

Training set composition:

nbad>0 reals: 20

nbad>0 bogus: 2040

nbad=0 reals: 7916

nbad=0 bogus: 10947

nbad=0 : 18863

nbad=1 : 15

nbad=2 : 28

nbad=3 : 21

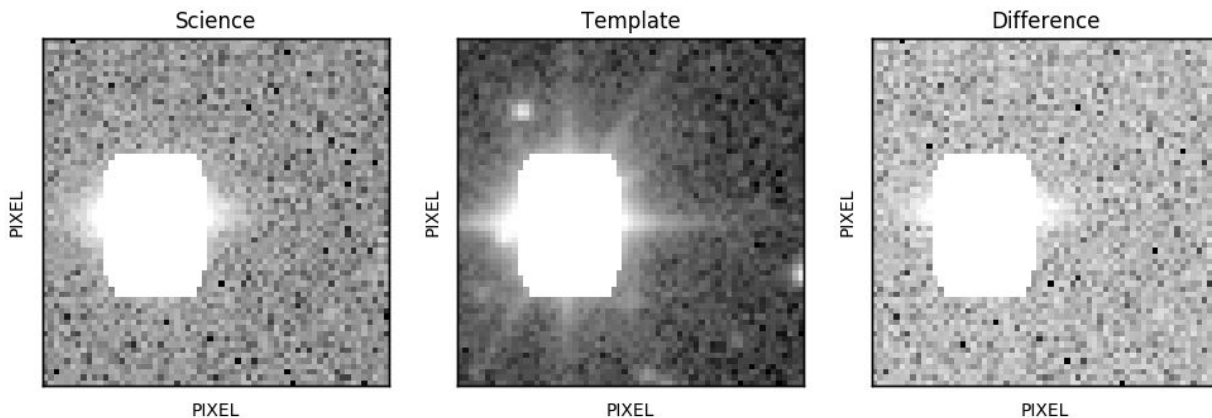
nbad=4 : 63

nbad=5 : 1904

nbad=6 : 21

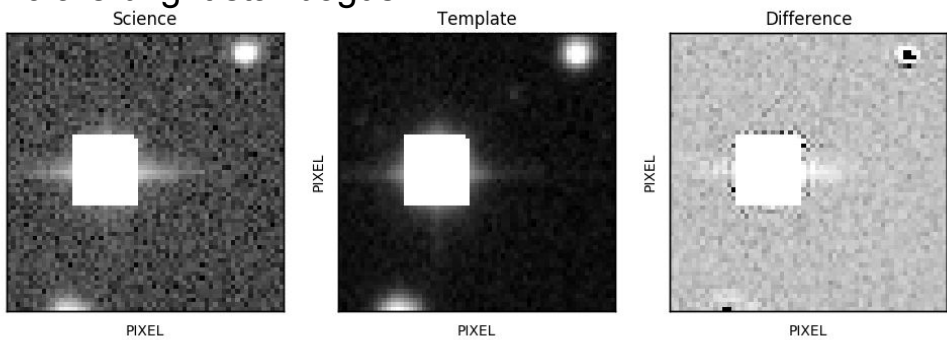
nbad=7 : 8

The 2040 nbad>0 bogus are completely dominated by bright stars.

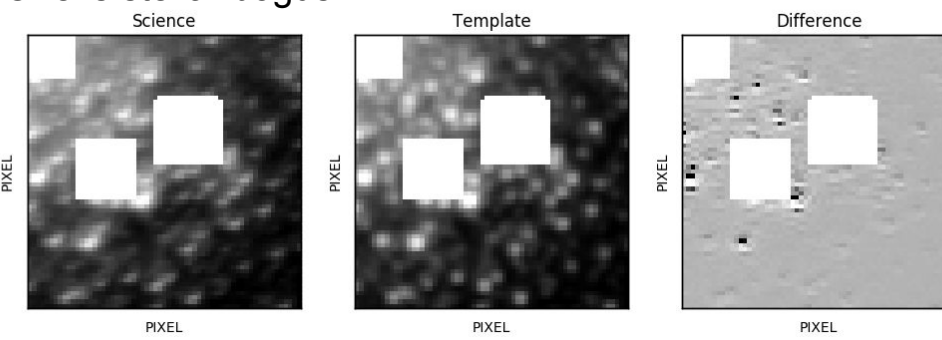


The 20 training set reals have major contamination.

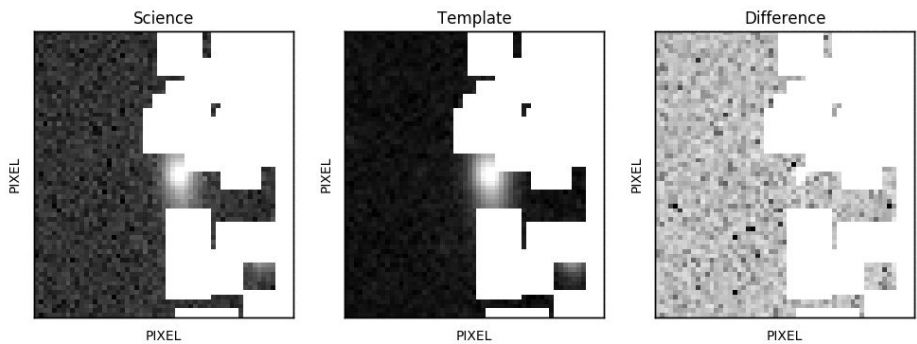
10 are bright star bogus.



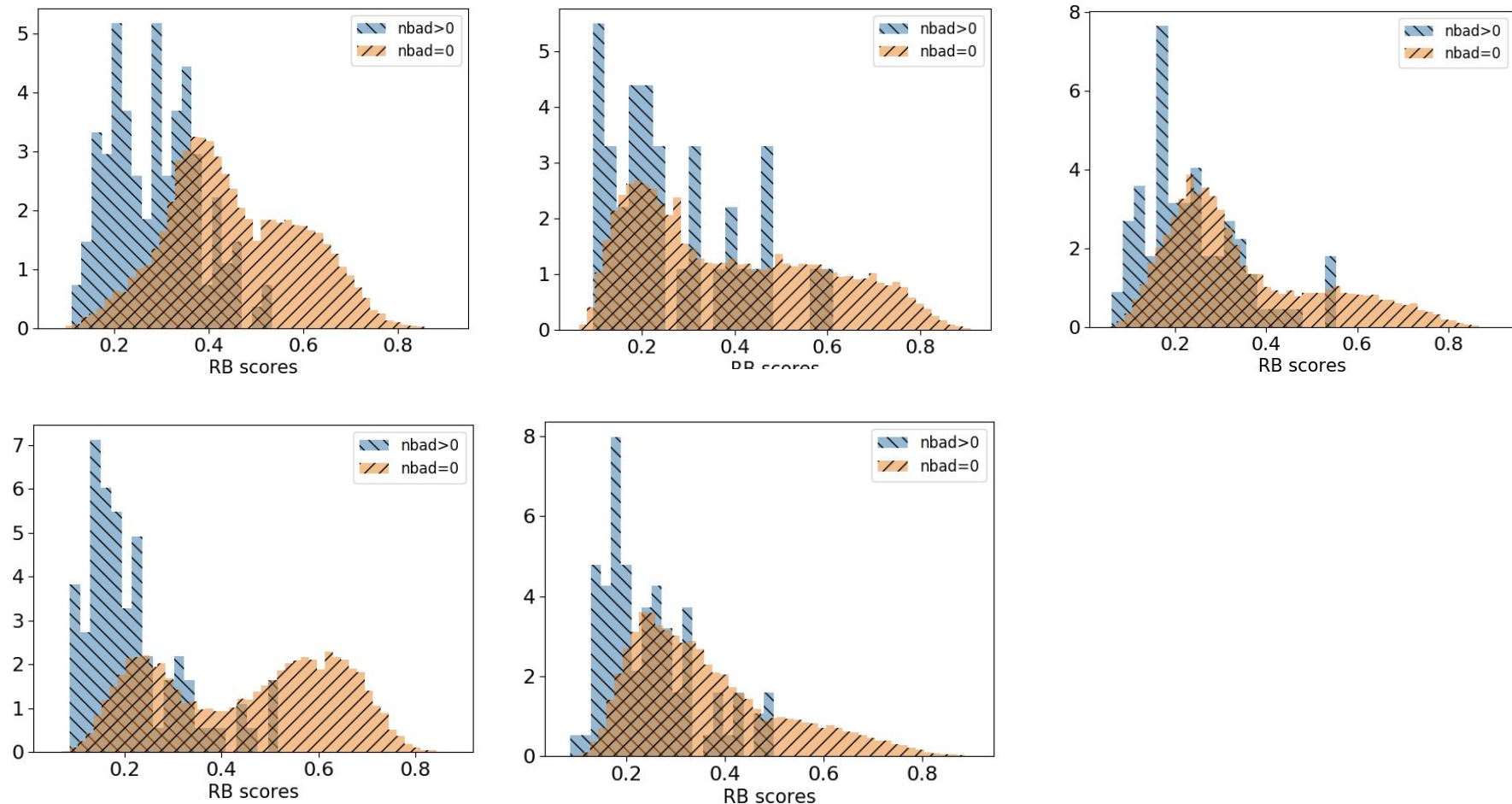
One is stellar bogus:



The other 9 are reals like this:

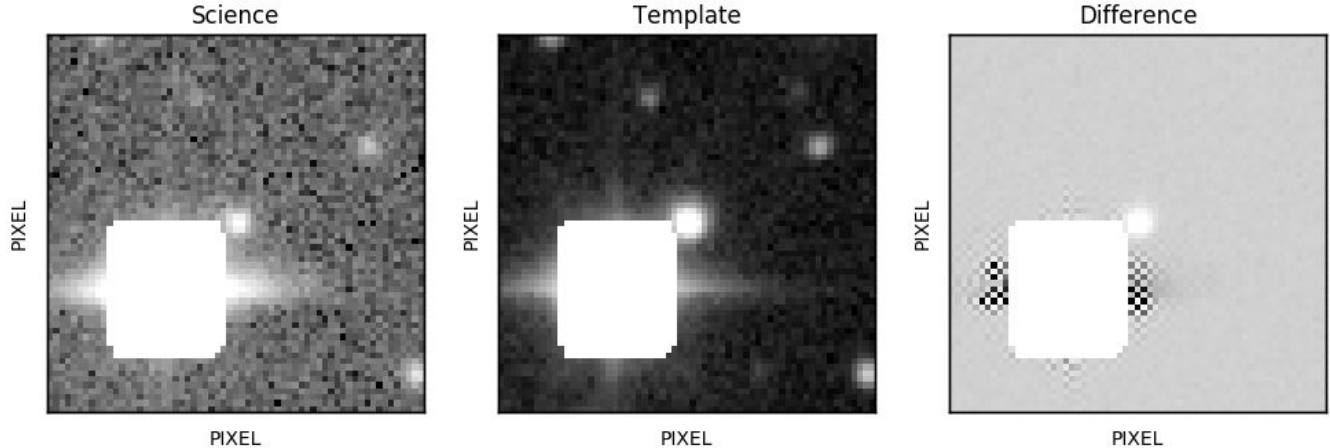


RB scores of recent nightly data broken down by nbad>0 and nbad =0:

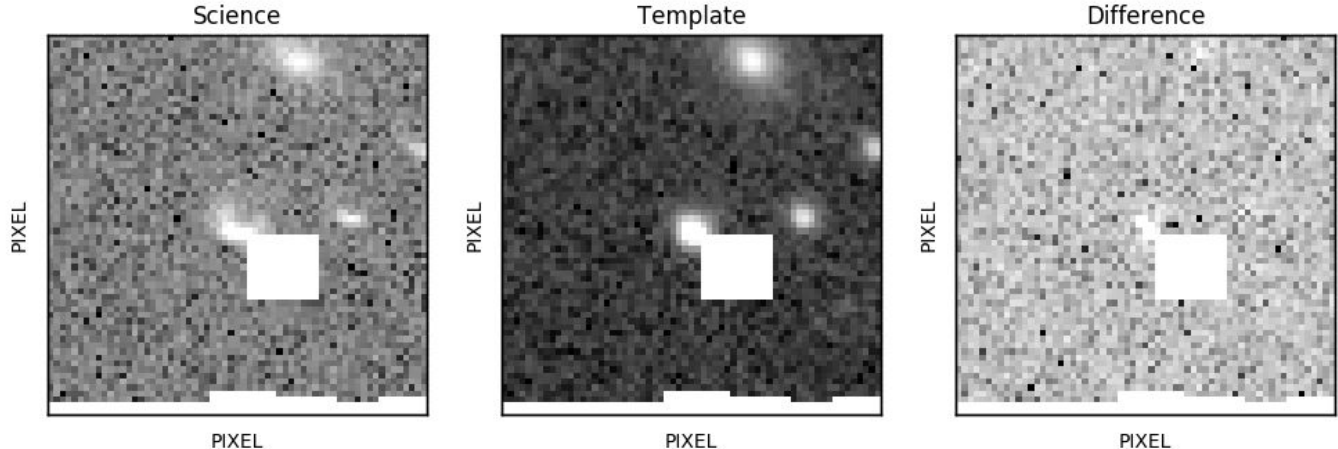


RB scores are high (0.4-0.6) when there is a real source at the centre, even if there is a bright star nearby.

E.g. RB score 0.61

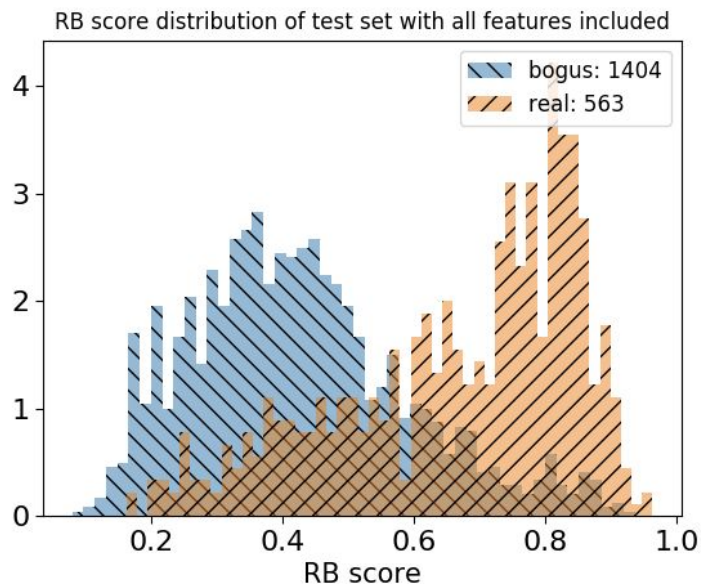


RB score 0.48

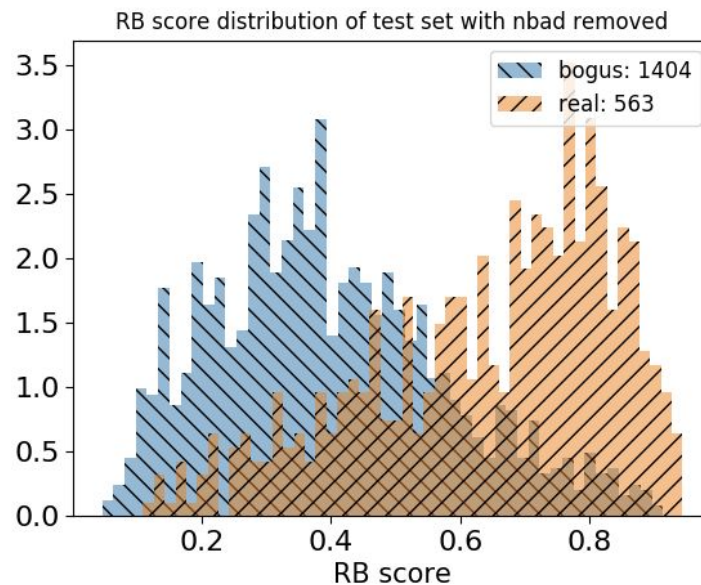


Removing nbad did not significantly change test set RB distributions:

With nbad:



Without nbad:



Removing nbad increases training set FNR at 0.01 FPR by 0.02 (but this is on the training set where nbad>0 is poorly represented).