

Aggregates

- In the archive, Chandra Science Papers (CSPs) are associated with a given observation, but
 1. Individual observations form proposals
 2. Proposals grouped together form larger projects, which often span different proposal types (e.g. CDF used DDT, GO, GTO time to accumulate the final exposure)
- AGGREGATES are groups of observations that form a coherent whole. This may be a proposal, or multiple proposals. These have been identified through the literature, asking people, and Obscat.
- Examples: CDFS (4 Msec), DEEP2 (3.7 Msec) , M101 (1.1 Msec)



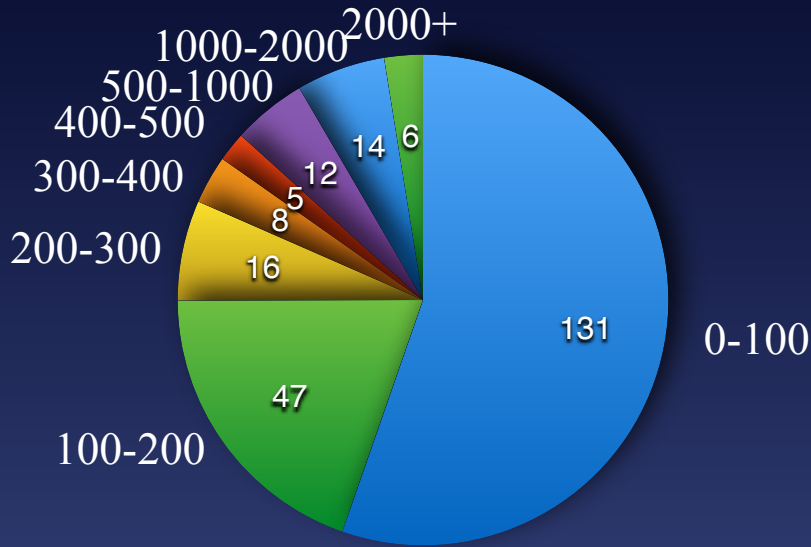
Weighted Publication Rate

For each aggregate:

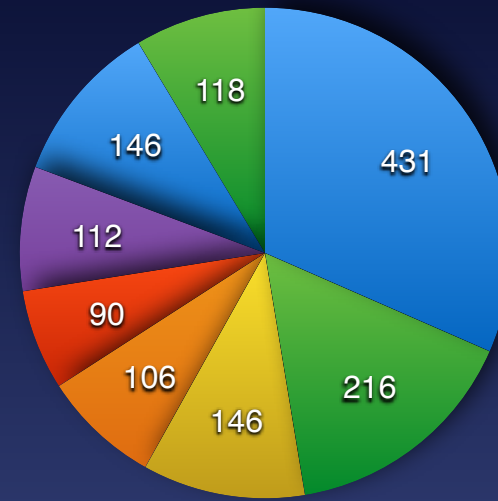
- Collect all CSPs which used data from the aggregate. Weight the paper by:
 - Percentage of time in paper attributed to the aggregate (to account for papers which use data from multiple aggregates)
 - Percentage of time in aggregate attributed to paper (to account for papers which do not use the entire aggregate)
- Determine the weighted publication RATE as
 - $(\text{sum of weighted CSPs}) / (\text{age of aggregate})$
- Use of publication RATE for each aggregate allows aggregates of different ages to be combined in a bin (e.g 0-100 ks, 200-300 ks etc).
 - appropriate for the Chandra archive: publication statistics indicate that the entire archive is still being used (i.e. the oldest data in the archive are still being published).



Effectiveness of LPs and XVPs: Publication statistics



Weighted publication rate as a function of aggregate size:
many more smaller aggregates!

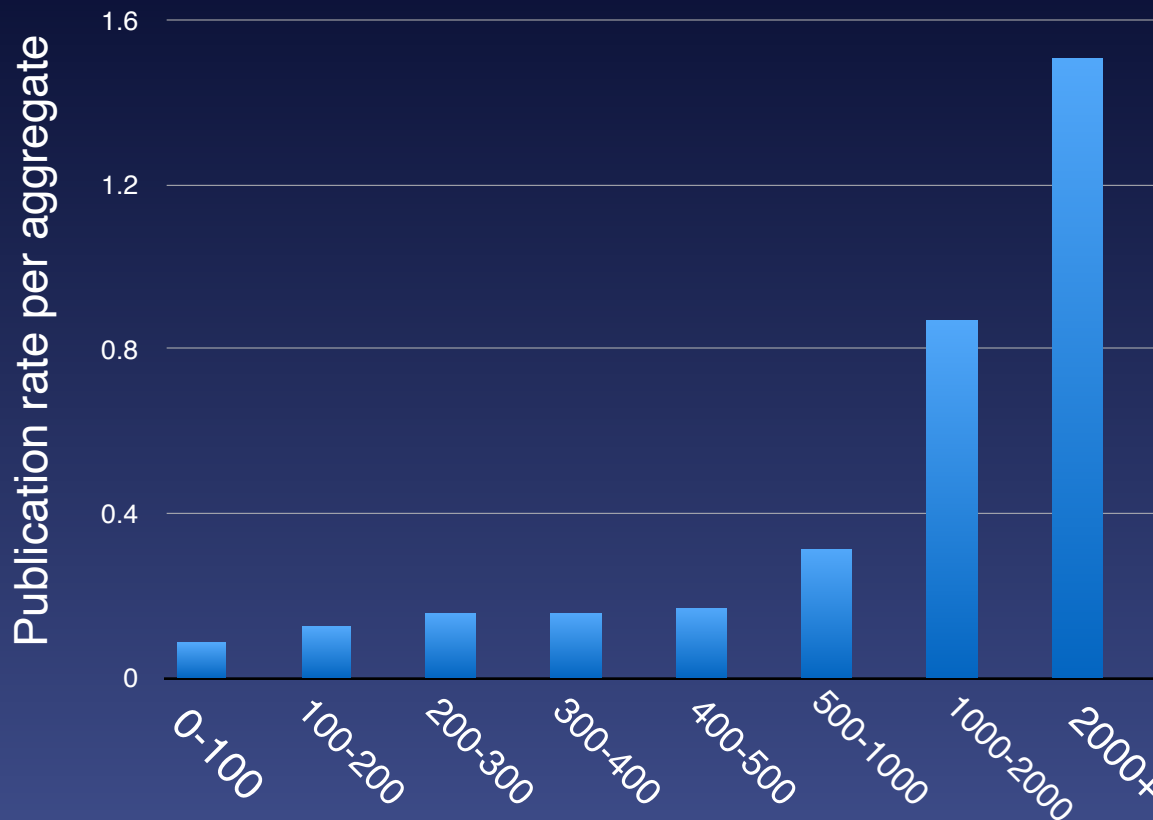


Weighted publication rate as a function of aggregate size, normalizing by total time.
Slices would be same size if publication rate was independent of aggregate size.

As expected, XVPs have publication rates lower than the smallest aggregates, but comparable when normalized by time.



Effectiveness of LPs and XVPs: Publication statistics

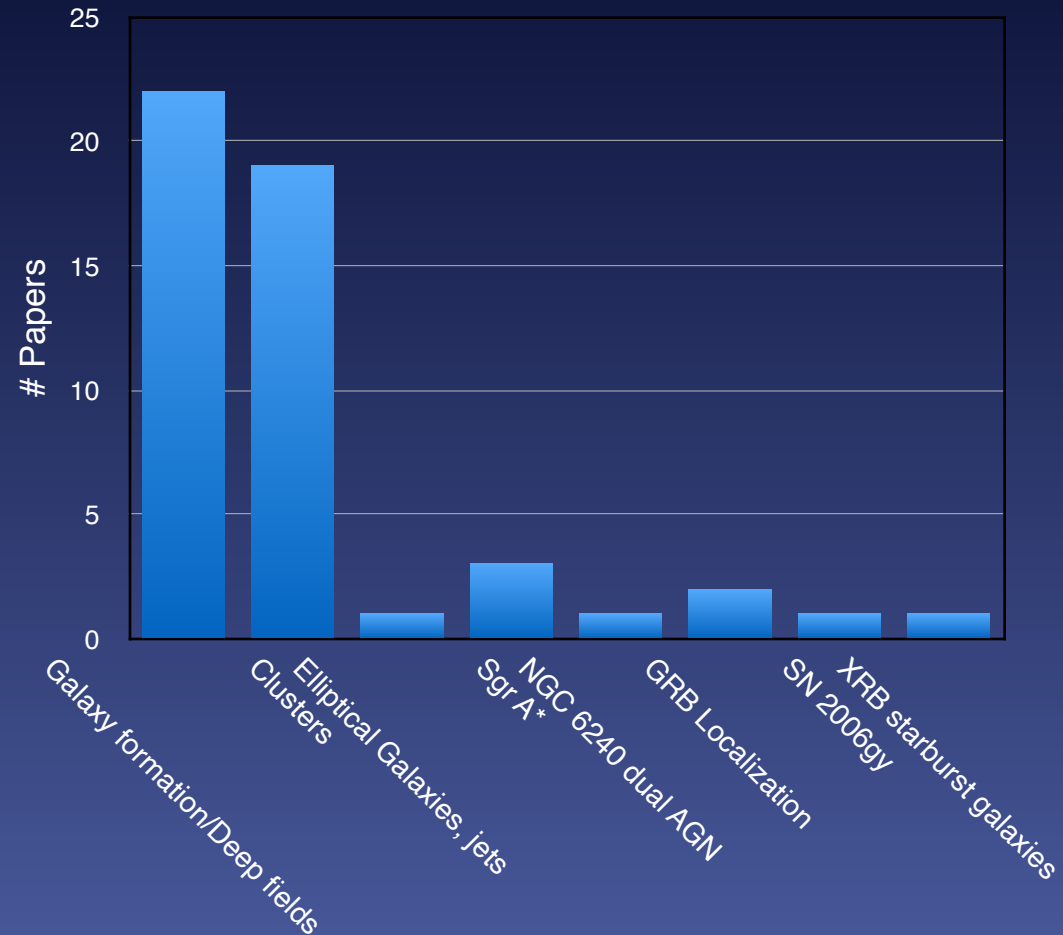


Publication rate per aggregate as a function of bin size.
Major caveat: most “Type 2” Chandra science papers have not been tied back to the original data sets.....*THESE NUMBERS MIGHT BE DIFFERENT NOW.*



Effectiveness of LPs and XVPs: highest impact papers

- Top 50 cited CSPs (no weighting!)
- spans years 2000-2010
- top paper has 803 citations, published in 2006
- 22 of the top 50 are galaxy formation/deep fields, with roots in one of the deep surveys
- 19 are cluster related science, 3 of which are cosmology papers from the 400d survey (others include shocks, cold fronts, cooling flows and the Bullet Cluster)
- Conclude: 50% of top 50 papers associated with XVP-sized aggregates!



Summary

- XVP-proxies earlier in the mission contribute a relatively small fraction of the absolute numbers of Chandra publications
- XVP-proxies have much higher publication rates than smaller aggregates: they produce a steady stream of publications. Actual numbers of papers definitely underestimated since not all CSPs Type 2 have been linked back to the original dataset.
- XVP-proxies feature prominently in the Top 50 CSP list.

