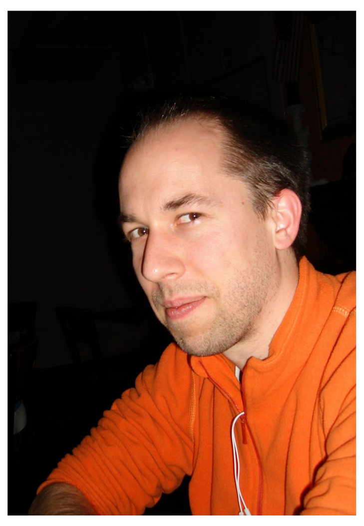
P2-5

Coverage and lacunarity of astronomical datasets

CENTRE DE DONNÉES ASTRONOMIQUES DE STRASBOURG

Sébastien DERRIERE
sebastien.derriere@astro.unistra.fr



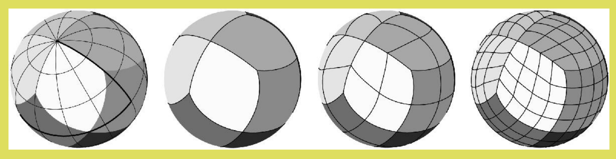
Abstract

We present the use of MOCs (Multi-Order Coverage maps) associated to HiPS datasets (Hierarchical Progressive Surveys) to assess the sky coverage of various datasets (image surveys and source catalogues), but also their lacunarity (presence of holes).

The goal is to :

- Estimate at best the sky coverage of a dataset
- Find a metric to quantify the scattering of data on the sky (distinguish contiguous datasets from patches of data scattered over large areas, or containing many holes)
- Apply this to source catalogues, without oversampling

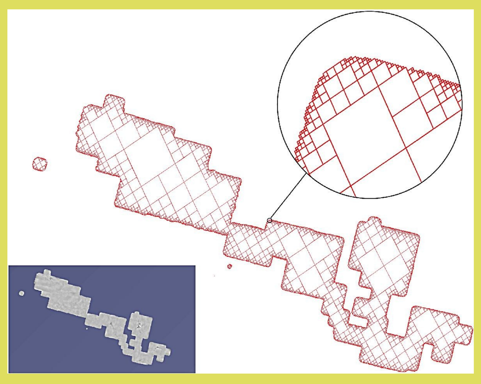
HiPS and MOCs



See Fernique et al. (2015)

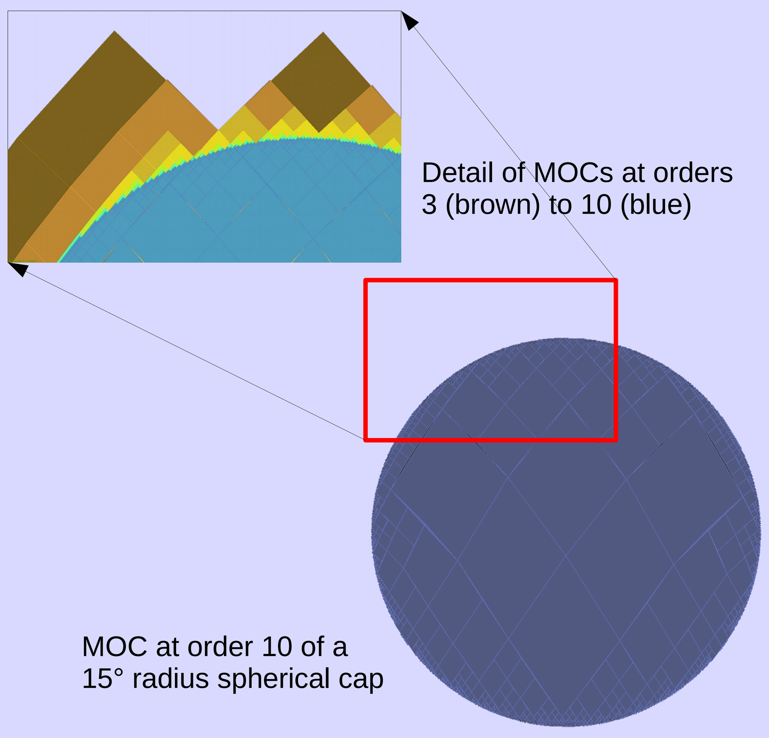
- Based on the HEALPix tesselation of the sphere (equal area)
- Progressive surveys (HiPS) and coverage maps (MOCs) allow to describe complex datasets
- Precision increases with order *k*

k	$N_{side} = 2^k$	N_{pix}	θ_{pix}	$k_{tile,512}$	$N_{tile,512}$	$\theta_{tile,512}$
0	1	12	58°.6			
1	2	48	29°.3			
2	4	192	14°.7			
3	8	768	7°.33			
4	16	3072	3°.66			
5	32	12 288	1°.83			
6	64	49 152	55′.0			
7	128	196 608	27′.5			
8	256	786 432	13′.7			
9	512	3 145 728	6′.87	0	12	58°.6
10	1024	12 582 912	3′.44	1	48	29°.3
11	2048	50 331 648	1′.72	2	192	14°.7
12	4096	201 326 592	51″.5	3	768	7°.33
13	8192	805 306 368	25″.8	4	3072	3°.66
14	2^{14}	3.22×10^9	12″.9	5	12 288	1°.83
15	2^{15}	1.29×10^{10}	6″.44	6	49 152	55′.0
16	2^{16}	5.15×10^{10}	3″.22	7	196 608	27′.5
17	2^{17}	2.06×10^{11}	1″.61	8	786 432	13′.7
18	2^{18}	8.25×10^{11}	0″.81	9	3 145 728	6′.87
19	2^{19}	3.30×10^{12}	0″.40	10	12 582 912	3′.44
20	2^{20}	1.32×10^{13}	0″.20	11	50 331 648	1′.72
21	2^{21}	5.28×10^{13}	0″.10	12	201 326 592	51″.5
22	2^{22}	2.11×10^{14}	50.3 mas	13	805 306 368	25″.8
23	2^{23}	8.44×10^{14}	25.1 mas	14	3.22×10^9	12″.9
24	2^{24}	3.38×10^{15}	12.6 mas	15	1.29×10^{10}	6″.44
25	2^{25}	1.35×10^{16}	6.29 mas	16	5.15×10^{10}	3″.22
26	2^{26}	5.40×10^{16}	3.15 mas	17	2.06×10^{11}	1″.61



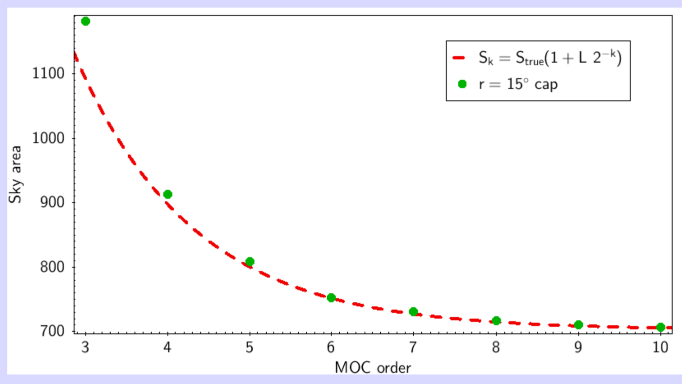
IVOA standards

MOCs of mock data



- MOCs can be used to estimate the sky coverage, by adding the areas of all HEALPix pixels.
- The total area is overestimating the true area, because pixels spread on the perimeter.
- The highest the order of the MOC, the better the precision.
- The area of a HEALPix pixel at order *k* is $\Omega_{pix} = (\pi/3)\ 4^{-k}$, and the estimated area from a MOC at order *k* is related to the true area by a relation of the form $S_k = S_{true}(1 + L \cdot 2^{-k})$ where L is a parameter related to the perimeter of the MOC.
- More complex MOCs will have larger values of L.
- We define $\mathcal{L}$ = log(L) as the lacunarity of the dataset : small values will indicate simple and contiguous data distribution, while large values will characterize scattered or irregular data.

The notion of lacunarity was introduced by Mandelbrot to describe the texture of fractals, and was extended (see for example Dong (2000)) to describe the distribution of gap sizes in geometric objects : homogeneous objects have low lacunarity, and heterogeneous objects have high lacunarity.



Computing the MOCs at different orders k for a r=15° spherical cap, and fitting the formula over orders k=7 to 10 gives an estimated S_{true}=704.5deg², very close to the exact area S=702.8deg², with $\mathcal{L}$ = 0.648.

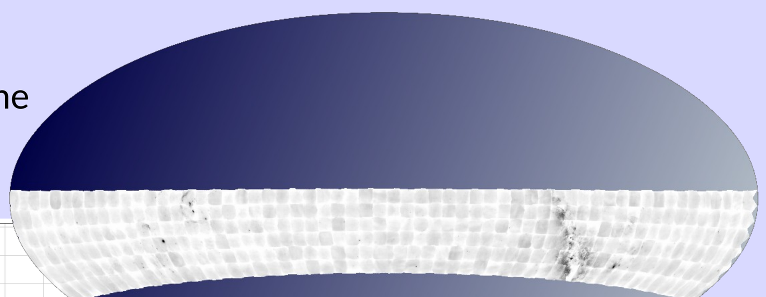
For a MOC corresponding to 3 disjoints caps, the fit yields S_{true}=2109.4deg², very close to the exact area S=2108.4deg², with $\mathcal{L}$ = 0.715.

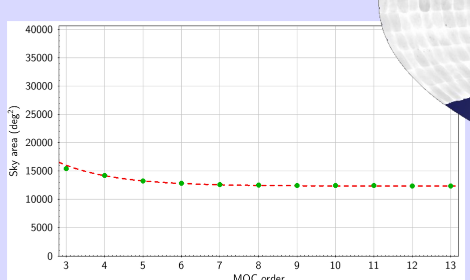
Application to image HiPS

For 3 example image HiPS, we compute MOCs from order *k*=3 down to a maximum order, and study the evolution of the MOC area with *k*.

CDS/P/MAMA/posse

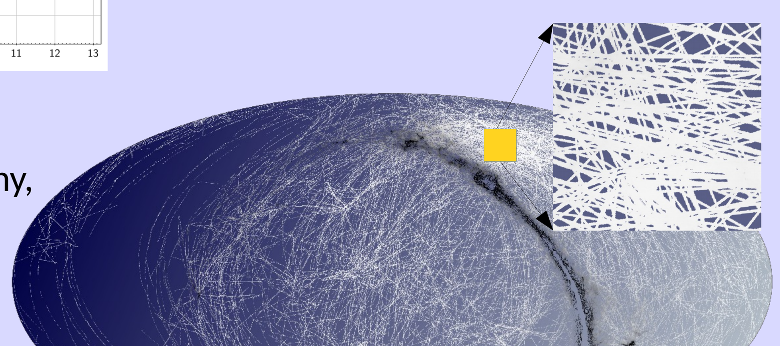
This survey corresponds to a contiguous fraction of the sky, with a simple boundary.
Areas derived from MOCs quickly converge to the corresponding sky coverage.

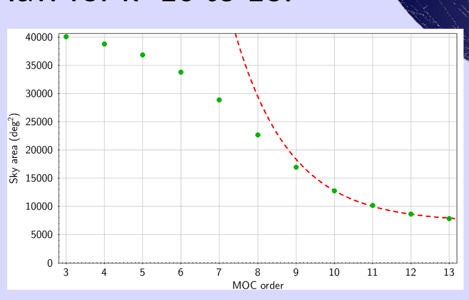




CDS/P/ISOPHOT/170

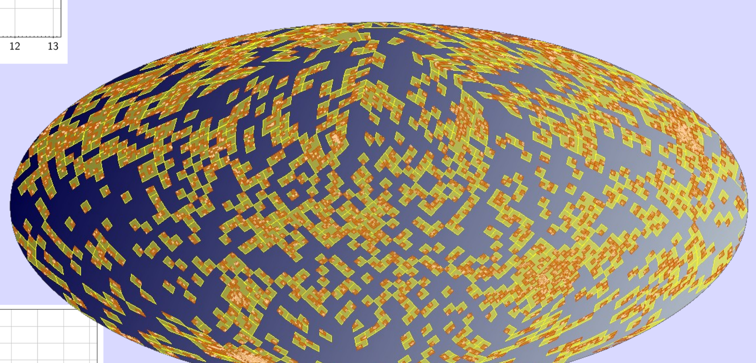
The ISOPHOT 170um serendipity survey is very patchy, with many unobserved holes in the slew survey.
At low orders, the MOC pixels merge and the total MOC area becomes nearly all-sky.
It is possible to fit the above law for k=10 to 13.

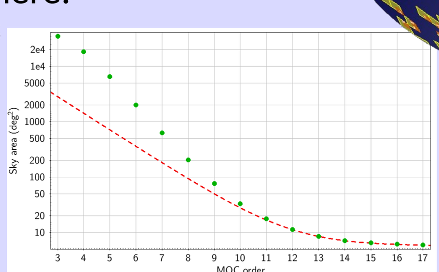




CDS/P/HST/V

This compilation of HST observations around the V band contains many small-area pointed observations scattered over the sky, as can be seen from the MOCs at order 5, 7 and 9 represented here.
The MOC area quickly decreases with *k*, and we need to reach order *k*=17 to start converging towards the true coverage.





Lacunarity increases for heterogeneous datasets.

Results :

Comparing the coverages obtained here to the ones quoted in the CDS MOC server, one sees that the `moc_sky_fraction` for the ISOPHOT survey, computed at k=10, is overestimated.

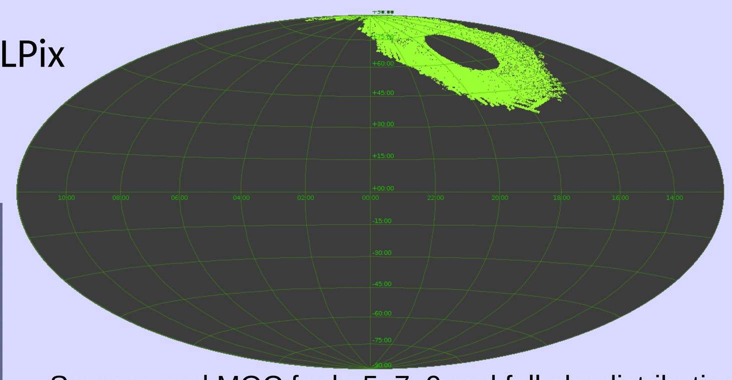
Survey	S_{true} (deg2)	$\mathcal{L}$	MOC Server (deg2)
MAMA/posse	12358	0.377	12425
ISOPHOT	7252	2.896	12771
HST/V	5.745	3.595	5.978

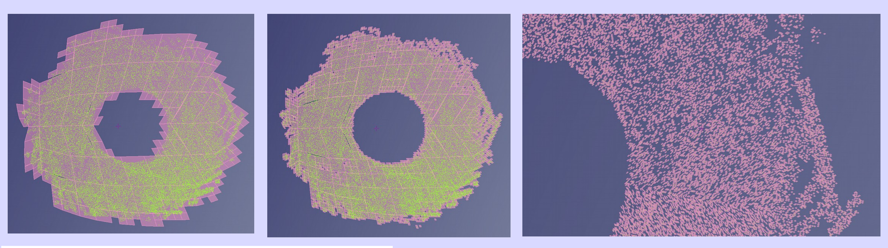
Application to catalogues

Estimating the sky coverage of a catalogue from the positions of its individual sources is more difficult than for an image survey : if the density of sources per HEALPix pixel becomes too low at high orders k, empty pixels cause an underestimation of the area.
To illustrate this, we study here two VizieR catalogues with a similar number of sources, but very different distributions on the sky.

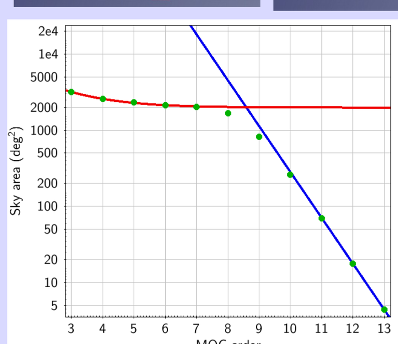
CDS/I/335/table1 : LUT Survey Catalogue Data Release 1 (Men et al. 2016)
N=86 467 sources.

When k is too large, there is a single source per HEALPix pixel, and the MOC area becomes $S_k = N(\pi/3)\ 4^{-k}$ (blue dashed line in the graphs).



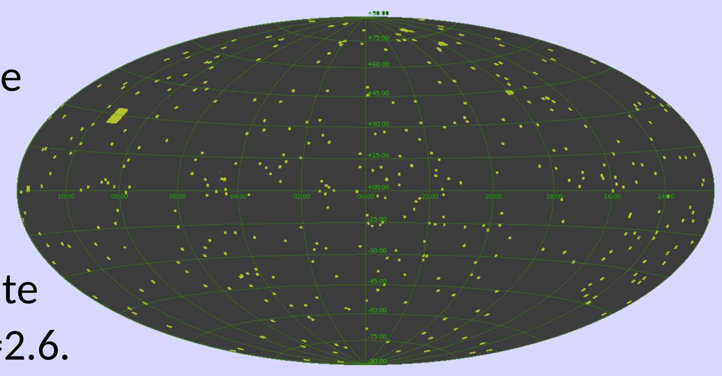


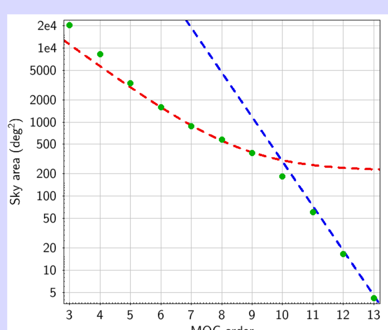
Sources and MOC for k=5, 7, 9 and full-sky distribution for I/335/table1.



k=3 to 6 allow to estimate S_{true} =1984 deg2 and $\mathcal{L}$=0.683. For k>7, empty pixels in the MOC give a wrong coverage.
The MOC server gives S=69.39 deg2, using k=11.

CDS/I/273A/erlcat : Extragalactic Reference Link Catalog (de Vegt et al., 2001)
N=89 422 sources.





k=6 to 9 allow to estimate S_{true} = 219.5deg2 and $\mathcal{L}$=2.6.
The MOC server gives S=60.39 deg2, using k=11.

Conclusions: MOCs can provide a very convenient way to estimate the sky coverage of image surveys or catalogues. The analysis of the variations of the coverage with the MOC order also provide information on the scattering within the dataset, for example using the lacunarity introduced in this study.

ADASS XXIX – Groningen – 2019 - CDS – Observatoire astronomique de Strasbourg

References :
Dong, P., Annals of GIS, 6, 20 (2000)
Fernique et al., A&A 578, A114 (2015)

Observatoire astronomique de Strasbourg | ObAS

cnrs

Université de Strasbourg