

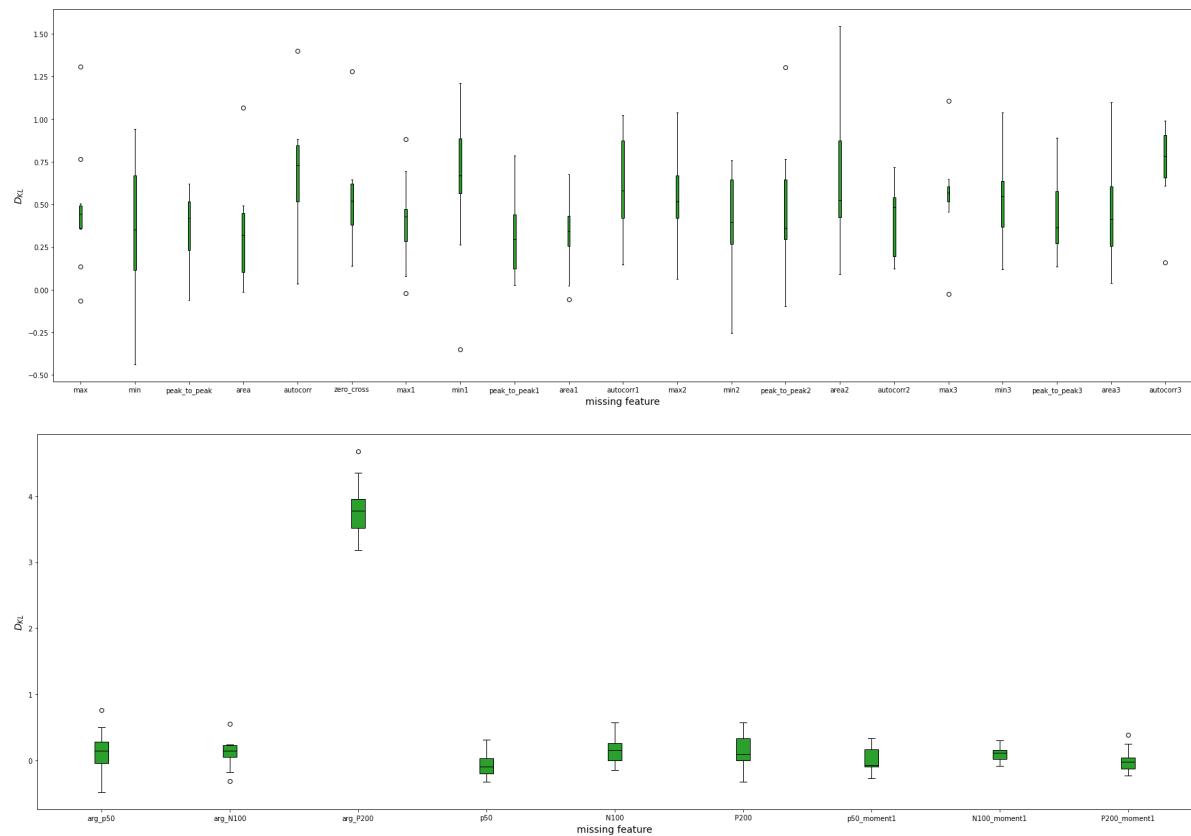
Week summary 9 Jan 22

Investigating Summary Features

KL-divergence between complete set of summary features and all but one missing (missing feature is indicated on x-axis):

$$D_{KL}(p(\Theta|x_1||p(\Theta|x)))$$

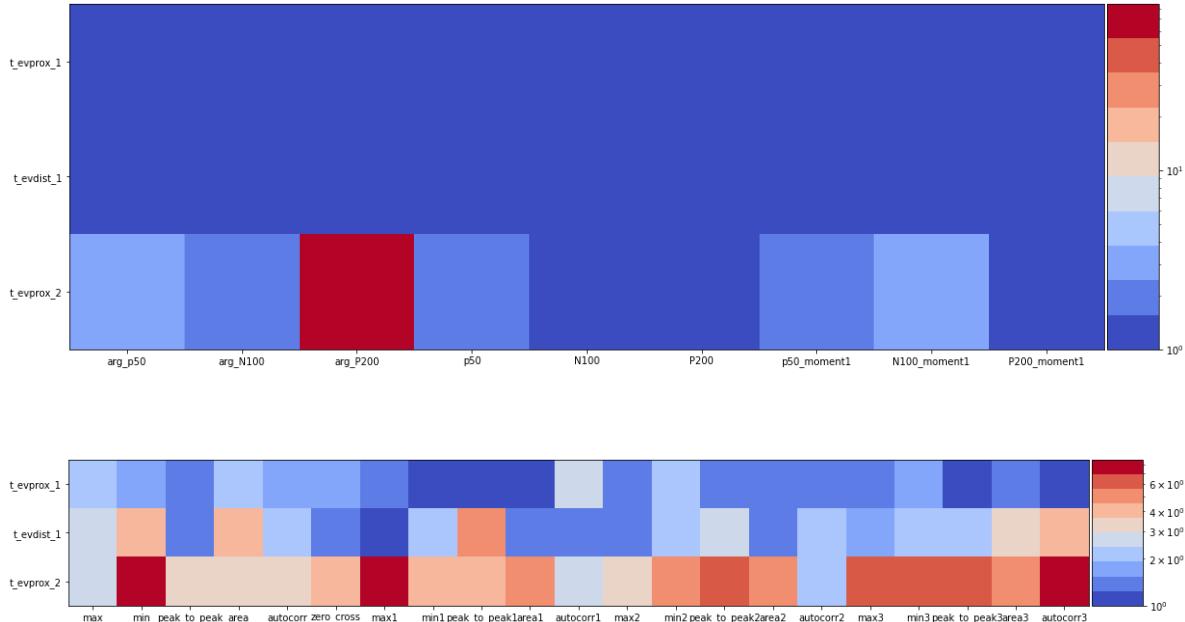
where $x = \{x_1, x_2\}$



Kritisch: - misst nur, ob das Feature 'überflüssig' ist - das könnte auch der Fall sein, wenn wie im obigen Fall viele Feature etwas 'ähnliches' messen...

Ratio of variances between the marginals of $p(\Theta|x_1)$ and $p(\Theta|x)$

- this is for the last posterior of the sequential approach such that evaluation has to be regarded with caution! That's also the reason why the first parameters do not have large changes in variance anymore! and it could be that some features were already very important for calculating the posterior for the first two parameters!



Any conclusions so far?

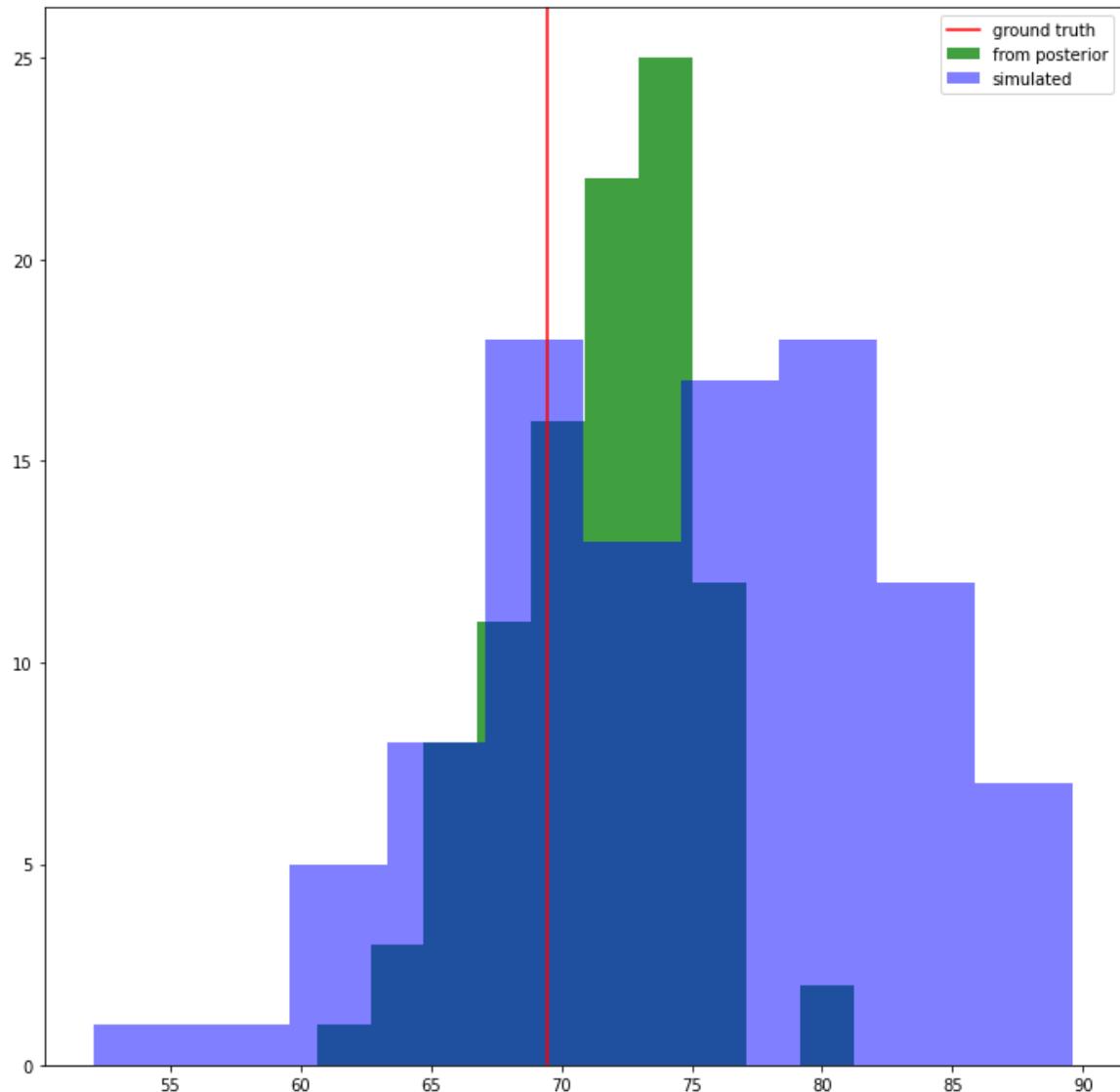
- to differentiate the value of summary features between parameters, we need to investigate without the sequential approach again!!!
- the overall max value and the first two autocorrelations seem to be less important than other features

Wäre es sinnvoll wie bei Milas Praktikumsarbeit feature ranges zu definieren, z.B:

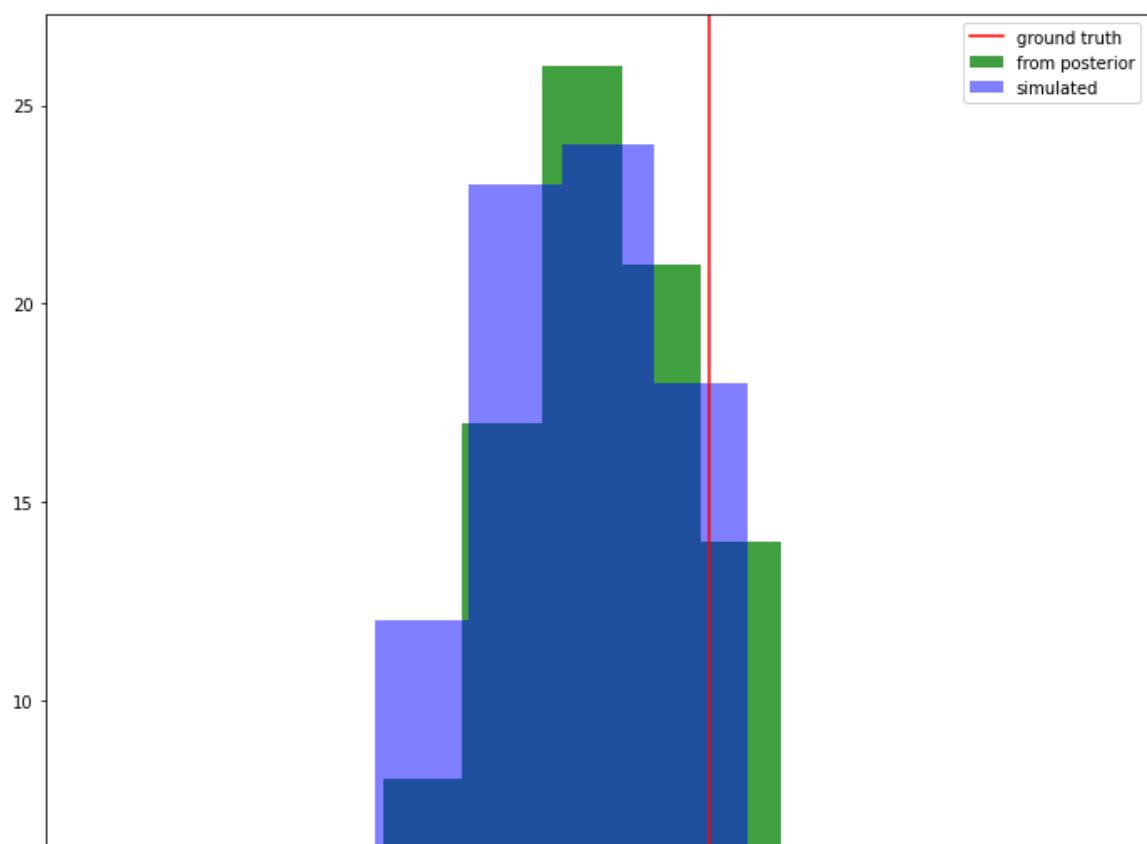
- number of zero crossings between 3 and 6
- positive peak value between 0 and 60
- (similar to what I did manually with the timing of the peaks..)

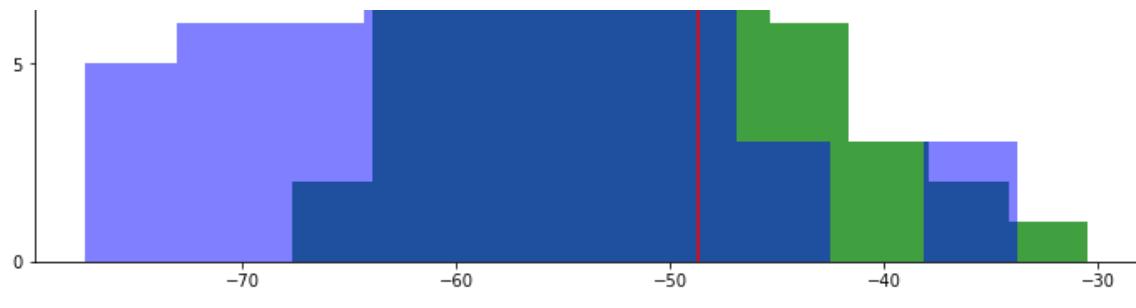
Looking at the histogram distribution of feature values (comparing prior and posterior values);

Histogram of summary stat "max"

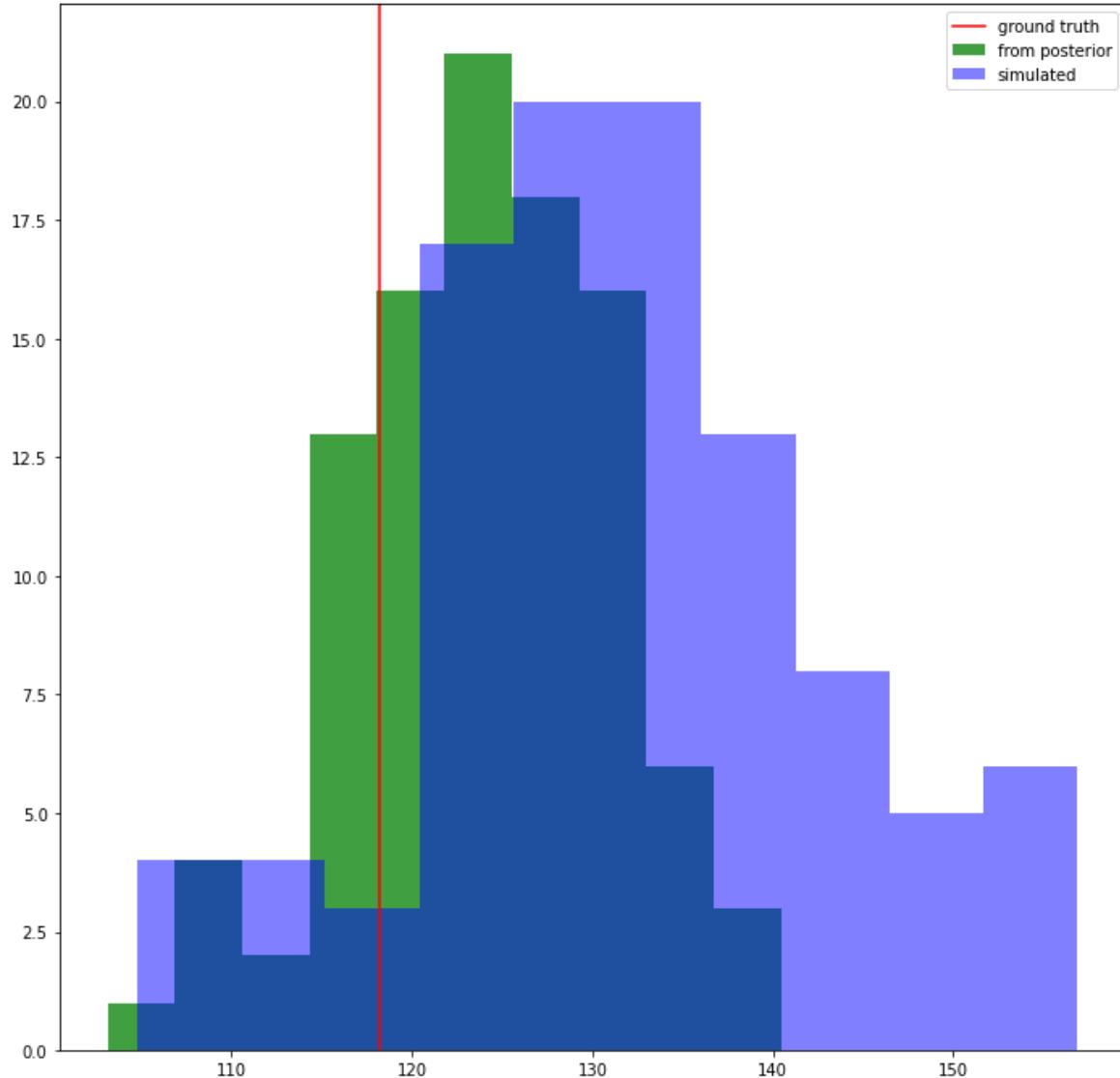


Histogram of summary stat "min"

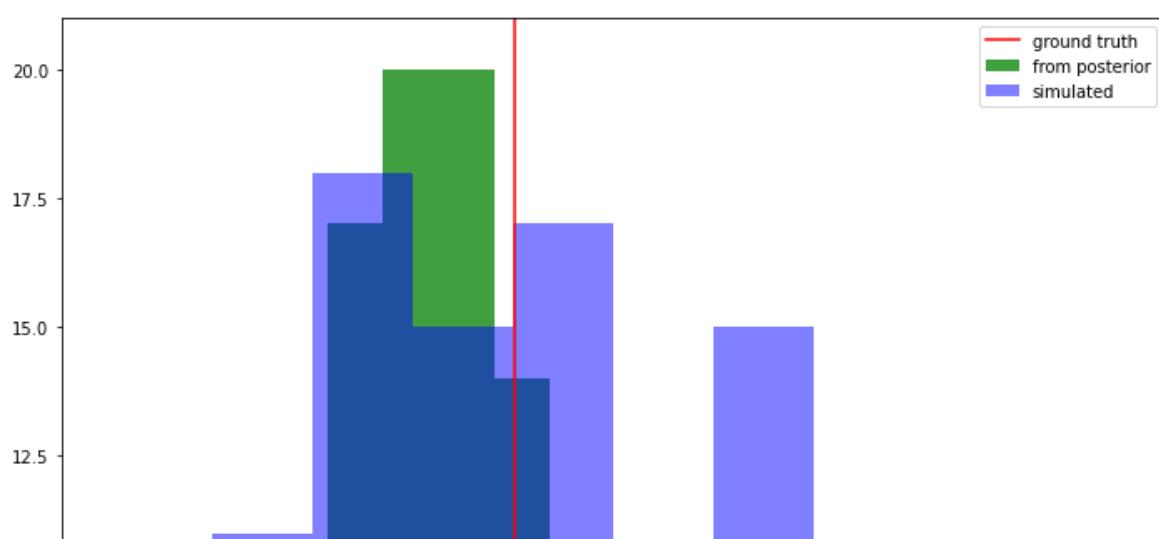


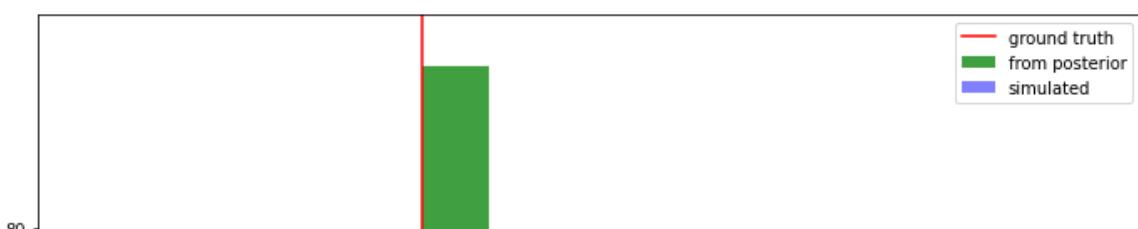
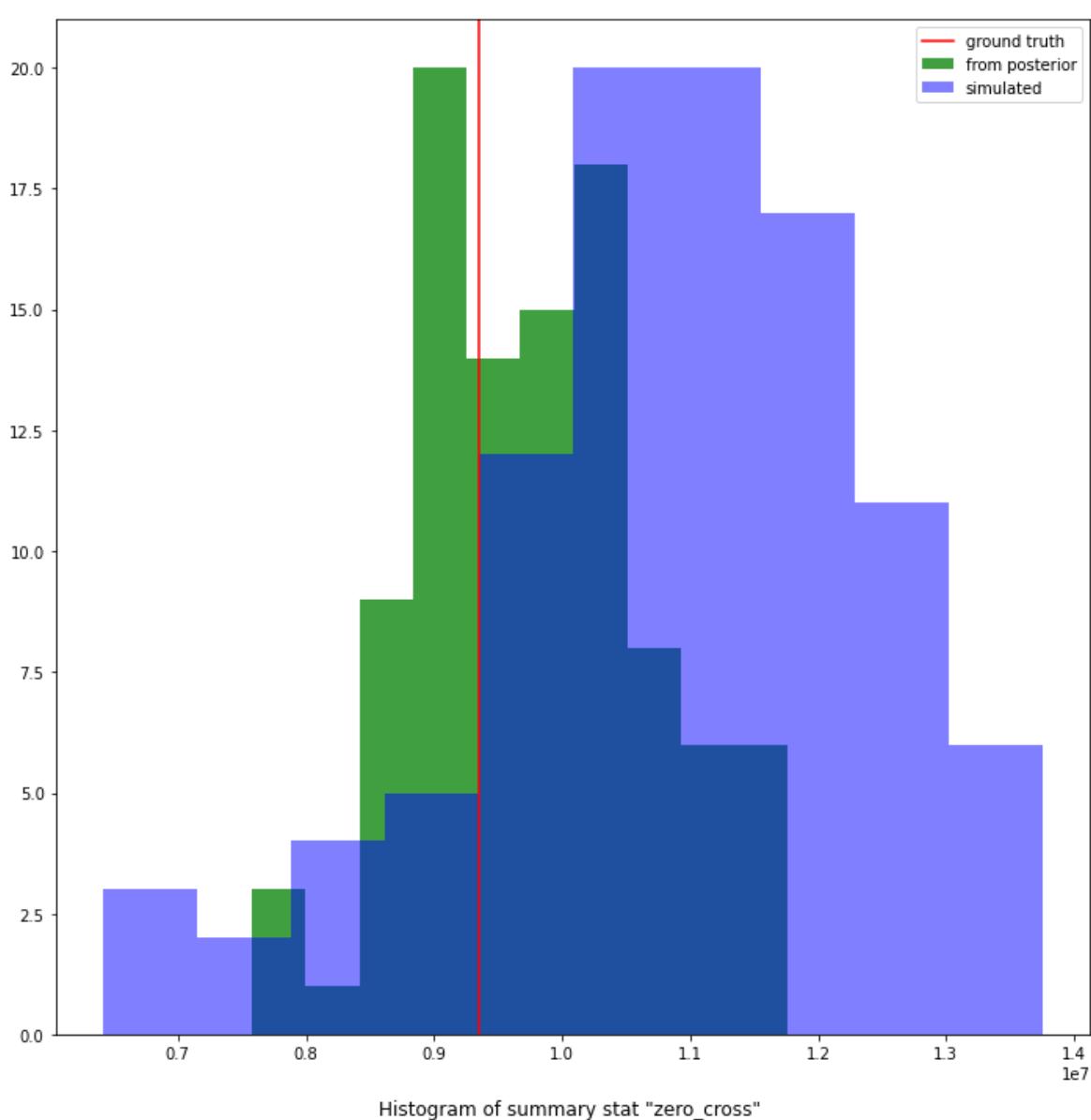
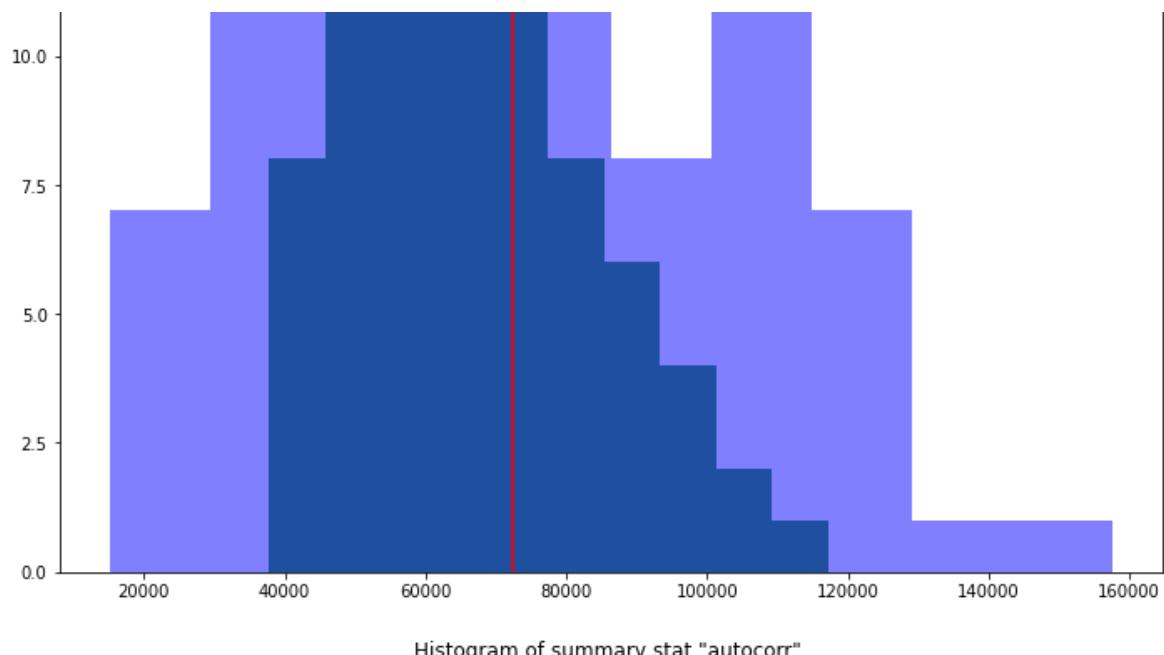


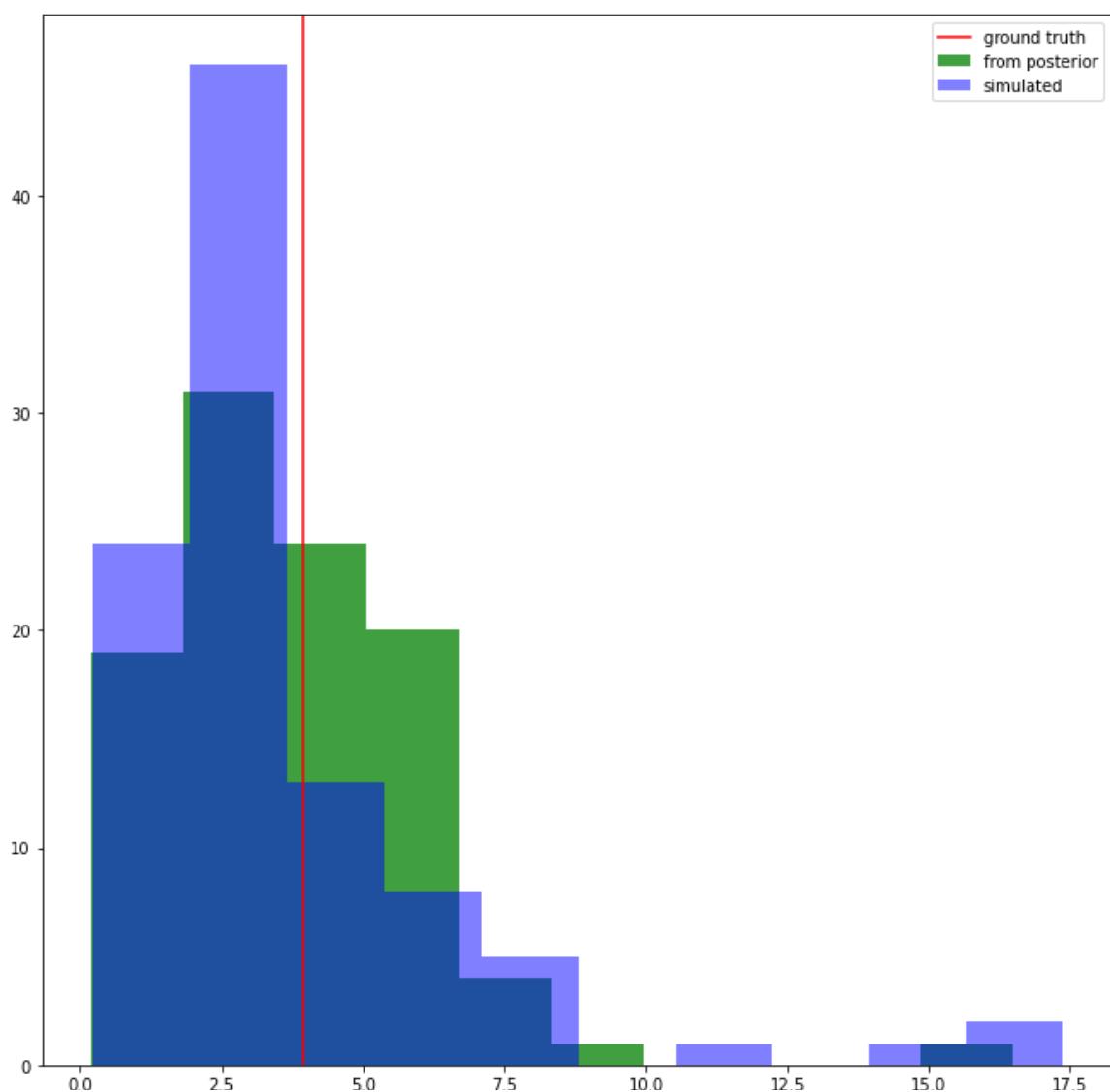
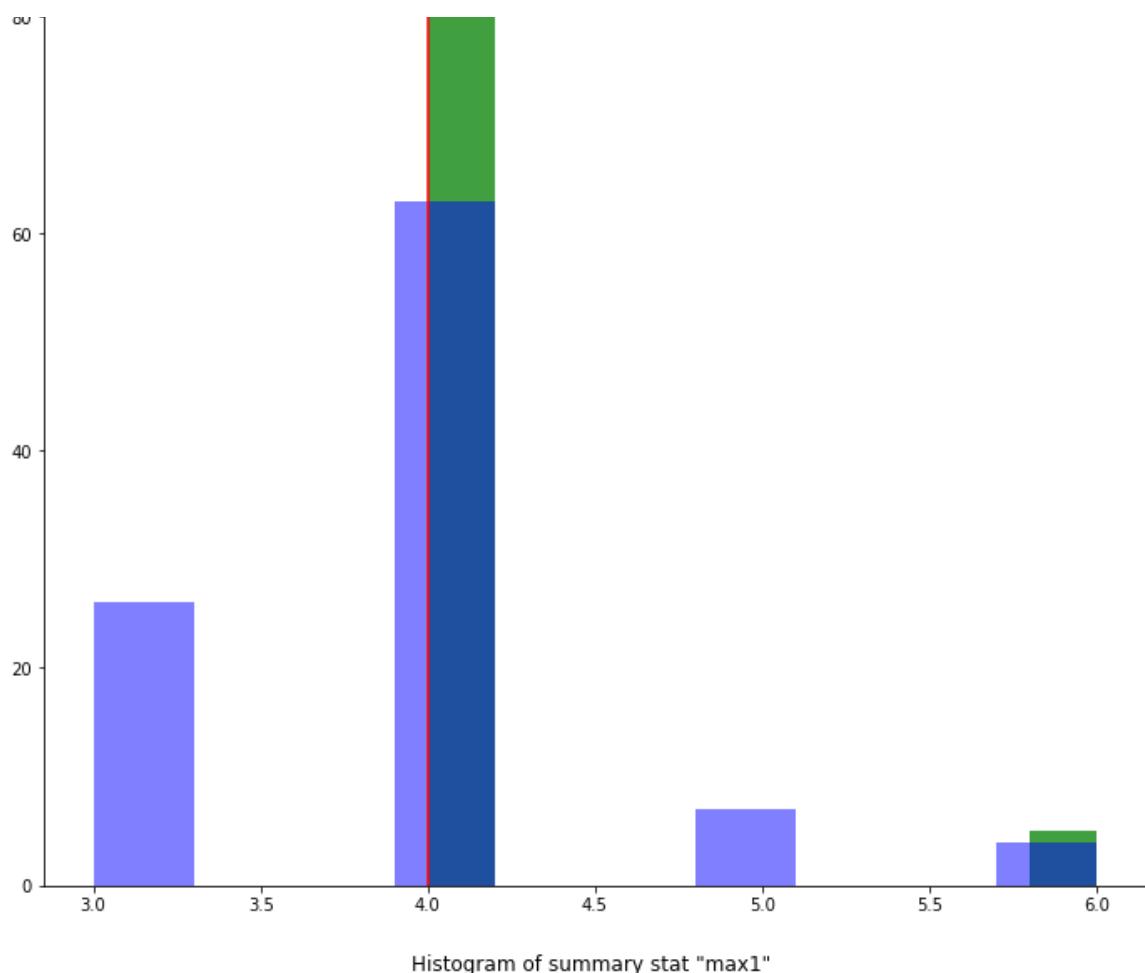
Histogram of summary stat "peak_to_peak"



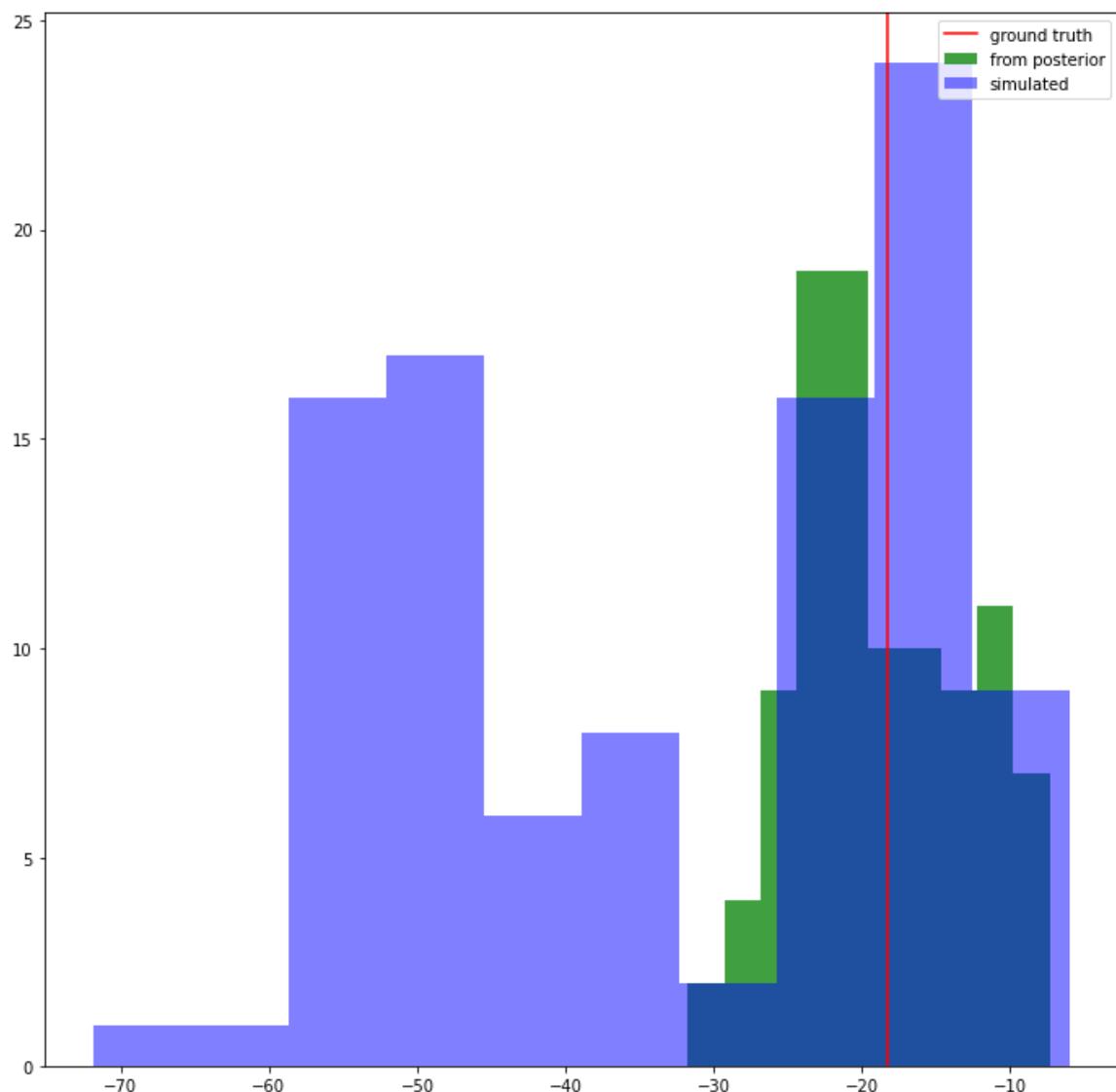
Histogram of summary stat "area"



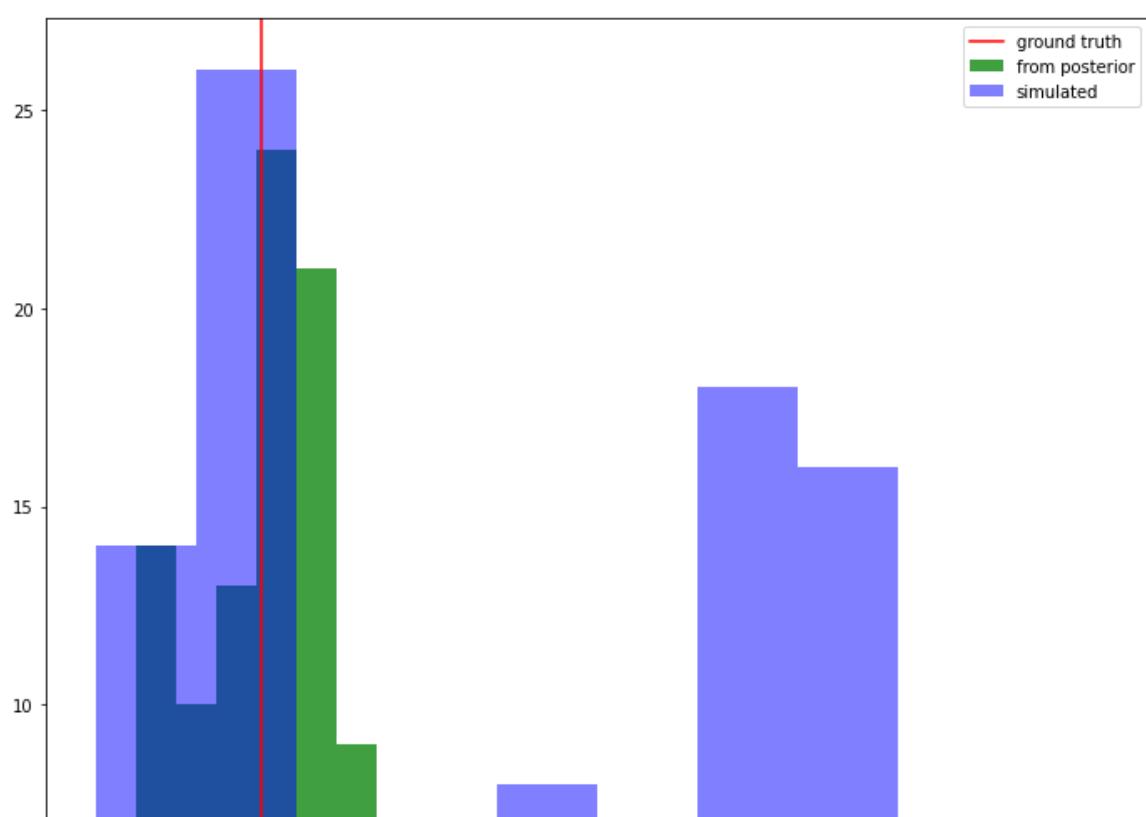


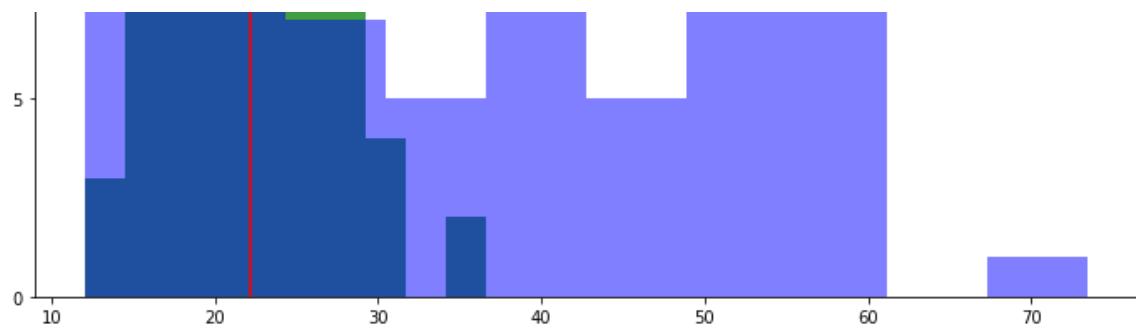


Histogram of summary stat "min1"

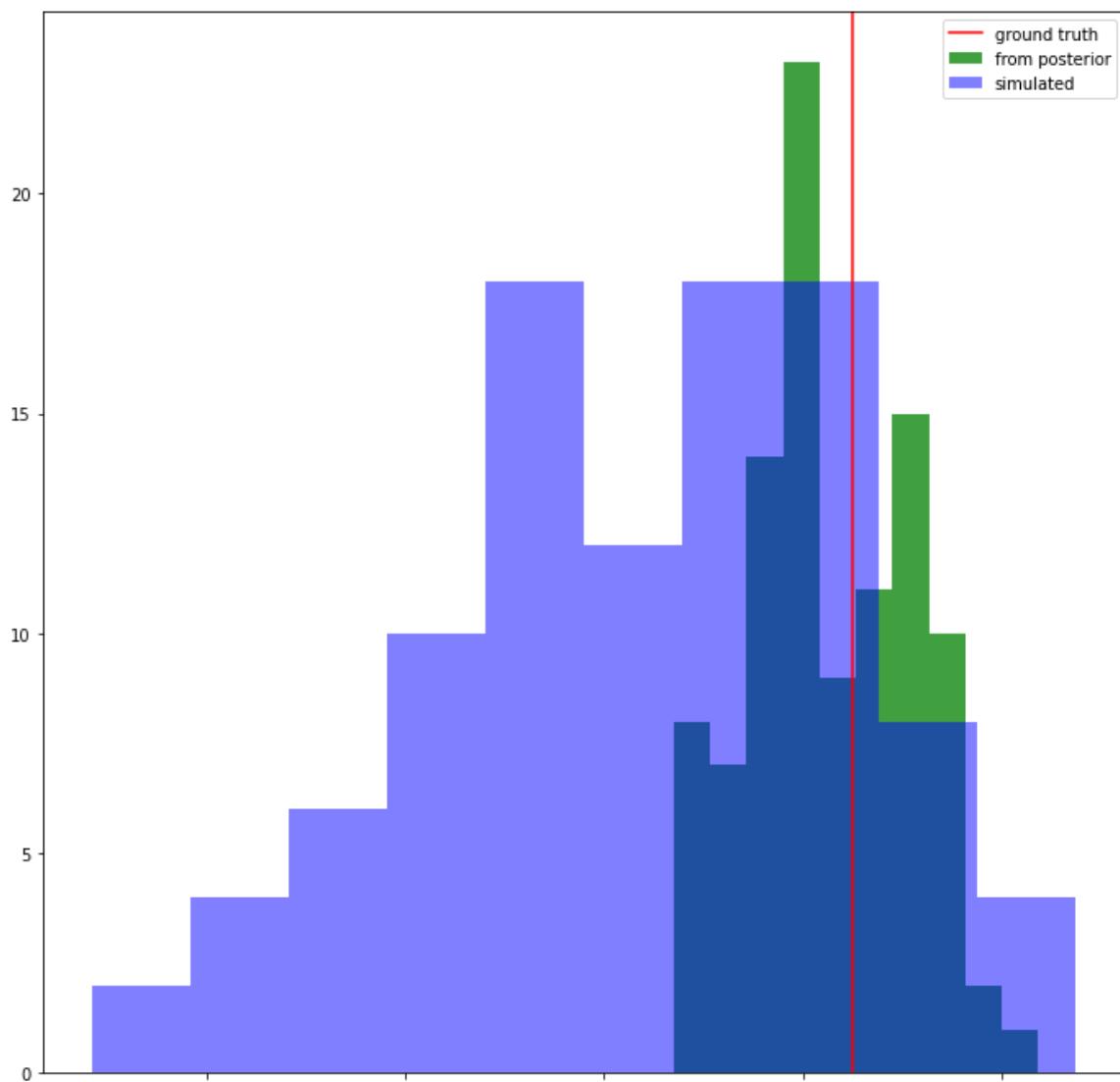


Histogram of summary stat "peak_to_peak1"

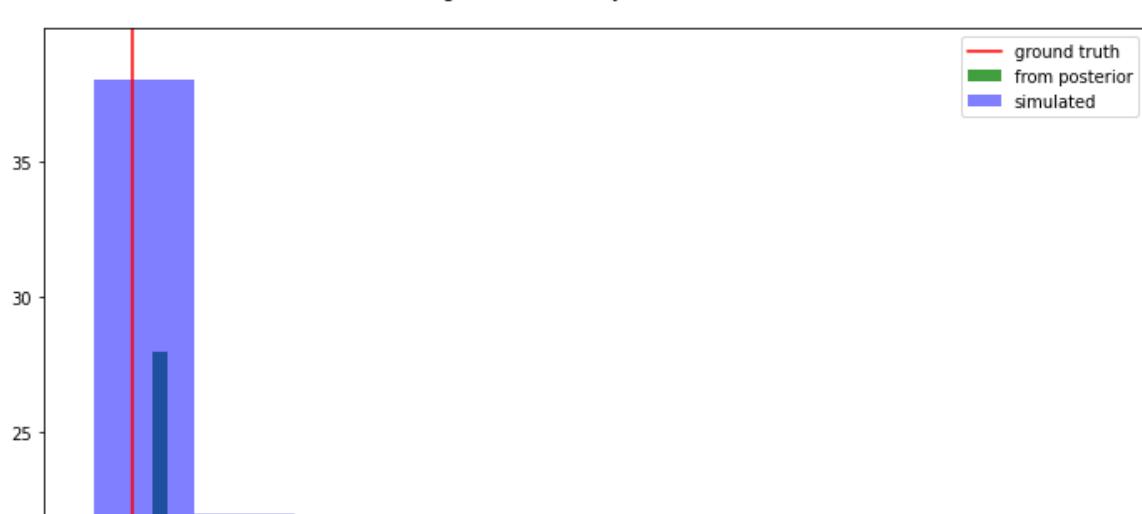


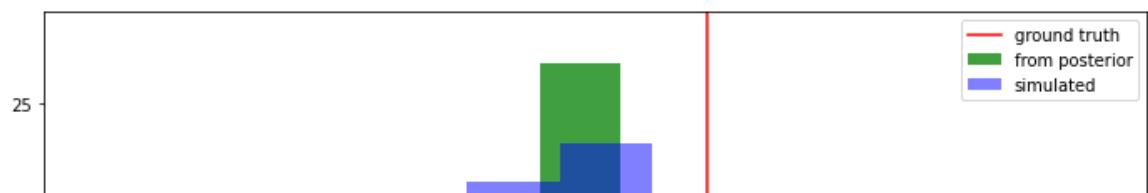
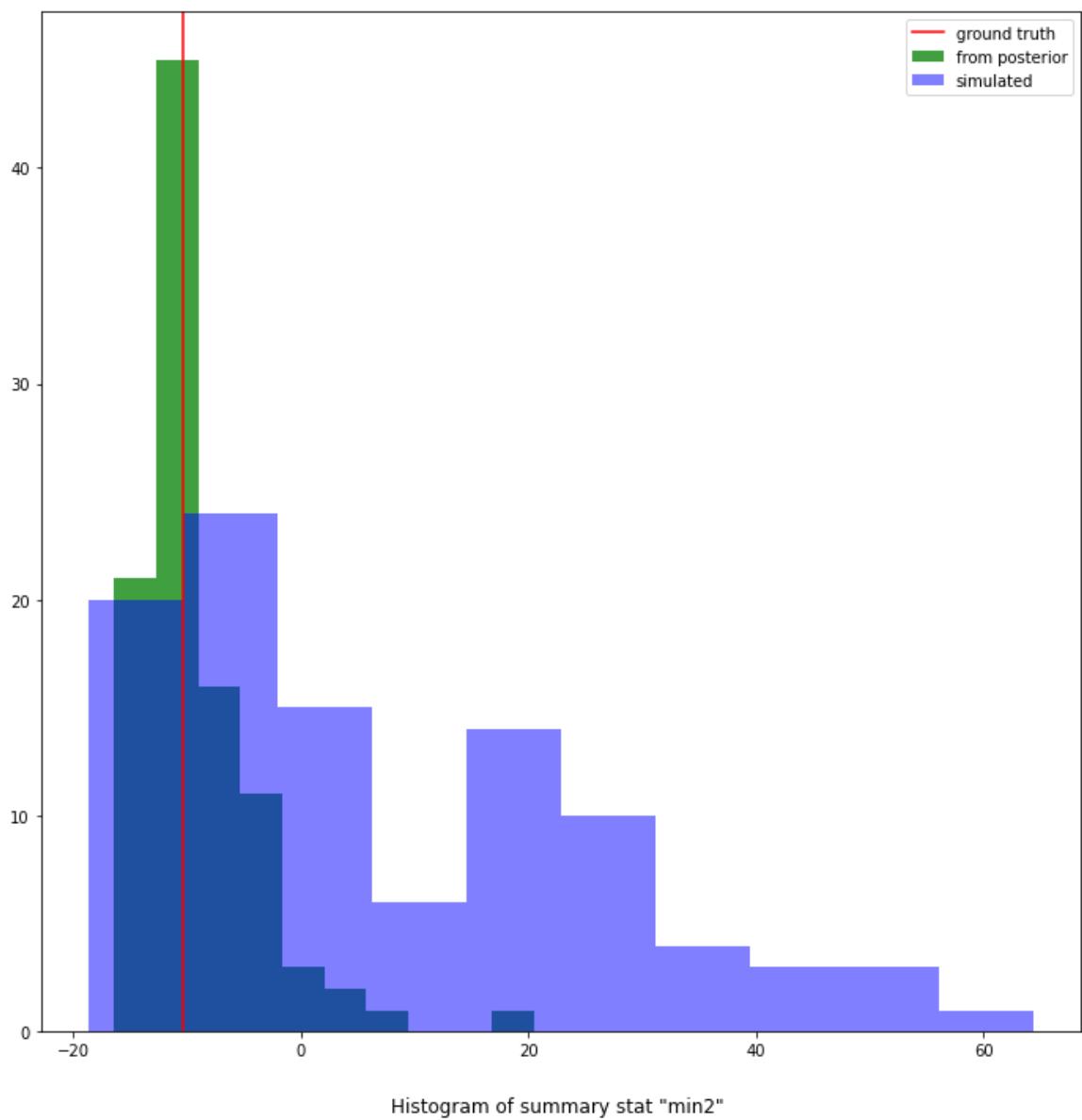
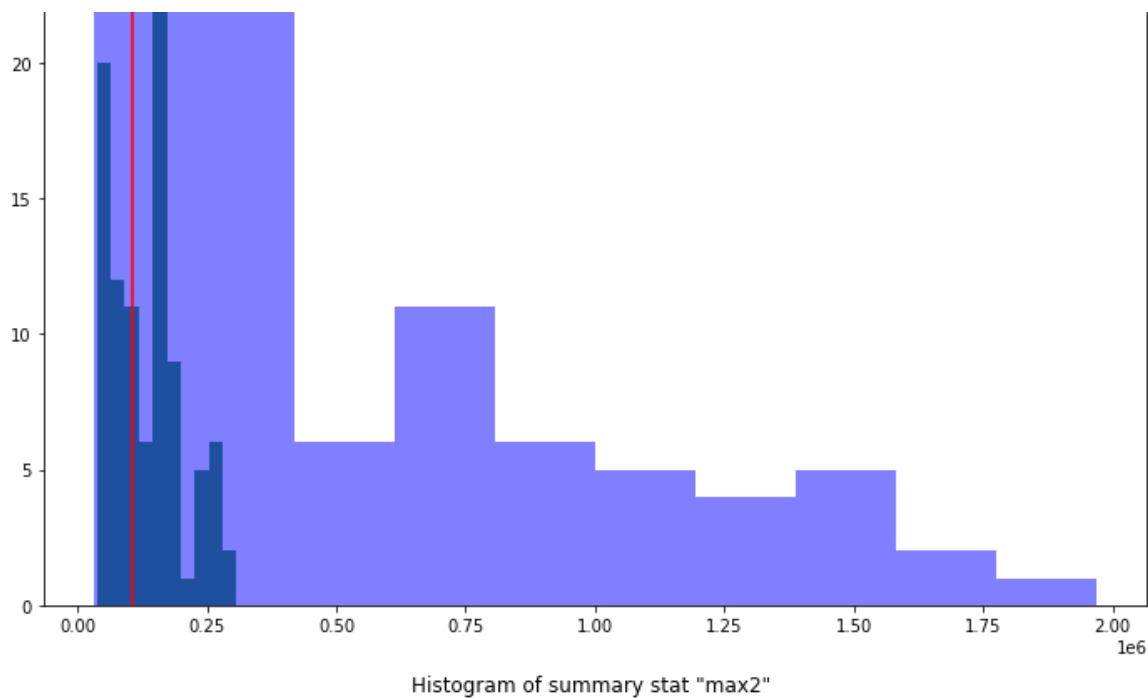


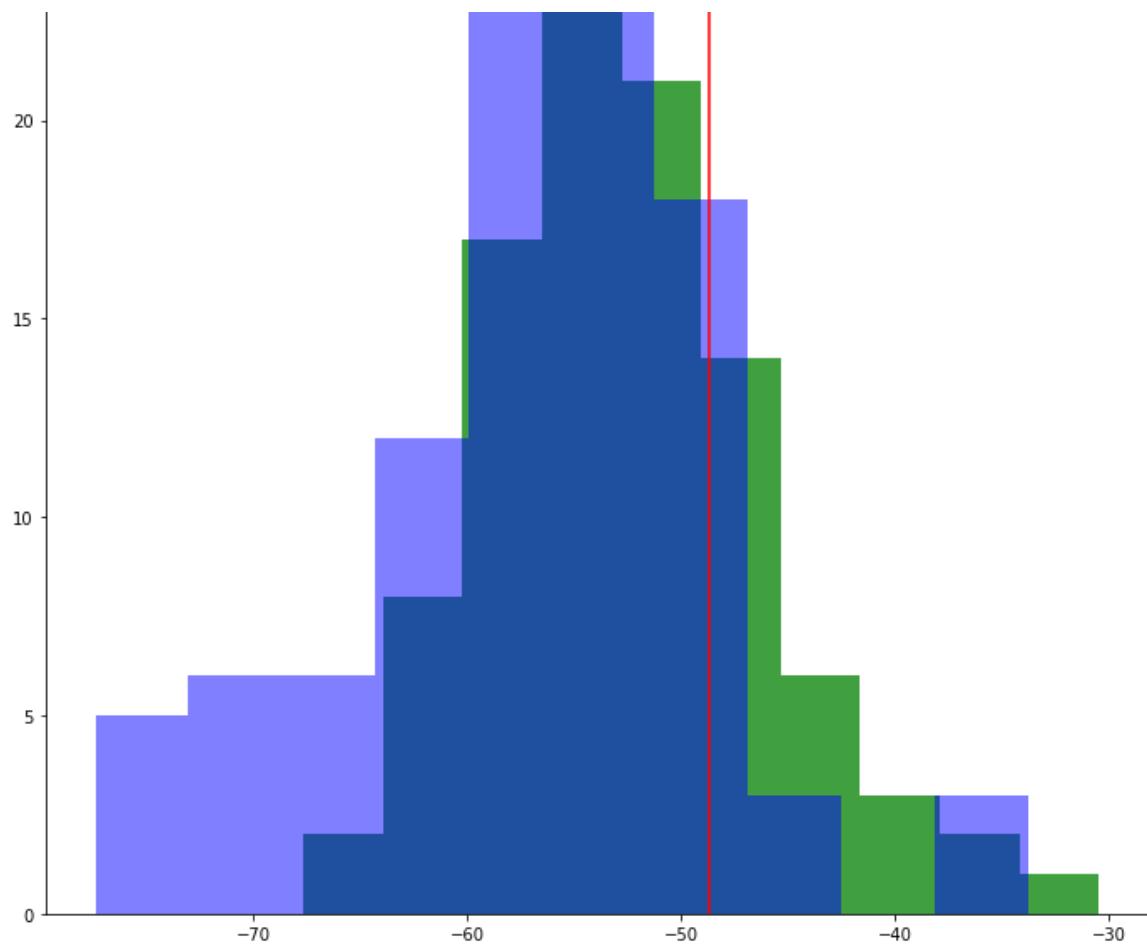
Histogram of summary stat "areal"



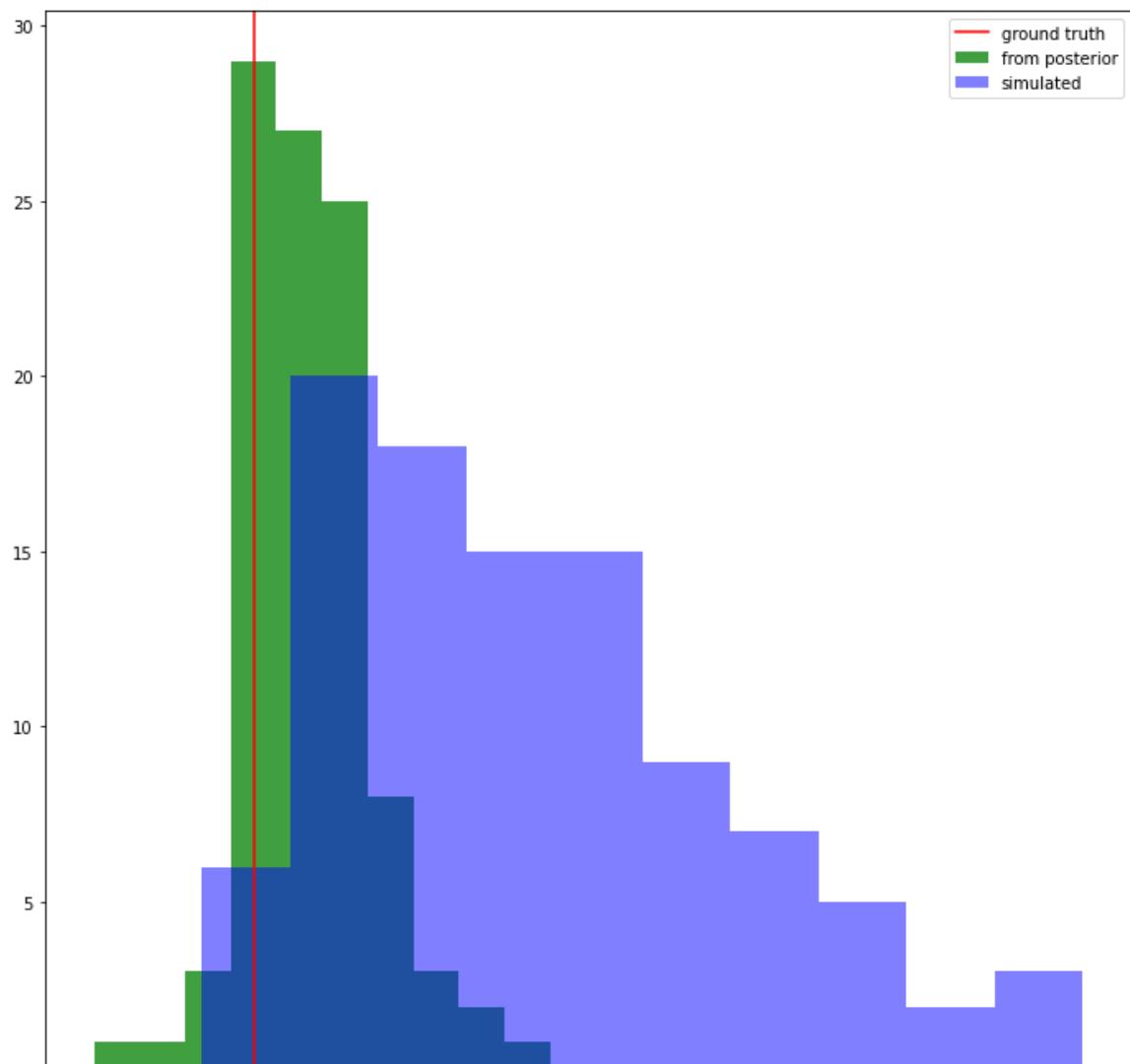
Histogram of summary stat "autocorr1"

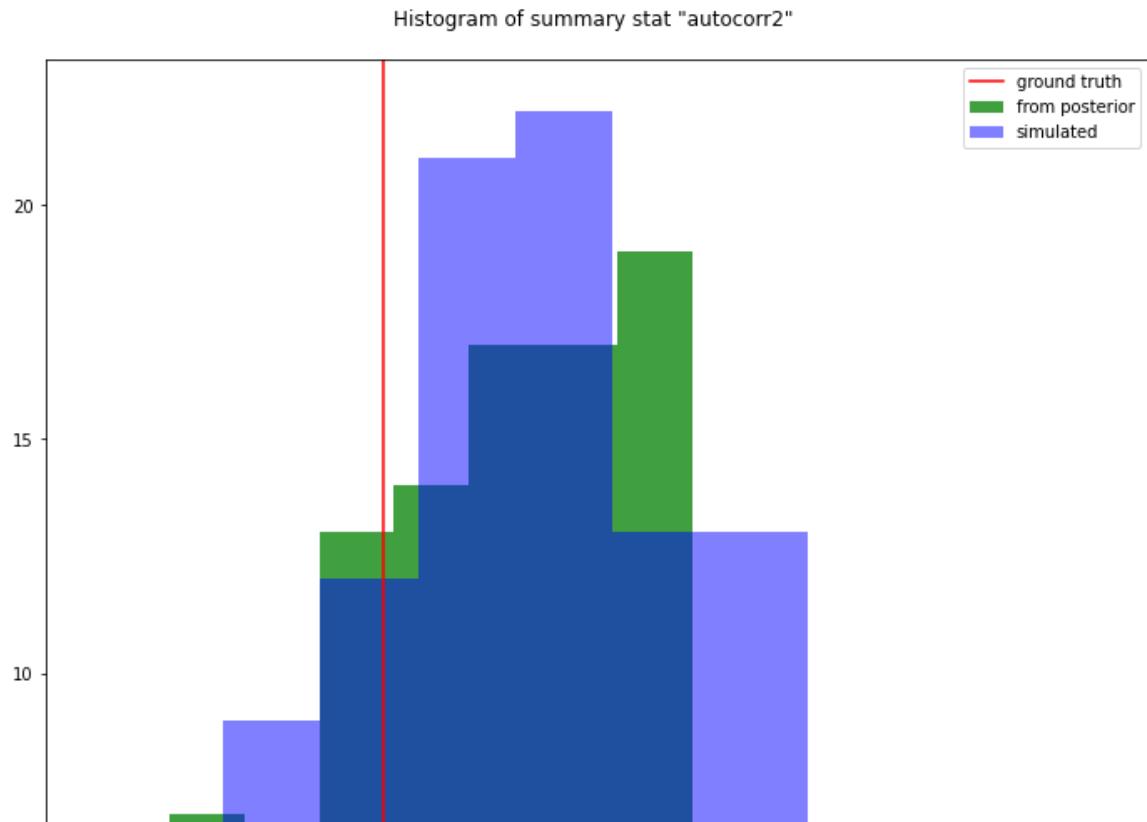
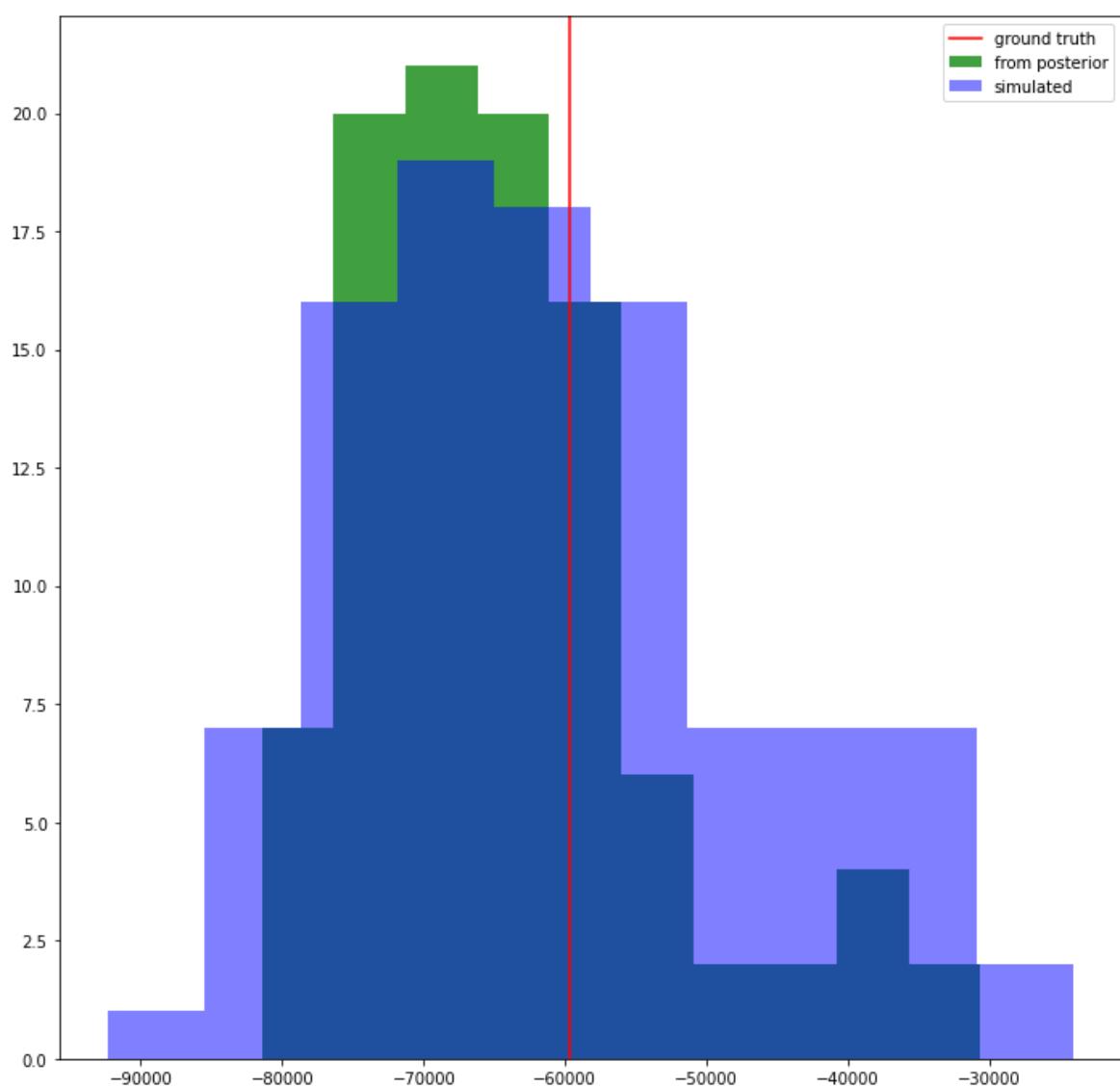


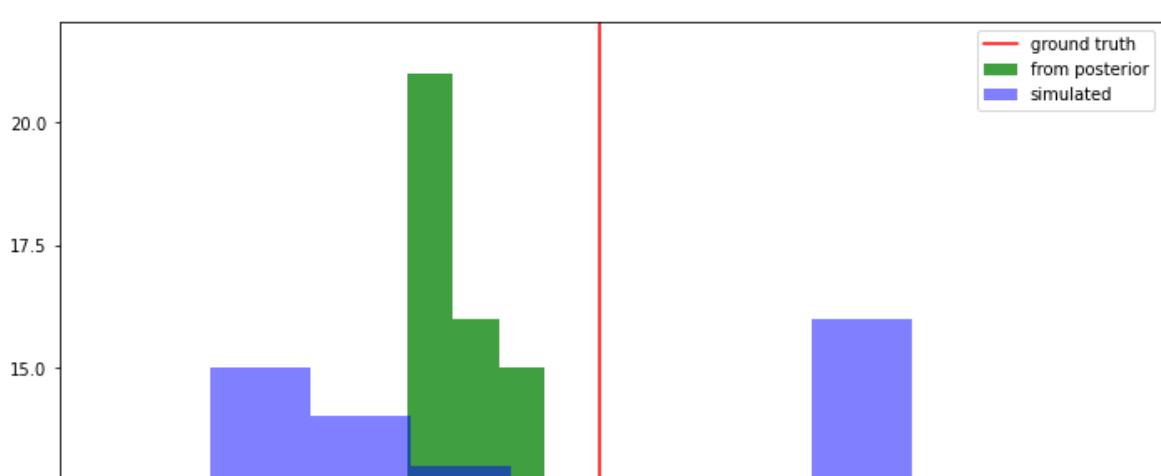
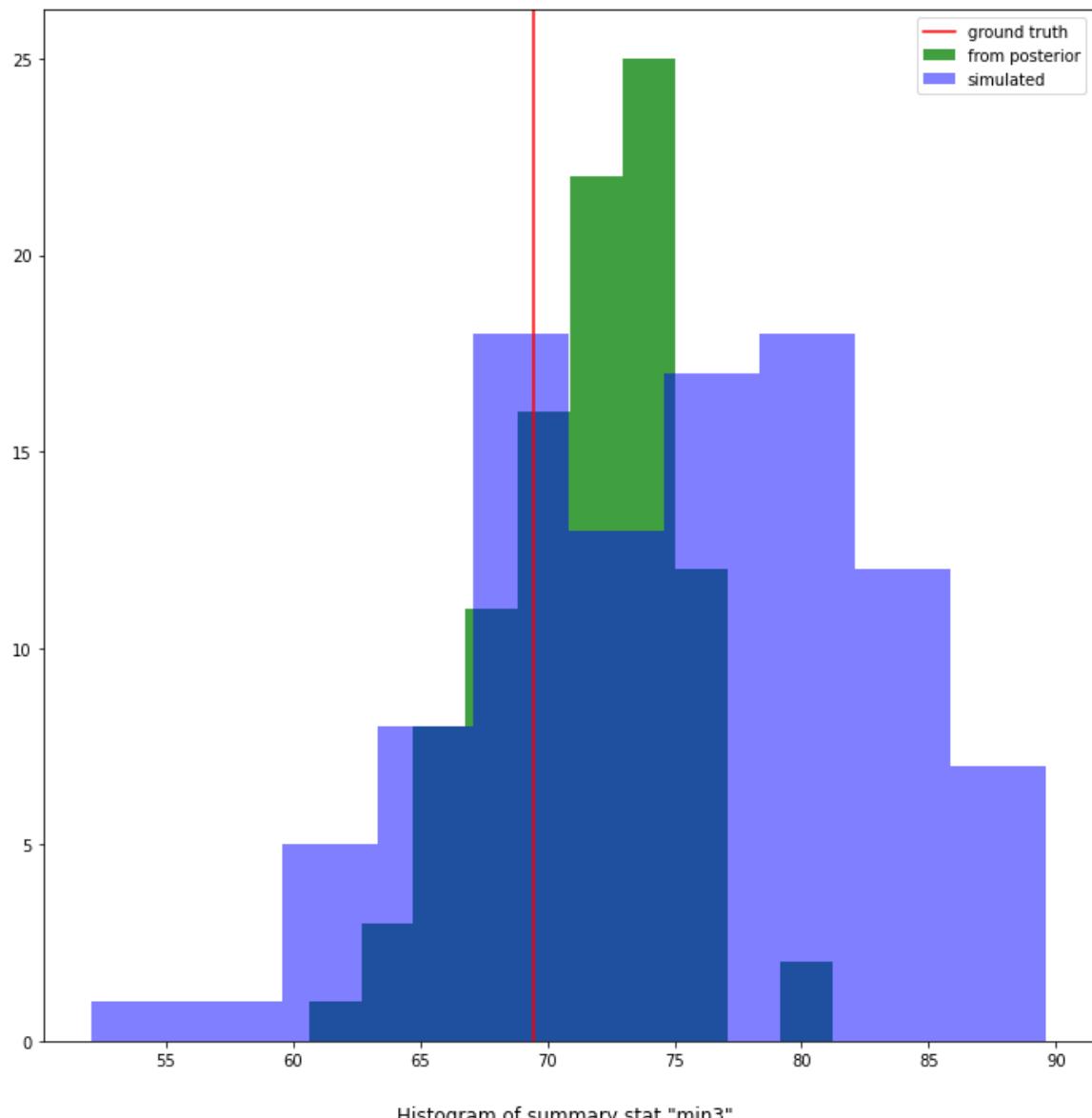
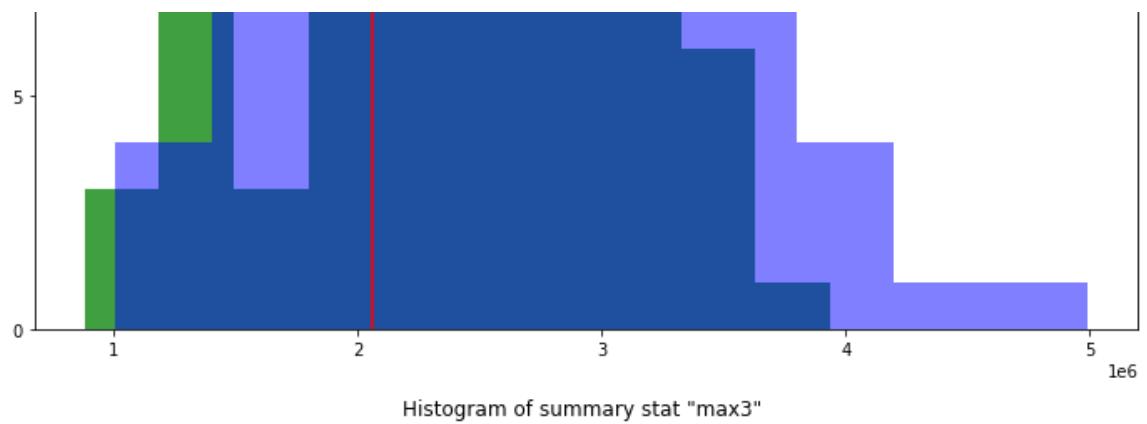


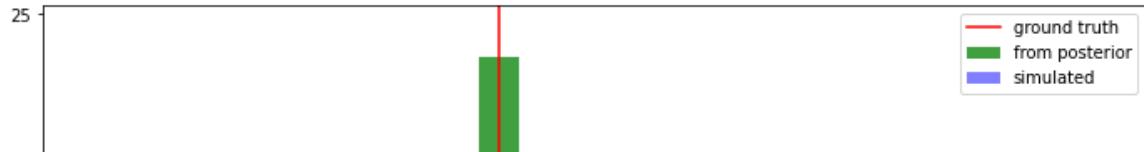
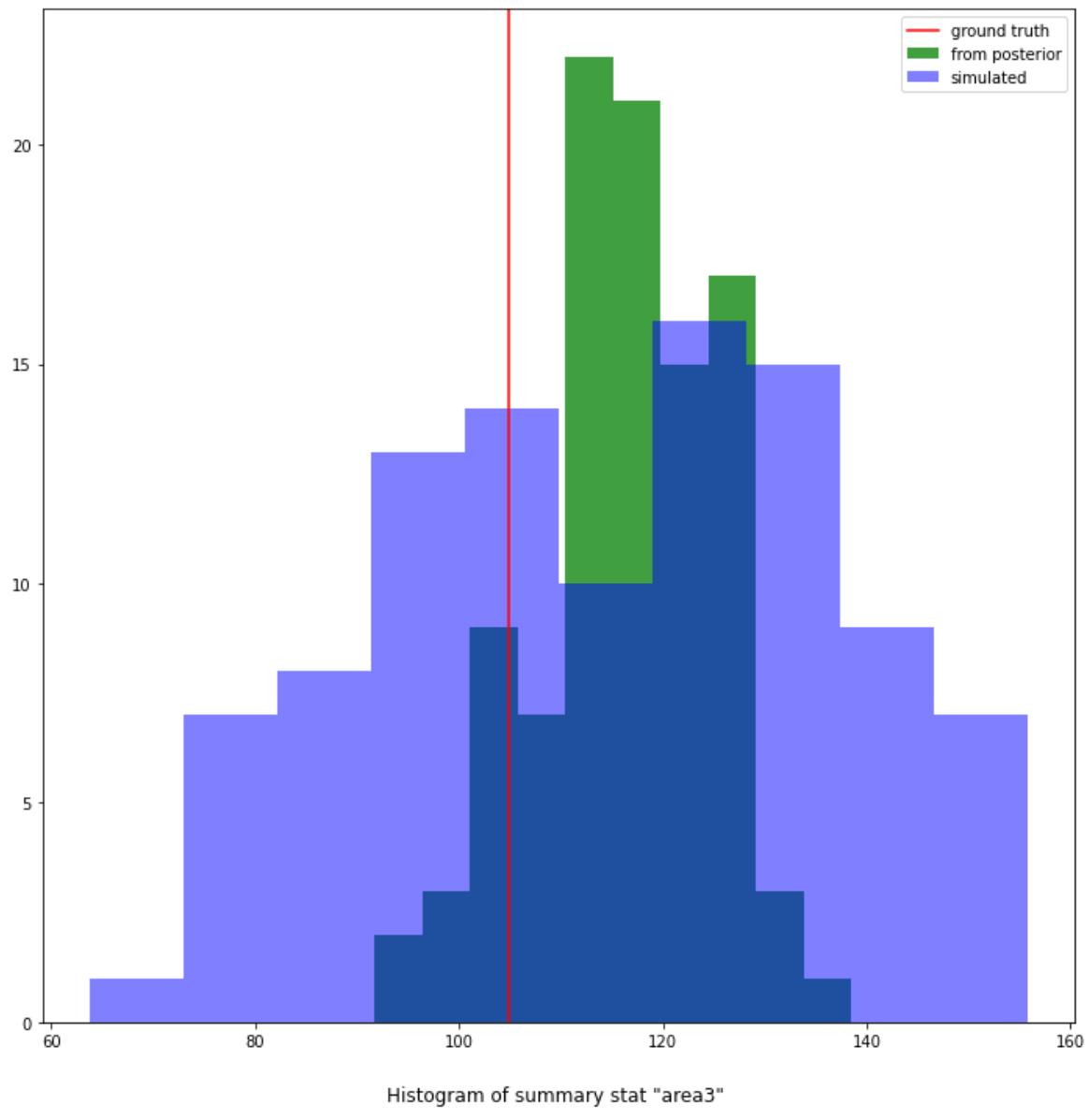
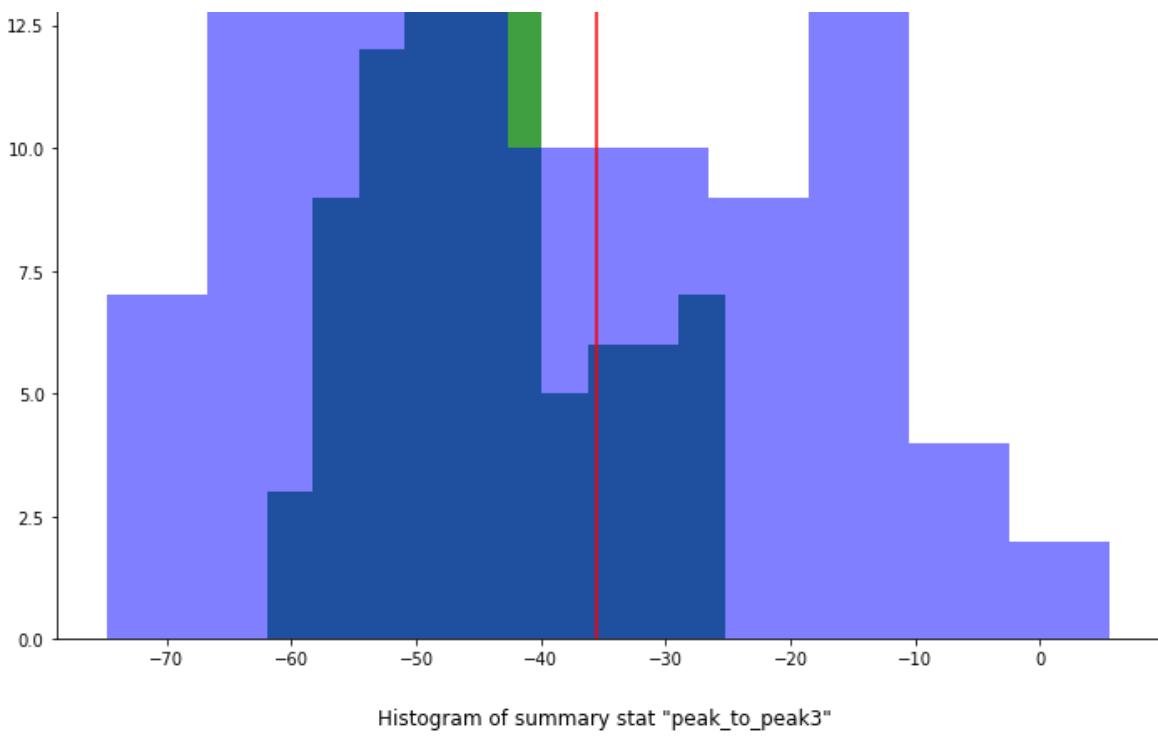


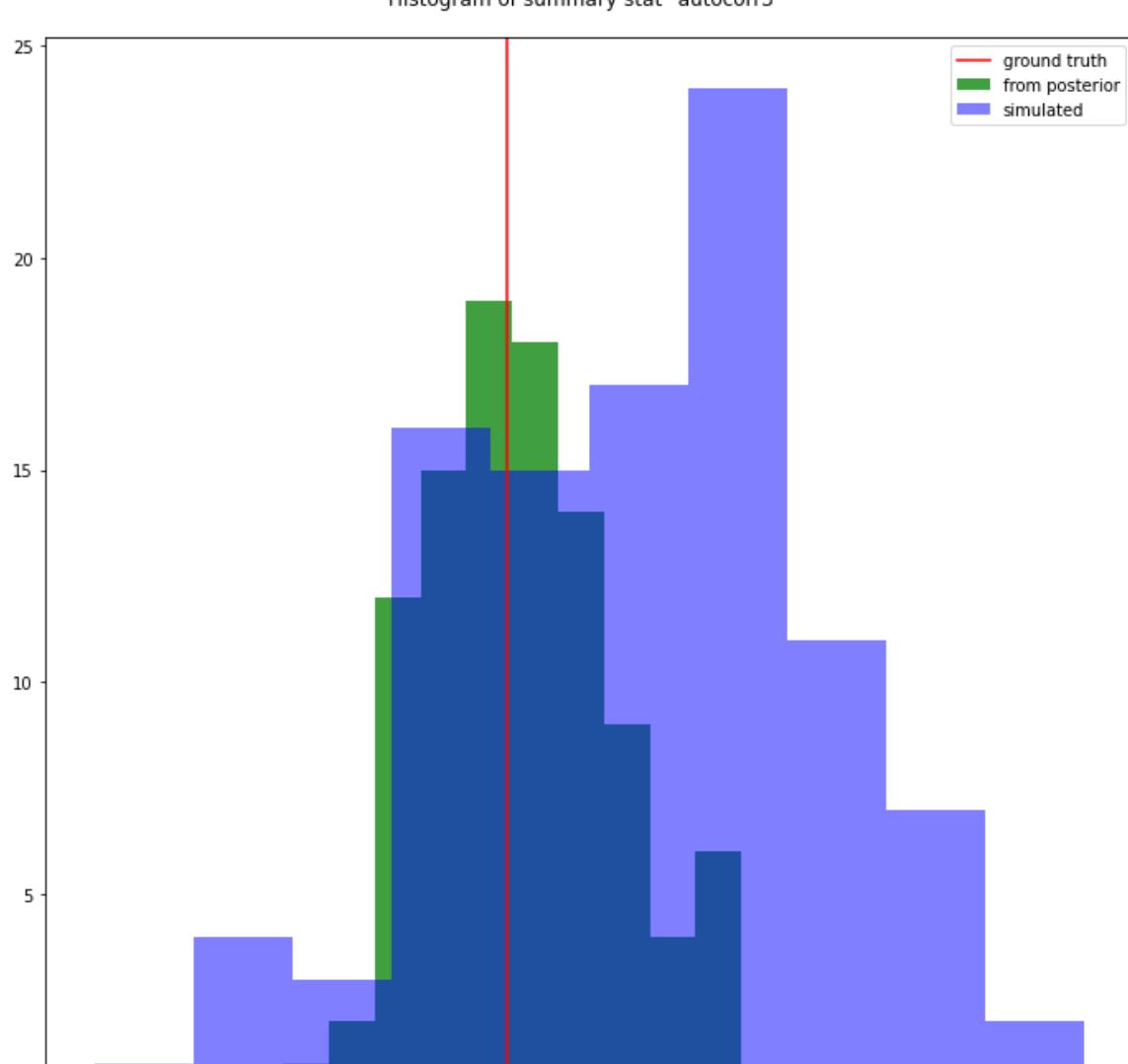
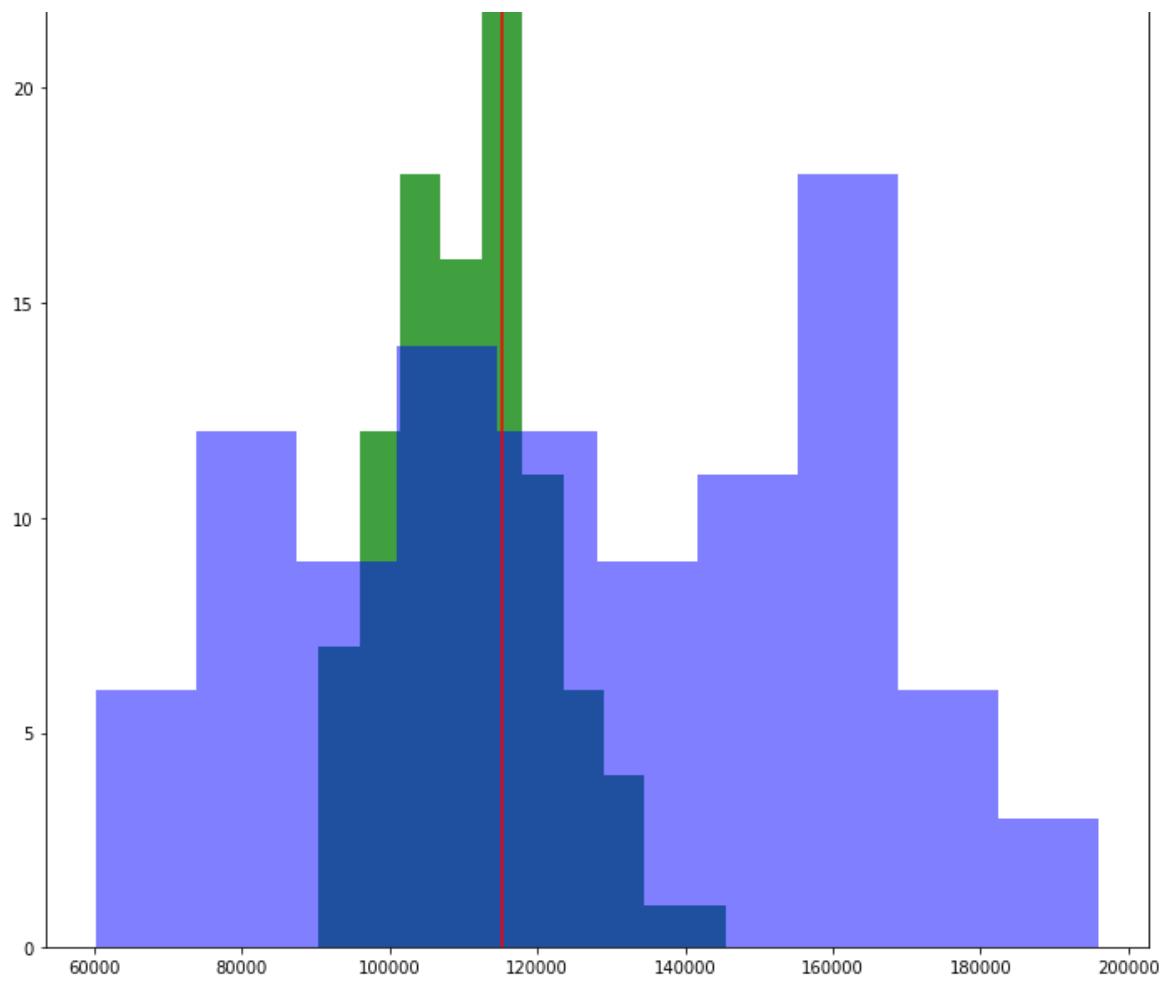
Histogram of summary stat "peak_to_peak2"

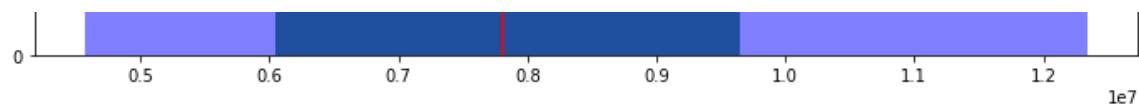






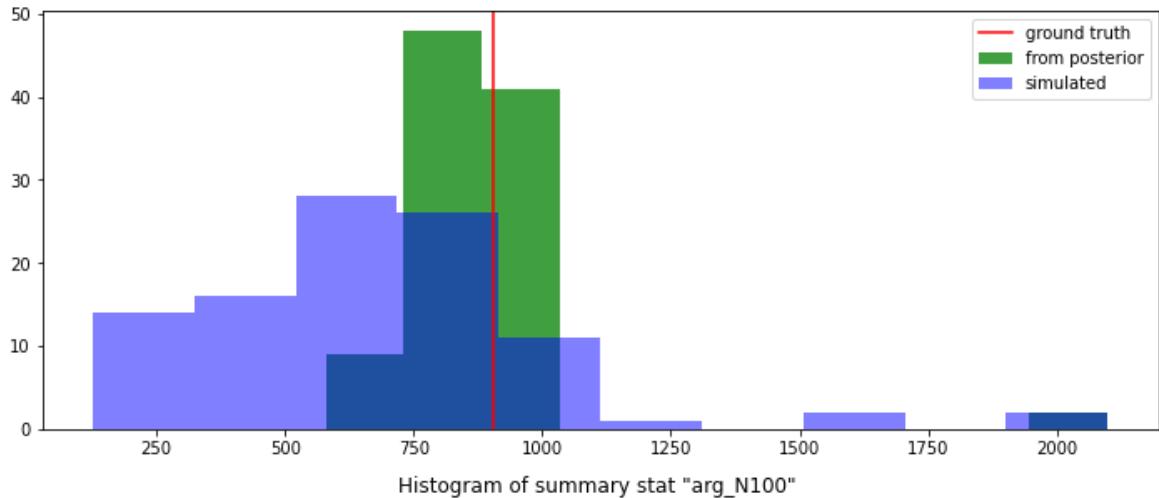




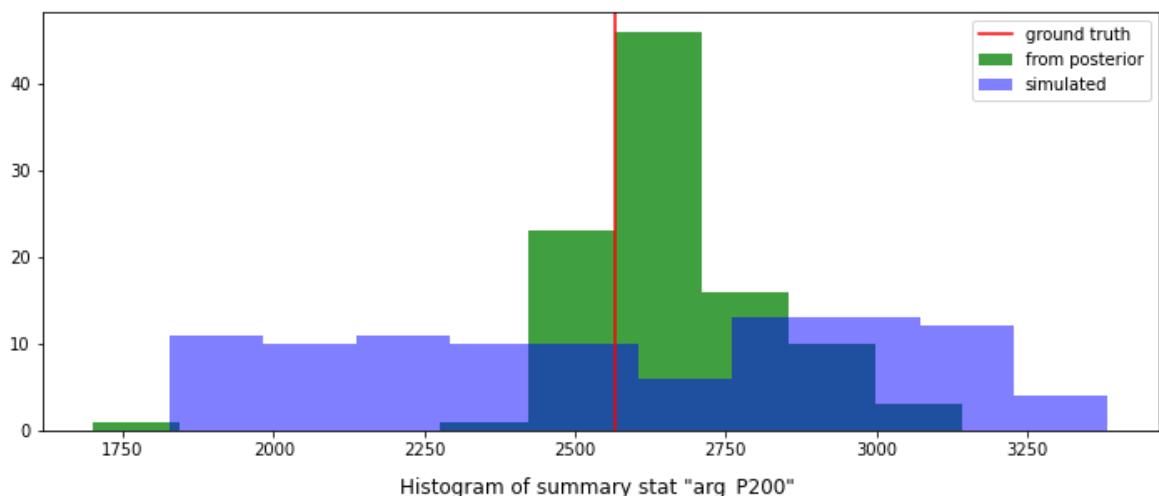


for the other 9 summary feature:

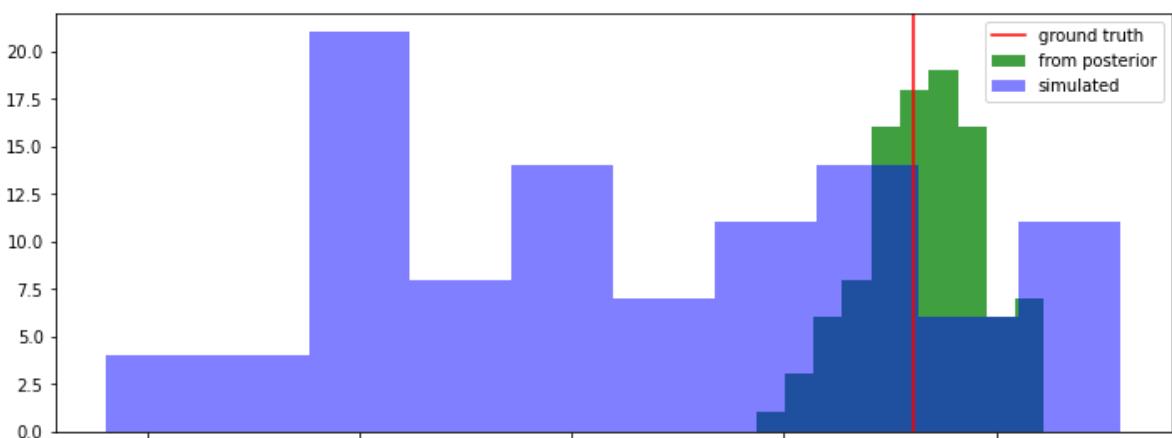
Histogram of summary stat "arg_p50"



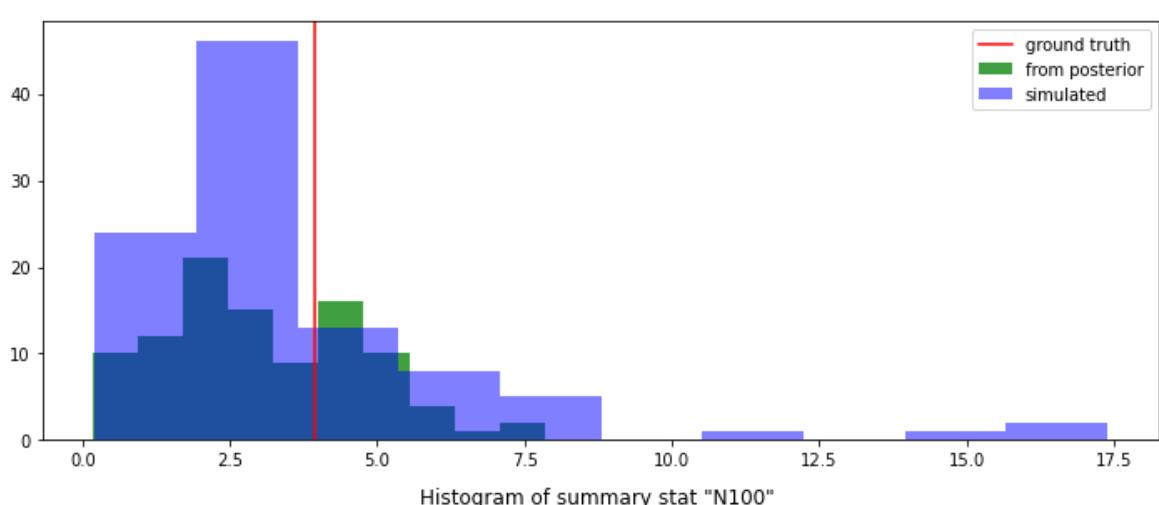
Histogram of summary stat "arg_N100"



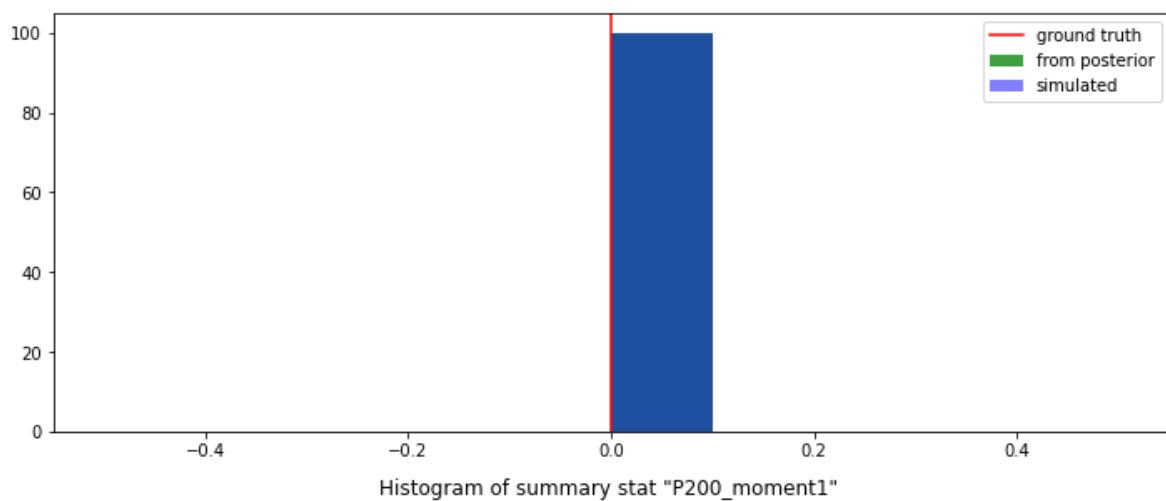
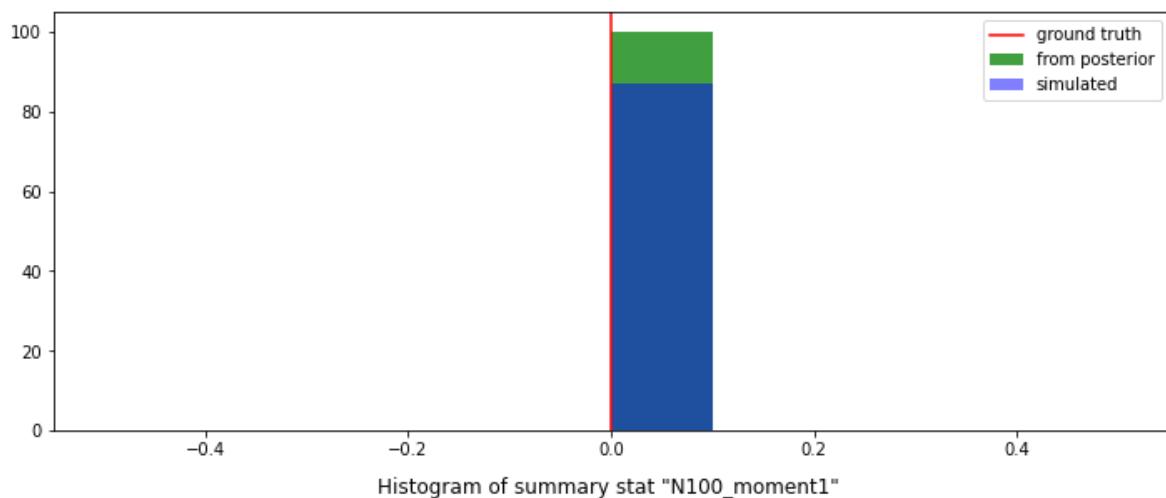
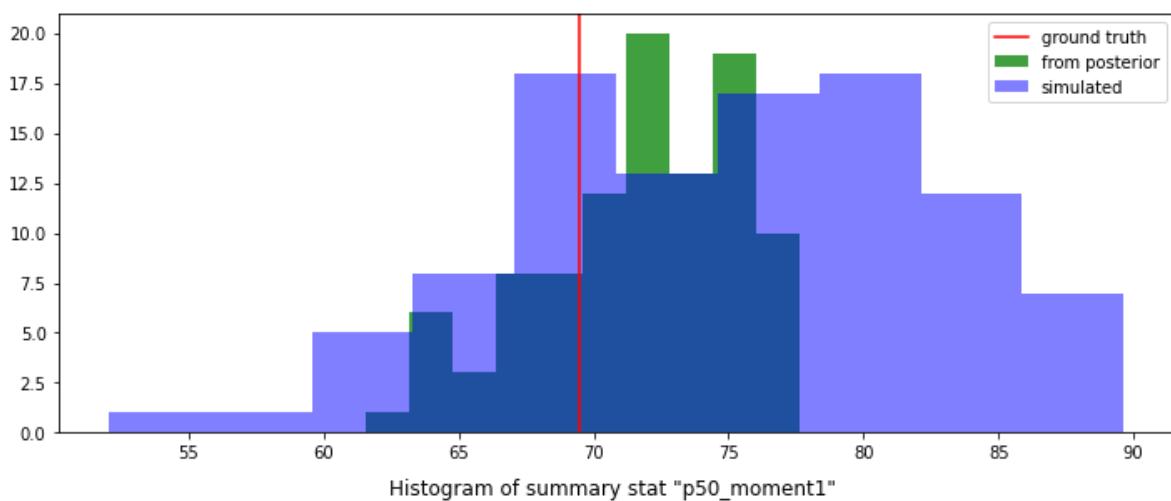
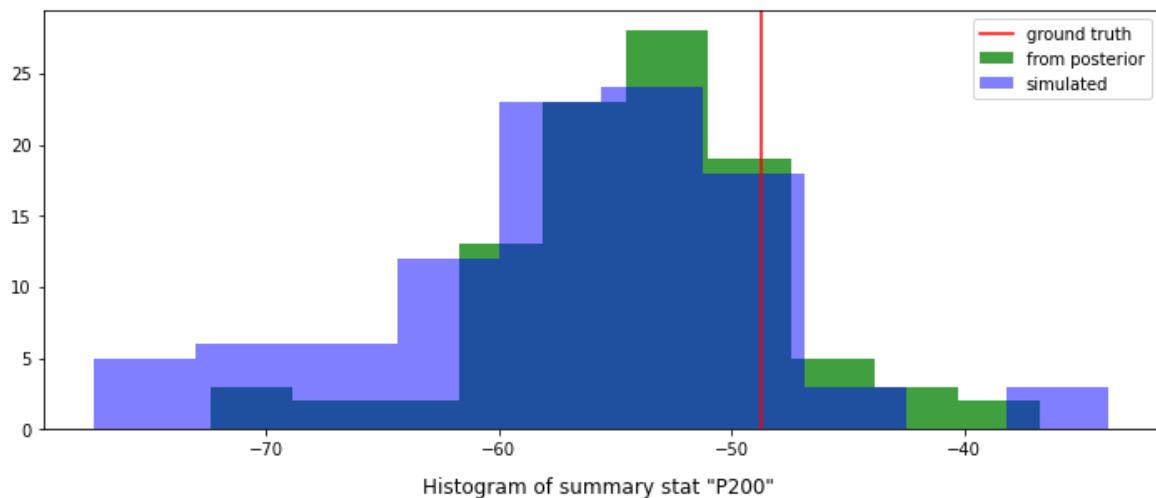
Histogram of summary stat "arg_P200"

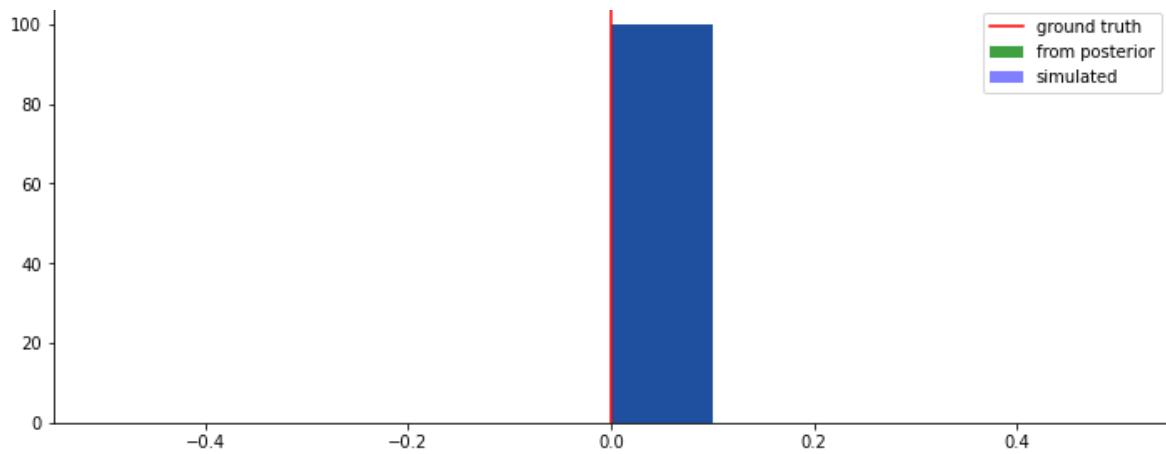


Histogram of summary stat "p50"



Histogram of summary stat "N100"





Weitere Punkte:

- ähnlich wie Mila in ihrer Lab rotation: definieren von feature ranges: z.B es darf zwischen 3 und 6 peaks geben oder ähnlich
- wie gut kann Evaluation überhaupt sein ohne wirklichen ground truth?