the number of opportunities to observe an object of the survey.

Number of targets: example 1



- Example 1: the science requires **$N_{\text{req}} = 10,000$ objects** in $A_{\text{req}} = 5,000 \text{ deg}^2$ area.
- For $\text{SSM} = 0.5$, the object completeness is 0.9 (very high completeness because of the paucity of objects)
- The required density $\text{Density}_{\text{req}} = N_{\text{req}} / A_{\text{req}} = 10,000 / 5,000 = 2 \text{ objects deg}^{-2}$
- $\text{Density}_{\text{targets}} = \text{Density}_{\text{req}} / \text{completeness} = 2.22... \text{ deg}^{-2}$
- Since the targets are taken from an imaging survey covering $A_{\text{targets}} = 10,000 \text{ deg}^2$, the actual number of targets to be provided is: $N_{\text{targets}} = \text{Density}_{\text{targets}} \times A_{\text{targets}} = 2.22... \times 10,000 = \mathbf{22,222 \text{ objects}}$
- The PI has to provide **22,222 objects** for the input catalogue, i.e. more than two times the scientifically required number
- In general, the rarer is a class of object, the higher is the needed completeness.
- The higher is the completeness, the more difficult is to reach it.

Number of targets: example 2



- Example 2: a follow-up of a sky survey with a very large number of potential targets (1,000,000). In that case the required area is as large as the footprint of the input catalogue ($A_{\text{tot}}=15,000 \text{ deg}^2$).
- The input survey has in average $\text{Density}_{\text{input}} = 1,000,000 / 15,000 \sim 67 \text{ objects deg}^{-2}$
- The science requires:
 - **$N_{\text{req}} = 200,000 \text{ sources}$** in $A_{\text{req}} = 15,000 \text{ deg}^2$
 - Each target needs ~ 30 min of exposure time in Grey condition. The required Fiber-hours is 100,000 hrs
 - $\text{Density}_{\text{req}} = 13.333 \text{ objects deg}^{-2}$
- For $\text{SSM}(0.5)=0.6$, the density of objects is increased by $1/0.6=1.666\dots$:
 - $\text{Density}_{\text{targets}} = 22.222 \text{ objects deg}^{-2}$
- The team has to select from their input catalogue $\sim 23 \text{ objects deg}^{-2}$ out from $\sim 67 \text{ objects deg}^{-2}$
- The team has to provide **$\sim 333,334 \text{ objects}$** with a density of $\sim 23 \text{ objects deg}^{-2}$ covering $15,000 \text{ deg}^2$ for a required 100,000 Fiber-hours
- The ETC run will give a more precise Fiber-hours

Catalogue file input: one entry



- NAME: name of the target (string).
- RA, DEC: double in degrees (Reference frame Gaia DR2)
- PMRA, PMDEC: double in mas/yr proper-motion of the target
- EPOCH: years, it is a string (example 2019.5)
- RESOLUTION: required resolution mode 1=LRS, 2=HRS
- SUBSURVEY: Sub-survey specification, a name to the subsurvey (string, see rules applicable to strings)
- TEMPLATE: Spectral template filename
- RULESET: Name of the Ruleset containing the spectral success criteria that should be applied to test if this target has been successfully observed.

More catalogue file input



- MAG: magnitude of the target in the passband filter system defined by MAG_TYPE
- MAG_TYPE: Identifier for the magnitude or flux density system (expressed as “Filter System”) in which each entry in the MAG column is measured. Only a limited set of standard “Filters” and standard “Systems” are accepted
- DATE_EARLIEST: Earliest date when this target can be observed in Julian Date (decimal days)
- DATE_LATEST: Latest date when this target can be observed in Julian Date (decimal days)

Extended sources (catalogue file)

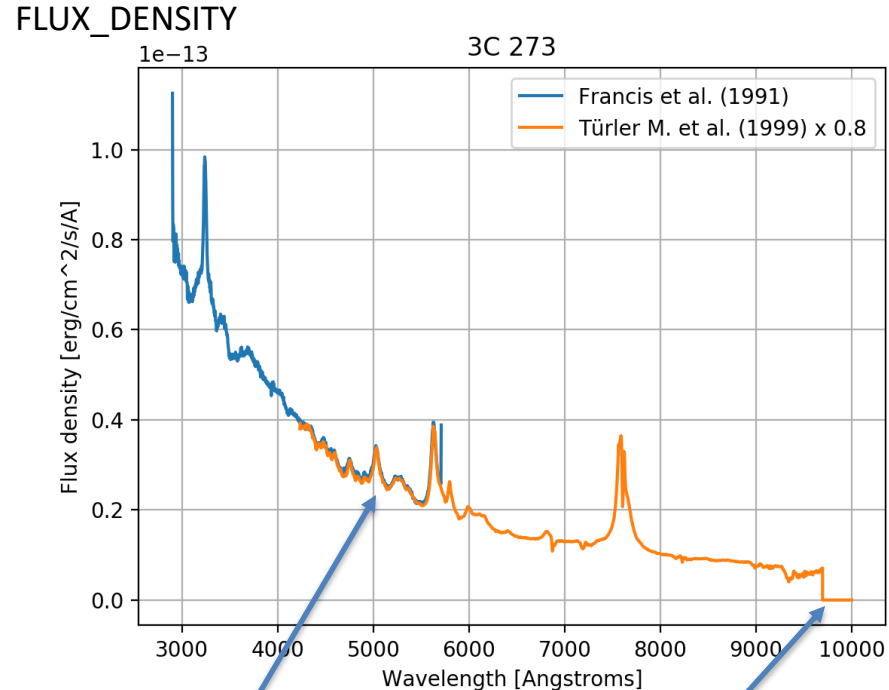


- Three types of target spatial shapes can be defined:
 - Flat source: the magnitude column will be considered as $\text{Mag}/\text{arcsec}^2$
 - Point source: total magnitude of a point source
 - Sersic 2D source (spherical symmetric source):
 - The Sersic sources are characterized by:
 - EXTENT_INDEX: Sersic index n (valid values from 0.5 to 10). Set to 0 for Point/Flat source)
 - EXTENT_PARAMETER : Sersic profile spatial extent r_{eff} (in radius in arcsec) of the target (valid values from 0.1 to 100 arcsec)
 - For Sersic sources, the flux as seen by the fiber (Fiber magnitude) is computed from the input magnitude (i.e. the total flux of the galaxy) by convolving the Sersic profile with a Moffat ($\alpha=2.5$) PSF profile
 - The PSF depends on the seeing as estimated from the object declination (in bins of 0.1 airmass)

Spectral templates



- A spectral template is the expected spectral energy distribution of a target. It can be used for several objects with varying magnitudes and/or spatial extent
- The ETC code will scale the required exposure time given the magnitude. The ETC computation time is proportional to the number of spectral templates.
- The wavelength coverage of the templates has to be larger than the low-resolution 4MOST wavelength coverage. In case no data is available in certain wavelength regions and the spectral success criterion is not applicable to these regions, flux density with zero values should be filled in.



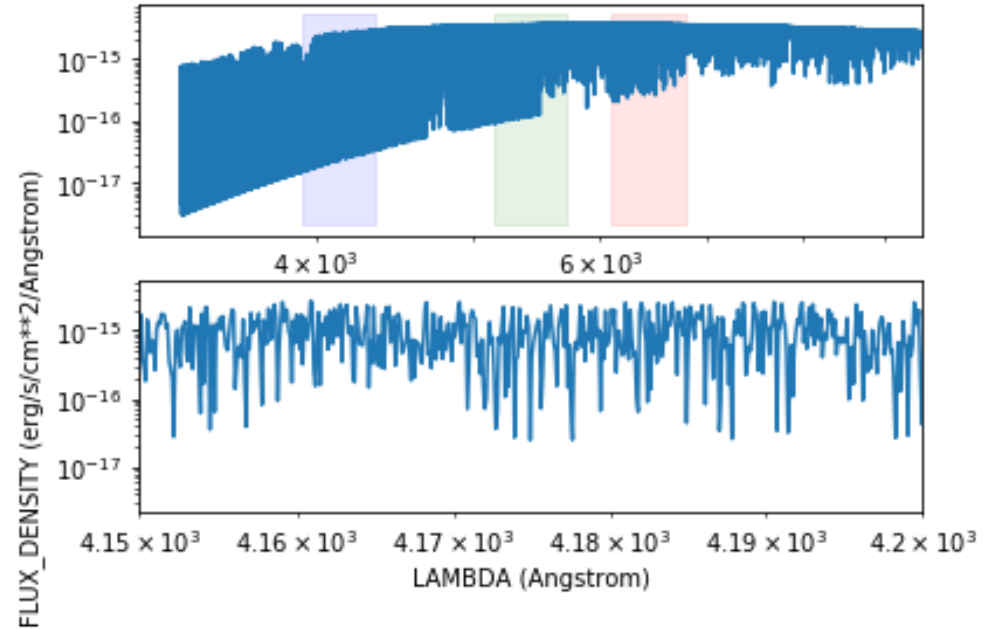
Merging of two input spectra into one unique template

Add zero fluxes

Spectral templates



Example of a HR spectral template
The 4MOST wavelength coverage is shown by the blue, green, and red boxes.

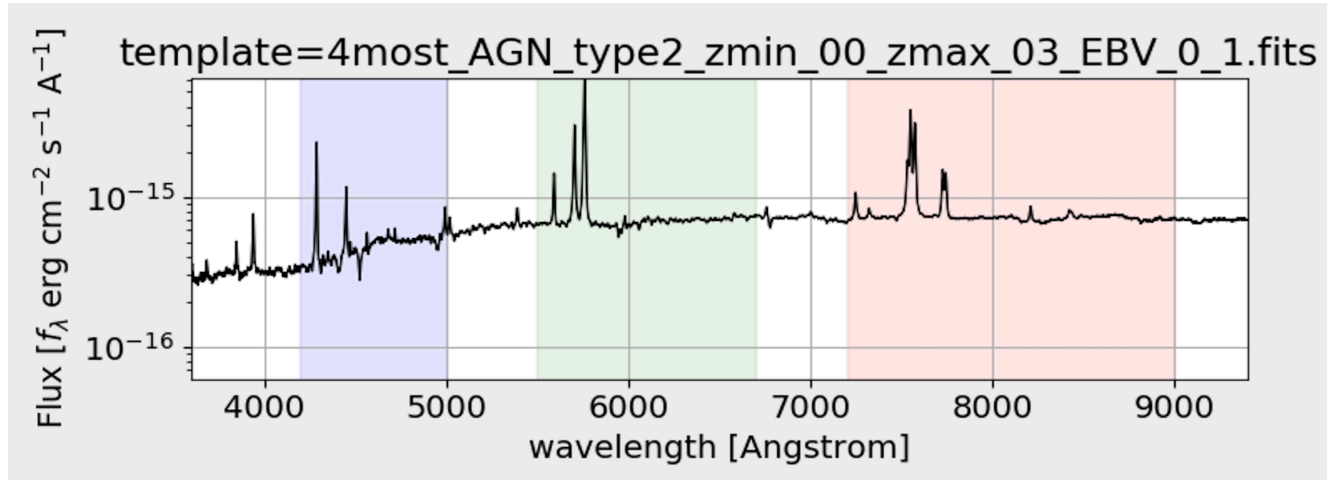


Spectral Success Criterion (SSC)



- SSC are required by a spectrum to be considered successfully observed
 - SSC (e.g. median SNR in a wavelength range) are used to judge the scientific worthiness, such that astrophysical quantities can be derived
- SSC have been designed to enable surveys to encode complex combinations of logical tests
- A SSC has two parts:
 - Rules: a rule is a simple expression applied to the spectra. For example, the median(SNR) between 4500 and 5000 Å. A rule should be applied within the wavelength coverage of each 4MOST band.
 - Rulesets: a ruleset is a combination of rules
- Rules and rulesets inputs follow specific format (see tutorial)
- Rules and rulesets can be expressed:
 - with many statistics: flux, SNR, ...
 - mathematical functions: min, max, mean, median, percentile, ...
 - Logical functions: and, or

SSC: Rules and Rulesets



Example 1: Detect the continuum in at **least one** of the 4MOST LR bands with SNR of at least 5

Rule_blue = median(SNR/Angstrom) between 4200 and 5000A

Rule_green = median(SNR/Angstrom) between 5500 and 6700A

Rule_red = median(SNR/Angstrom) between 7200 and 9000A

The ruleset will be

ruleset = (Rule_blue>5) or (Rule_green>5) or (Rule_red>5)

SSC: Rules and Rulesets



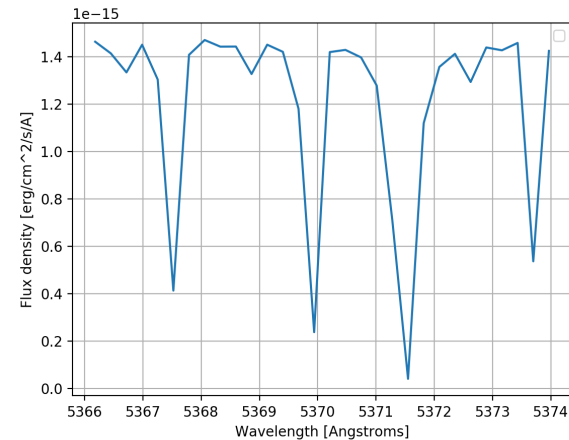
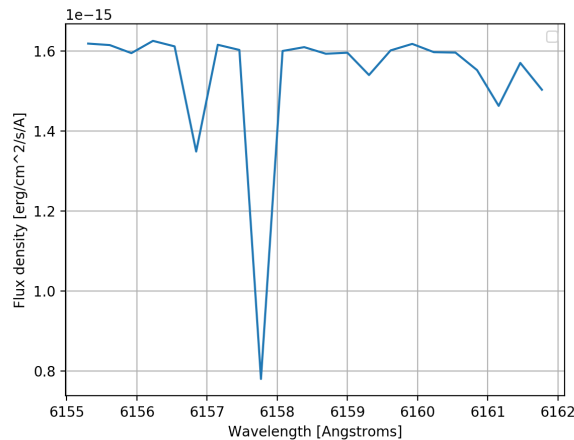
Example 2: Detect the FeI 6157.7A (in the HR red band 6100–6790A) + FeI 5373.71A (in the HR Green band 5160–5730A) with high SNR

- Rule_red = median(SNR/Angstrom) between 6155 and 6162A
- Rule_green = median(SNR/Angstrom) between 5366 and 5374A

The ruleset will be

- AverageStar = (Rule_green>200) & (Rule_red>200)
- BrightStar = (Rule_green>300) & (Rule_red>300)

This ruleset will be true only if the individual rules are true



Some assumptions in the ETC



Typical conditions used: see Genoveva Micheva's talk ETC algorithms and assumptions

- Three different moon illumination conditions ("FLI"). In particular, we define, following ESO's definitions:
 - dark ('d') $FLI < 0.4$,
 - grey ('g') $0.7 > FLI \geq 0.4$,
 - bright ('b') $FLI \geq 0.7$.
- airmass: each target is assigned a typical airmass with which its observation could be performed. The assigned airmass scales with the declination of the target.
- The maximum exposure time is 30 minutes. Exposures longer than 30 minutes are sum of 30 minutes sub-exposures. The dark current and readout noise apply to each sub-exposure.
- Sersic parameter ranges:
 - Sersic index n between 0.3 and 10
 - Sersic effective radius (arcsec) between 0.1 and 100
- All objects with exposure time below some threshold will be set to the threshold value:
 - Bright time : minimum 7 min
 - Grey time : minimum 10 min
 - Dark time: minimum 10 min

ETC and Scope of Survey run



- After uploading and validating all the input files:
 - A ETC run can be requested. The run time varies from few tenths of minutes to a few hours. Typically, the ETC computation time scales approximately with:
 - the number of templates
 - the sampling resolution of the input templates
 - the number of rulesets
 - the magnitude range for each class of objects
 - the number of resolution modes
 - the number of targets
- The airmass and fiber magnitude used for the calculation are provided as extra columns
- The ETC will generate a copy of the input catalogue with three added columns corresponding to the exposure time for dark (TEXP_D), grey (TEXP_G), and bright (TEXP_B) sky condition.
- After a ETC run has been successfully completed, a scope of survey (SOS) run can be requested.
- A SOS run provides a report with the requested **SoS fiber-hours** (lower limit to the actual fiber-hours since this is not an Operation Simulation) given the input for the three sky conditions for SSM=0.5 and SSM=0.9

Scope of Survey Fiber Hours



$$SoS fh_{B, G, D}(SSM = 0.5) = \frac{A_{req}}{A_{tot}} \sum T_{exp B, G, D}(N[SSM = 0.5])$$

- $N[SSM=0.5]$ ($N[SSM=0.9]$) targets required to reach $SSM=0.5$ ($=0.9$) are randomly selected from the catalogue.
- $T_{expB,G,D}$ are the exposure times in Bright, Grey, and Dark conditions
- The exposure times of the $N[SSM=0.5]$ and $N[SSM=0.9]$ targets are summed up and multiplied by the factor A_{req}/A_{tot} .
- A_{tot} is the footprint of the input catalogue and is calculated by the ETC.

Scope of survey output 1



Survey	Subsurvey	R	Tmax [min]	sky	N_{targets}	Frac $t=t_{\text{max}}$	Frac $t=t_{\text{min}}$	Areq [deg ²]	Atot [deg ²]	fA	NSIDE
SX	SX_ex1	2	120	D	1184420	0	0.41	14000	19149	0.73	32
SX	SX_ex1	2	120	G	1184420	0	0.41	14000	19149	0.73	32
SX	SX_ex1	2	120	B	1184420	0	0.26	14000	19149	0.73	32

Resolution:
LR =1, HR=2

Tmax set by the PI

Dark, Grey, Bright

Total number of targets (in the catalogue)

Areq < Atot
(from the catalogue)
fA = Areq/Atot

Healpix NSIDE used to estimate Atot


Fraction of targets in the catalogues that
are too bright or too faint (here 0%)

Scope of Survey output 2

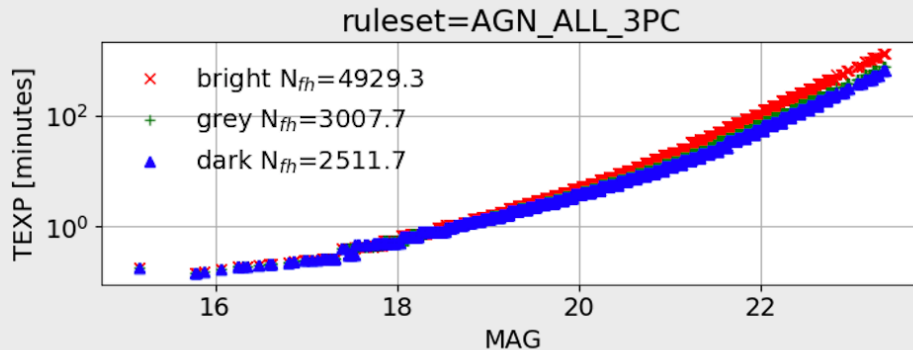
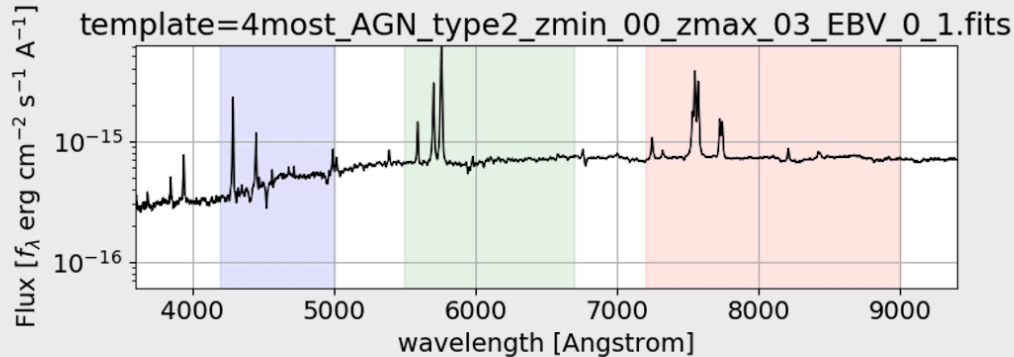


Survey	Subsurvey	R	Tmax [min]	sky	Total FH SSM=0.5	Total FH SSM=0.9
SX	SX_ex1	2	120	D	1.93×10^5	2.46×10^5
SX	SX_ex1	2	120	G	2.09×10^5	2.67×10^5
SX	SX_ex1	2	120	B	2.44×10^5	3.12×10^5

FH: fiber hours



ETC & SOS run output

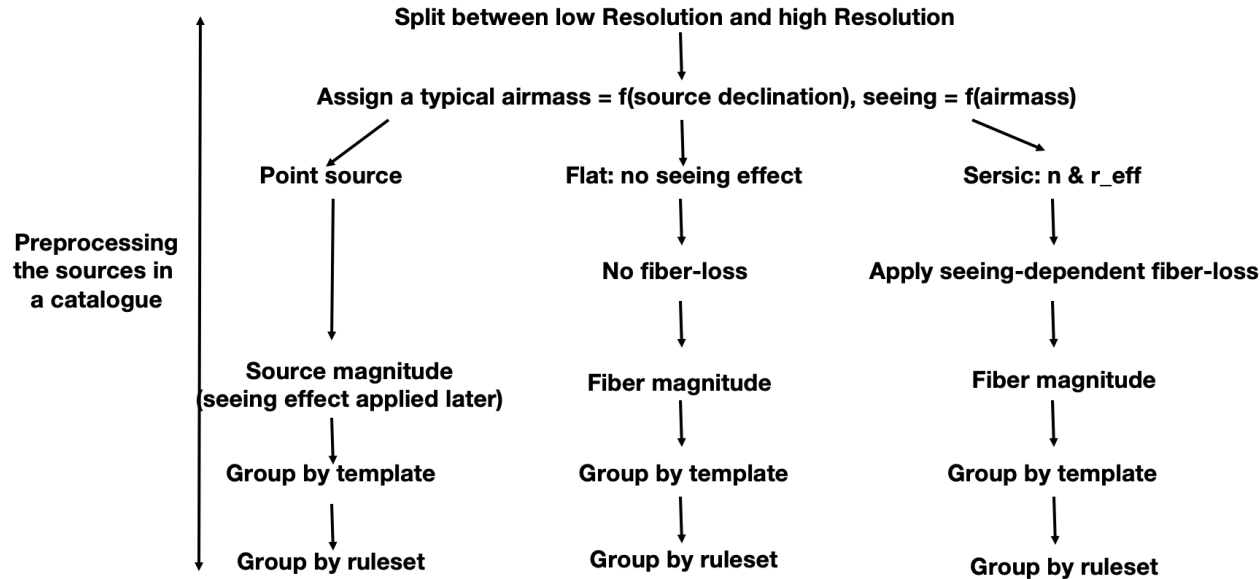


The output given a specific template, ruleset and rules applicable to a few thousand objects is shown.

N_{fh} is the fiber hours for bright, grey, and dark time.

N.B. These plots are not provided on the web interface





The preprocessed catalogue will have the sources grouped by resolution and subgrouped by target shape, template, and ruleset

Two columns are added: typical airmass and Fiber magnitude

