Revised seeing databases for `opsim`

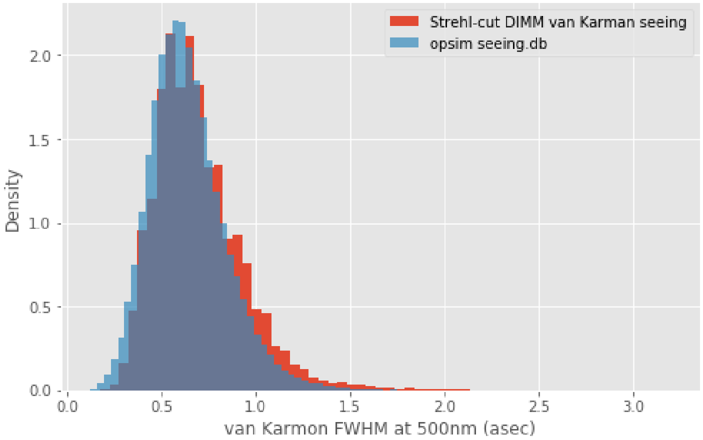
Eric H. Neilsen, Jr.

November 2019

This work was inspired by looking at the seeing distribution of LSST survey simulations made by `opsim`, and noticing that the seeing in the simulated surveys did not show the seasonal variation I expected based on experience with observing with DECam on the Blanco telescope, just a few miles away. In 2018, I obtained a data-set from the Cerro Pachon (Gemini South) DIMM, wrote an short python application, `simsee`, that could use it to generated a seeing database for `opsim` using DIMM data offset in time (and transformed from a Kolmogorov-based FWHM to one based on a von Kármán turbulence model), filling in gaps using an auto-regressive (AR) time-series model derived from that data. Reruns of `opsim` using this first seeing data set generated by `simsee` showed the seasonal seeing variation that I expected, and worse seeing overall.

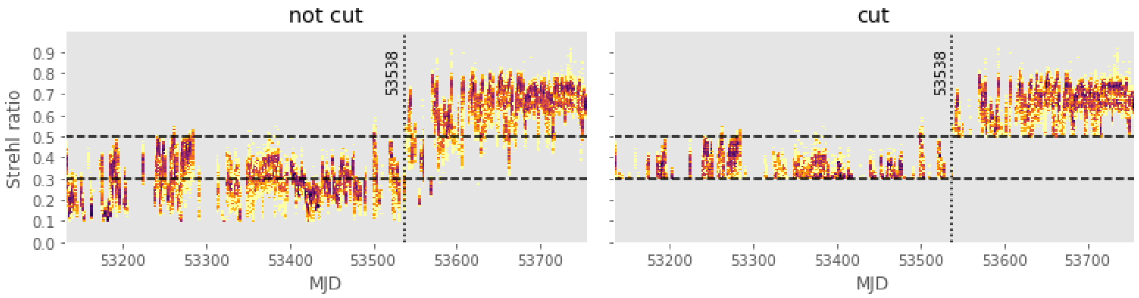
This still did not provide any understanding of the origin of the difference between my model and the `opsim` model. At the summer, 2019 LSST project and community workshop (PCW), Chuck Claver let me know that he had performed a cut on the Strehl ratio on the DIMM he used for the `opsim` model (to avoid using any data during which the DIMM had instrumental issues, such as being misaligned or out of focus), and provided me with both the cut data and the data cut by the Strehl ratio. The measurements in the uncut data set, running from 2004-05-06 to 2006-01-20, matched those from the Cerro Pachon DIMM data set for these dates: the `opsim` model was based on data from the same DIMM, but used a much shorter date range, and had been cut based on the Strehl ratio.

To confirm that the Strehl-cut data set provided was the one used as the basis for the `opsim` `seeing.db` database, I converted the FWHM values contained therein to von Kármán - derived FWHM values, and over-plotted histograms of the result an the contents of `seeing.db`:



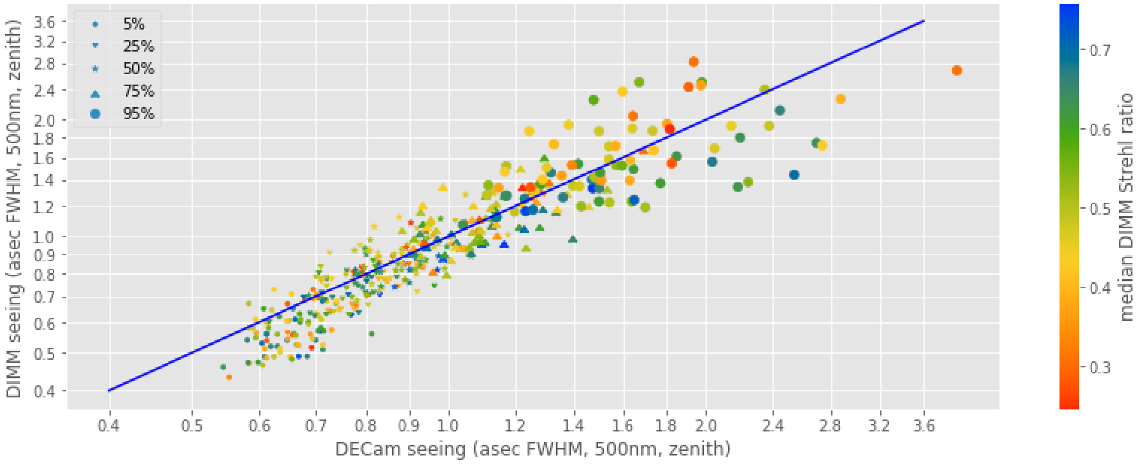
The histograms match well enough that it is plausible that the model used to derive the second (in blue) was derived from the first (in red).

2D histograms showing both the cut and uncut data sets made it clear what values of the Strehl ratio were used:



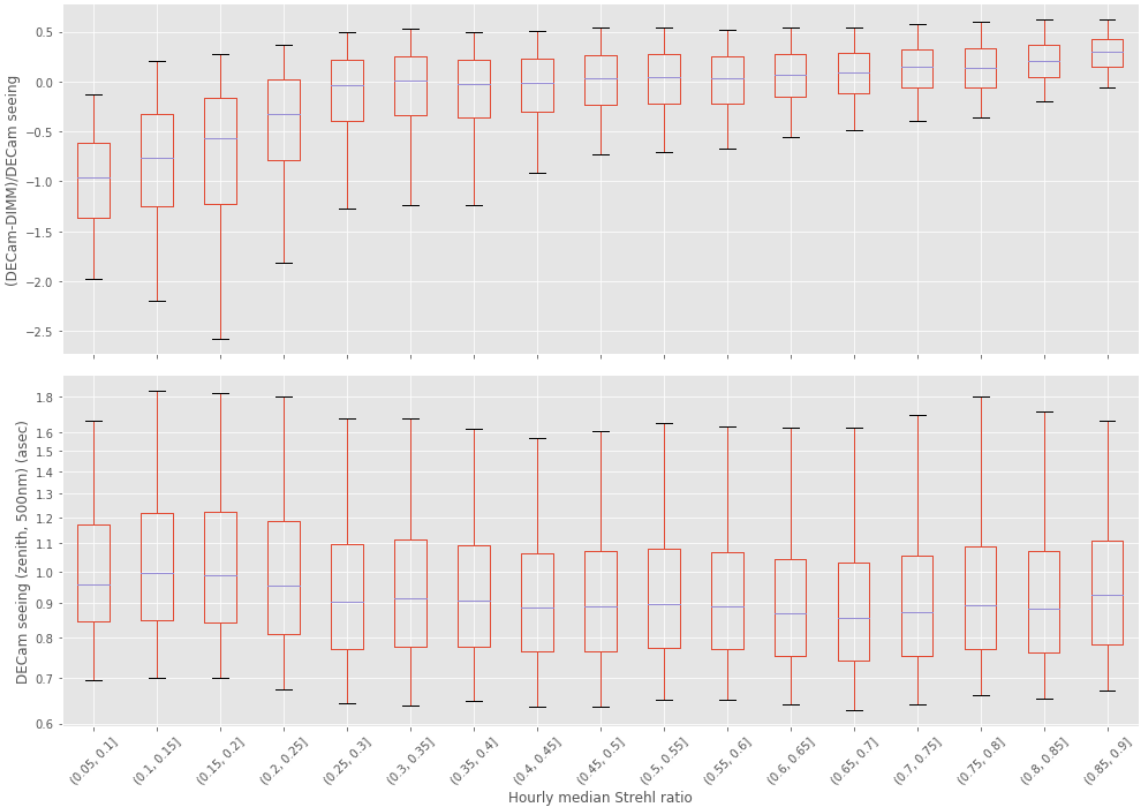
The “MJD” in the time axis above is the modified Julian date. It is clear that two different cut-off values for the Strehl ratio were used: 0.3 before MJD 53538 (2005-06-17), and 0.5 thereafter.

To understand what a good cutoff for the Strehl ratio would be, I then compared the DIMM data set to seeing measurements from the DECam camera, mounted on the Blanco telescope a few miles away. The overall distributions of DIMM data and DECam data (corrected to 500nm, zenith) were similar:



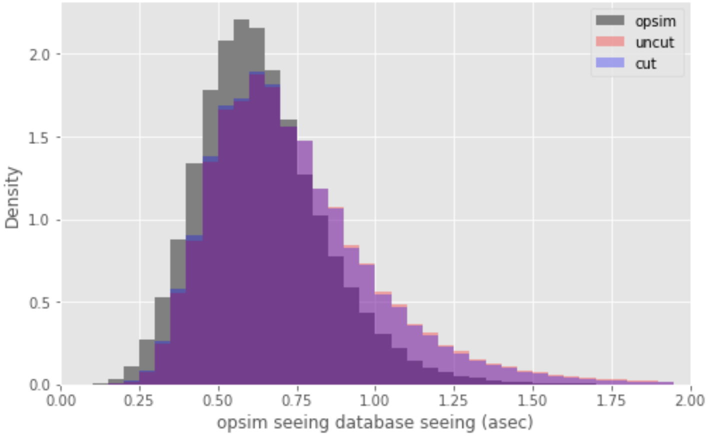
Monthly quantiles of the DECam corrected seeing match the Cerro Pachon DIMM quantiles for the corresponding months well, and there are no obvious deviations with the median Strehl ratios for the months. (The DECam seeing data “bottoms out” around 0.6", probably due to the DECam images becoming undersampled.)

Dividing the data into smaller (one hour) bins and examining the distributions in bins of Strehl ratios, two effects become apparent:



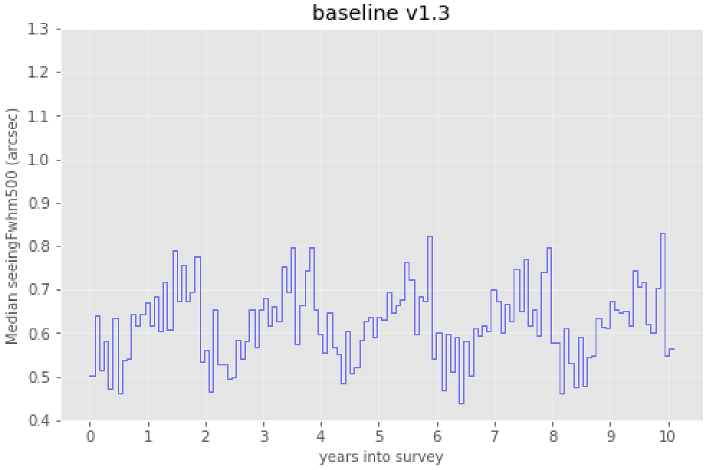
On hours during which the median DIMM Strehl ratio exceeds 0.25, the DIMM and DECam seeing are well matched, and the seeing distribution is independent of Strehl ratio. At lower Strehl ratios, however, the DIMM data does indeed overestimate the seeing, suggesting that a cut on the Strehl ratio of 0.25 is a good idea. However, the DECam seeing is worse at these times, so applying this cut will disproportionately remove times of genuinely poor seeing, so the application of this cut will bias the data set in the opposite direction.

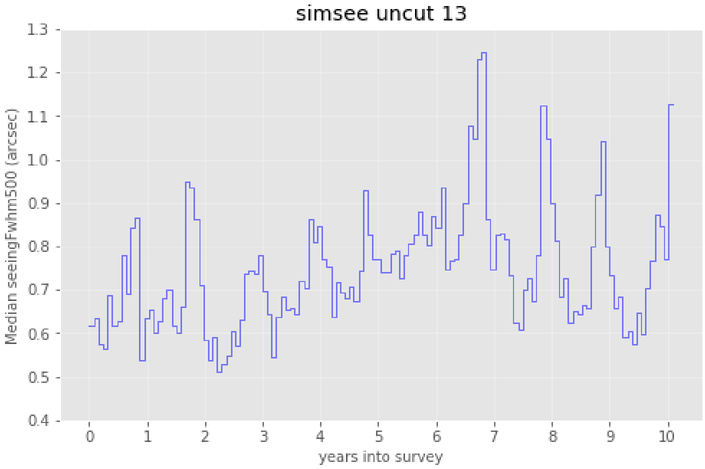
Therefore, I generated models and `opsim` seeing databases from both cut and uncut data using a utility call `simsee`. Fortunately, in the long baseline of DIMM data used, the fraction of data failing the cut is small, so the actual differences between the two models is minor. Both are significantly worse the original `opsim` model:

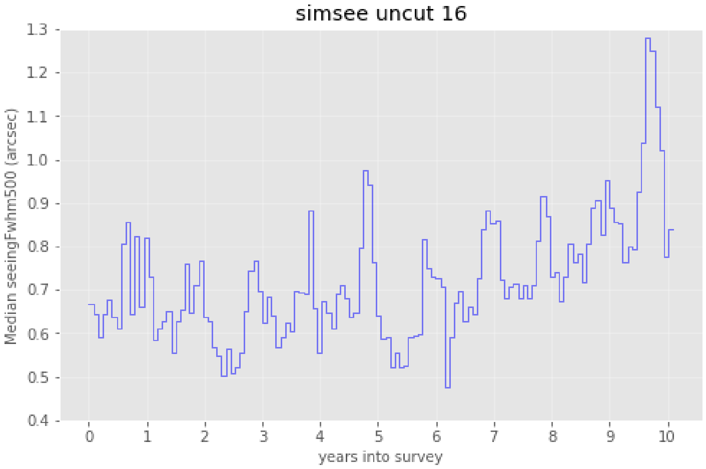


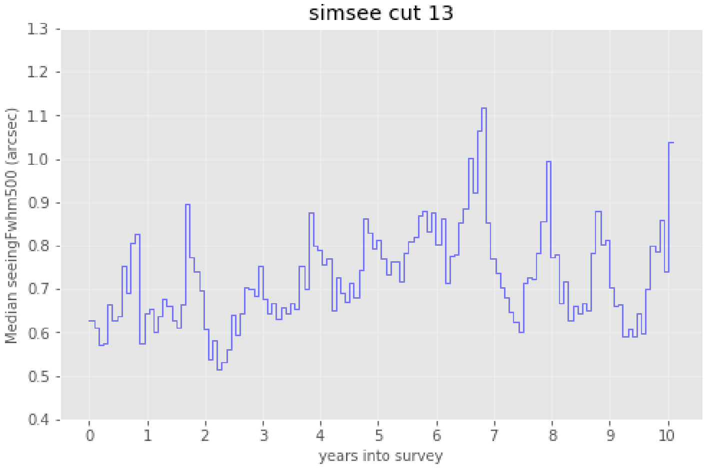
The `simsee` cut and uncut database are close enough that they almost completely overlap in the above histograms.

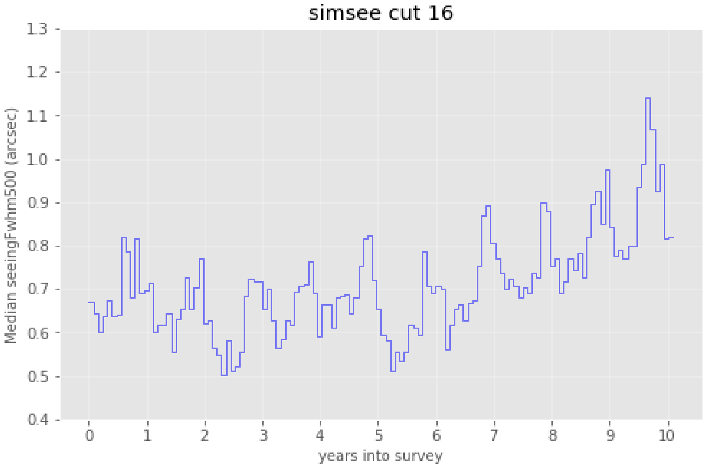
The differences between models are ever more pronounced when plotted with time:


baseline v1.3


simsee uncut 13


simsee uncut 16

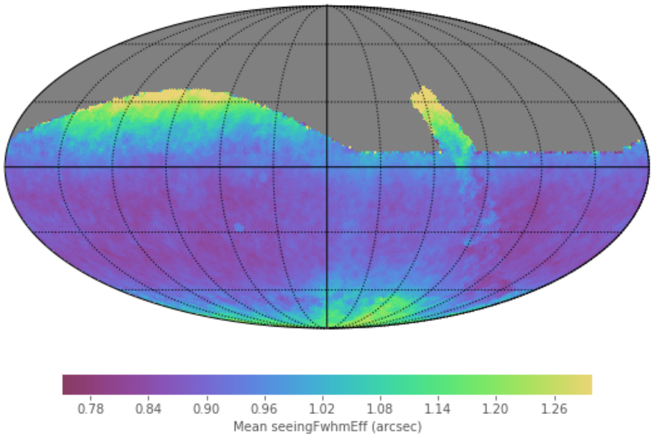

simsee cut 13


simsee cut 16

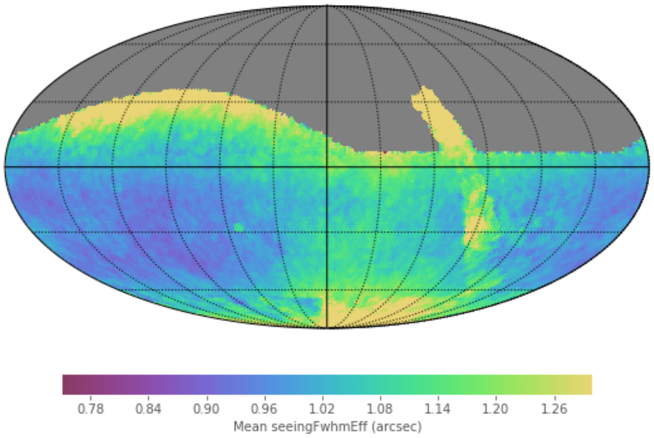
Each `simsee` seeing database is created by applying an offset to the DIMM data (after correction from Kolmogorov to von Kármán seeing (following Tokovinin (2002)), with missing windows of data filled in with artificial data generated by a model derived from the data. The "13" simulations apply and offset of 13 years between DIMM data and simulated LSST data, such that LSST years 2022 through 2032 are filled in with DIMM data from 2009 through 2019, and the "16" simulations apply and offset of 16 years, such the LSST years 2022 through 2035 correspond to 2006 though 2019. Note that the "16" seeing database can therefore be used to simulate more than just the nominal 10 LSST years of data collection, but the "13" simulations cannot.

I then ran `opsim` v1.3 baseline simulations using both the default seeing database and each of the ones shown above. The *i* band seeing maps (on a common color scale) look like this:

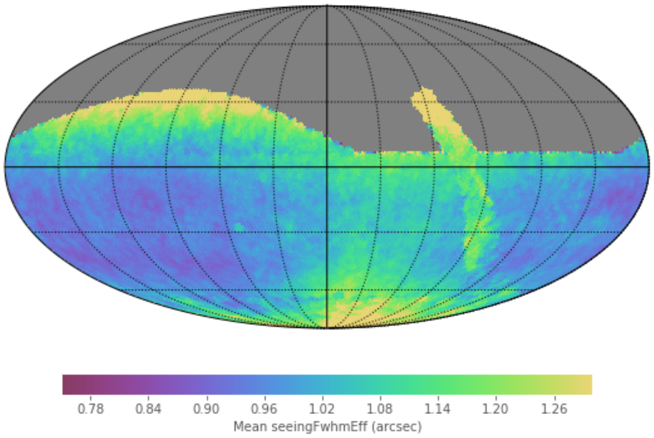
baseline v1.3



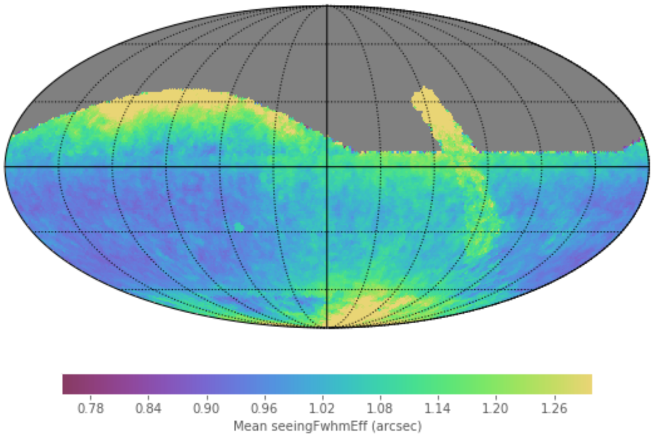
simsee uncut 13



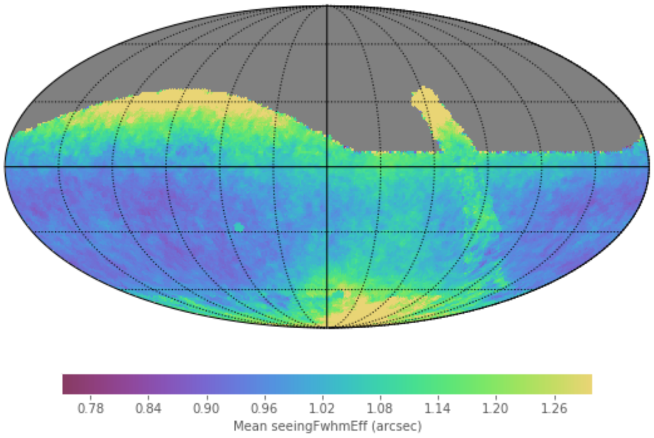
simsee uncut 16



simsee cut 13

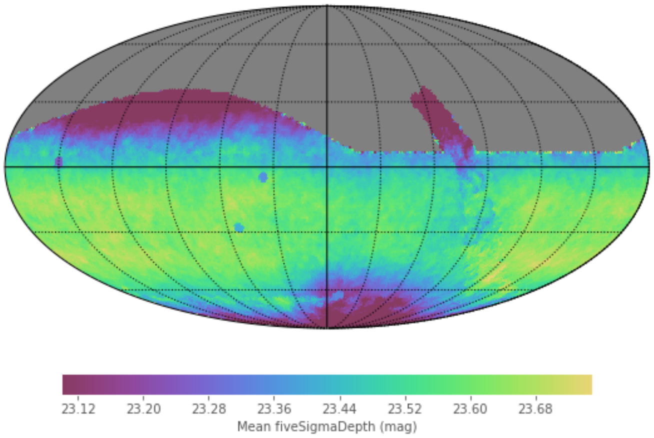


simsee cut 16

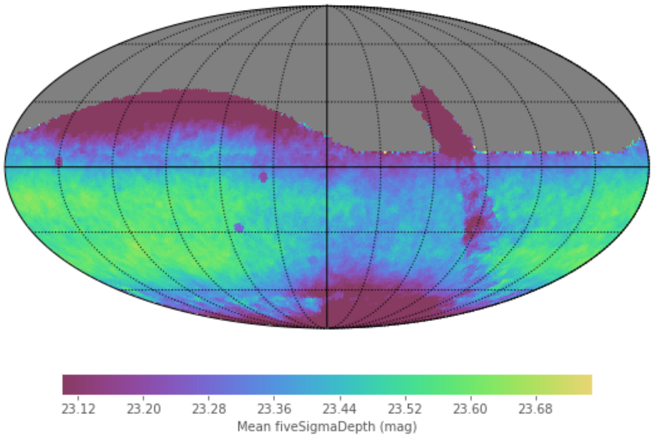


The depth maps have corresponding variation. In *i* band:

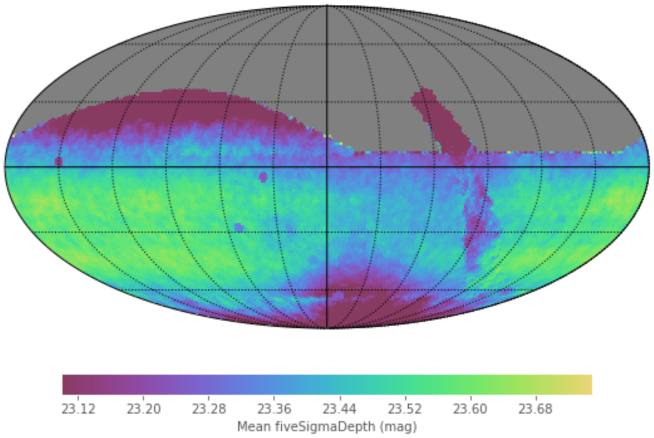
baseline v1.3



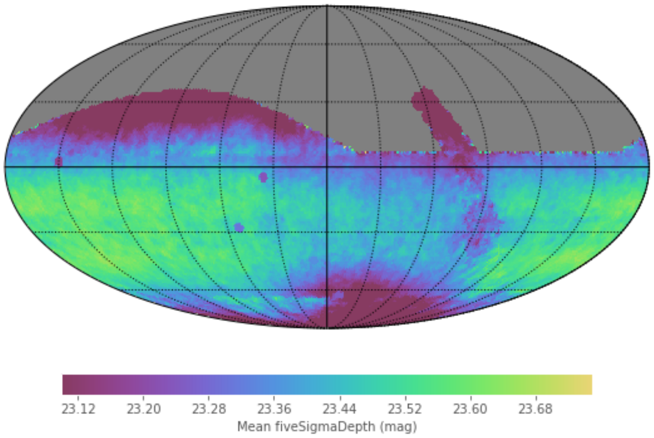
simsee uncut 13



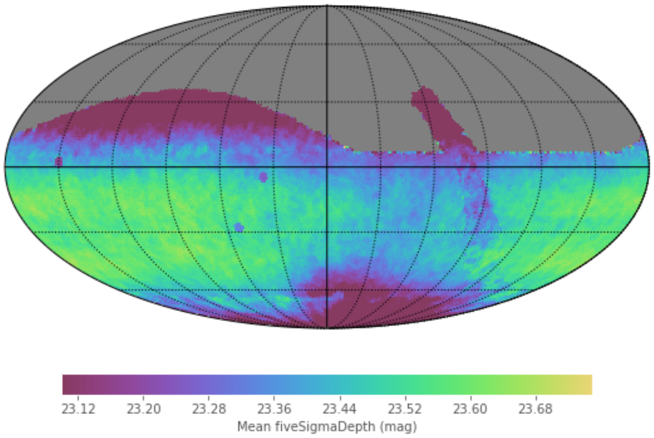
simsee uncut 16



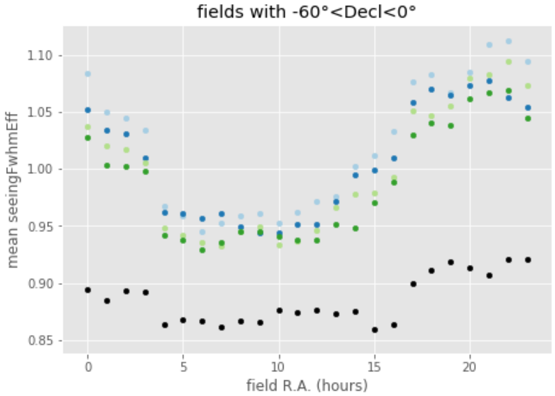
simsee cut 13

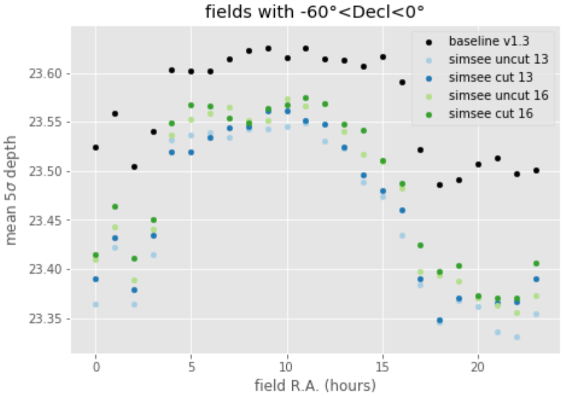


simsee cut 16



The new seeing database introduce a much more pronounced variation in depth with R.A. Plotting only area in the central (to the footprint) declinations:





If we look at the mean and inter-quartile ranges (IQRs) by hour of R.A. and band, again looking just at the footprint with central declination ($-60° < \delta < 0°$), corresponding to the above plot, the mean shifts and the IQR is larger in all bands:

	mean						IQR					
band	g	i	r	u	y	z	g	i	r	u	y	z
simulation												
baseline v1.3	0.97″	0.89″	0.91″	1.04″	0.86″	0.87″	0.03″	0.03″	0.04″	0.04″	0.06″	0.04″
simsee cut 13	1.11″	1.01″	1.05″	1.18″	0.98″	0.98″	0.14″	0.09″	0.12″	0.07″	0.12″	0.10″
simsee cut 16	1.10″	0.99″	1.03″	1.19″	0.97″	0.97″	0.13″	0.09″	0.11″	0.15″	0.10″	0.09″
simsee uncut 13	1.13″	1.02″	1.06″	1.22″	0.99″	1.00″	0.17″	0.12″	0.13″	0.15″	0.15″	0.13″
simsee uncut 16	1.11″	1.00″	1.04″	1.20″	0.97″	0.98″	0.15″	0.10″	0.13″	0.16″	0.10″	0.11″

Note that the FWHM changes by about 10% over the baseline in all bands, with all `simsee`databases, and the IQR increases by a factor of about three (or more).

If we look at all fields, and do not bin by hour of R.A. or impose a restriction on declination, we see similar behaviour, although the fractional change in IQR is less extreme:

	mean						IQR					
band	g	i	r	u	y	z	g	i	r	u	y	z
simulation												
baseline v1.3	1.00″	0.91″	0.94″	1.07″	0.87″	0.90″	0.31″	0.26″	0.28″	0.35″	0.23″	0.25″
simsee cut 13	1.14″	1.03″	1.07″	1.20″	0.99″	1.01″	0.39″	0.33″	0.35″	0.41″	0.32″	0.32″
simsee cut 16	1.13″	1.01″	1.05″	1.21″	0.98″	0.99″	0.39″	0.33″	0.36″	0.43″	0.31″	0.32″
simsee uncut 13	1.16″	1.05″	1.09″	1.24″	1.00″	1.03″	0.40″	0.34″	0.37″	0.44″	0.33″	0.34″
simsee uncut 16	1.13″	1.02″	1.06″	1.23″	0.98″	1.00″	0.41″	0.34″	0.37″	0.43″	0.31″	0.32″