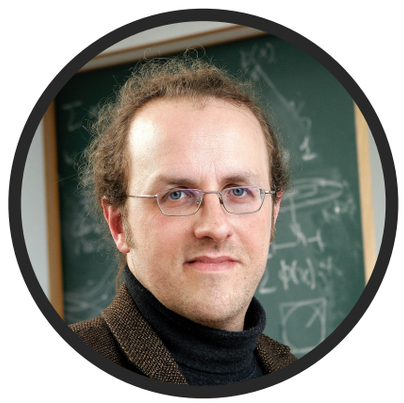
the search for Single Transits

Dan Foreman-Mackey

NYU→UW // github.com/dfm // @exoplaneteer // dfm.io



David W. Hogg
NYU

Bernhard Schölkopf
MPI-IS

Population Inference

treatment of false positives,
dependent parameters,
uncertainties & selection effects

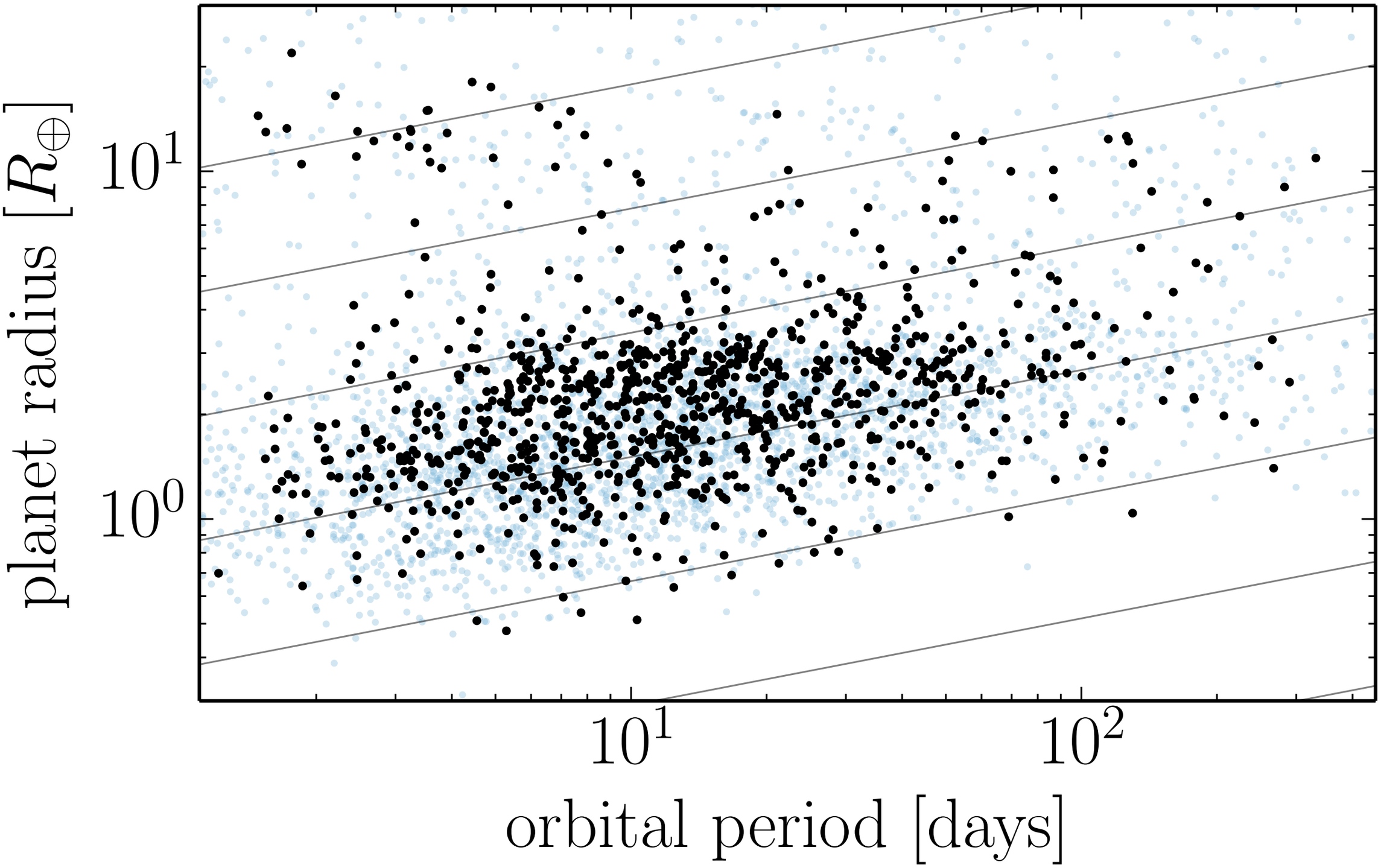
open source tools
applicable to all existing
& future exoplanet missions

Flexible & **robust** **inference** of the **exoplanet population**

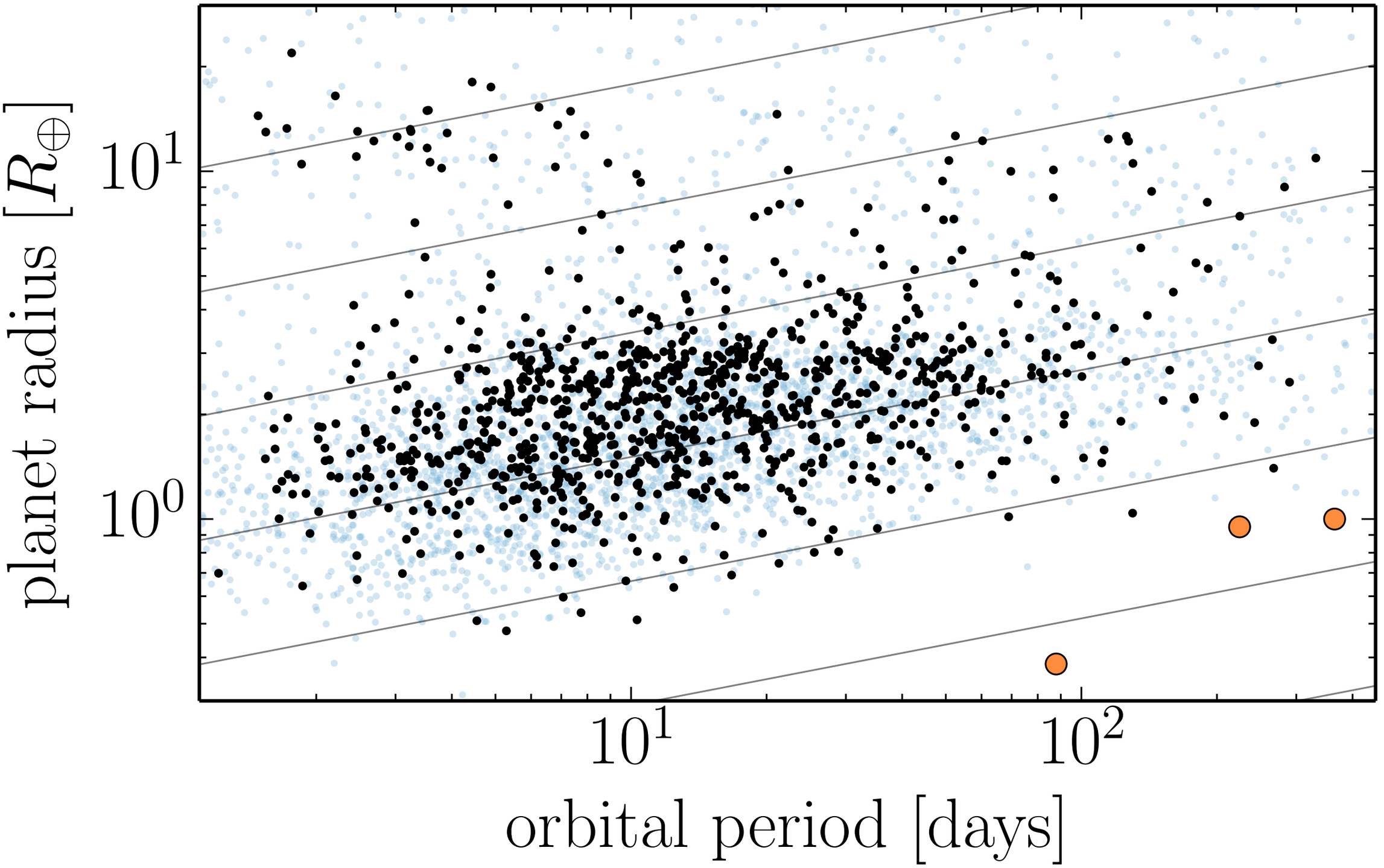
occurrence rate
period, radius, mass,
eccentricity, multiplicity,
mutual inclination, *etc.*

Ingredients of a population inference

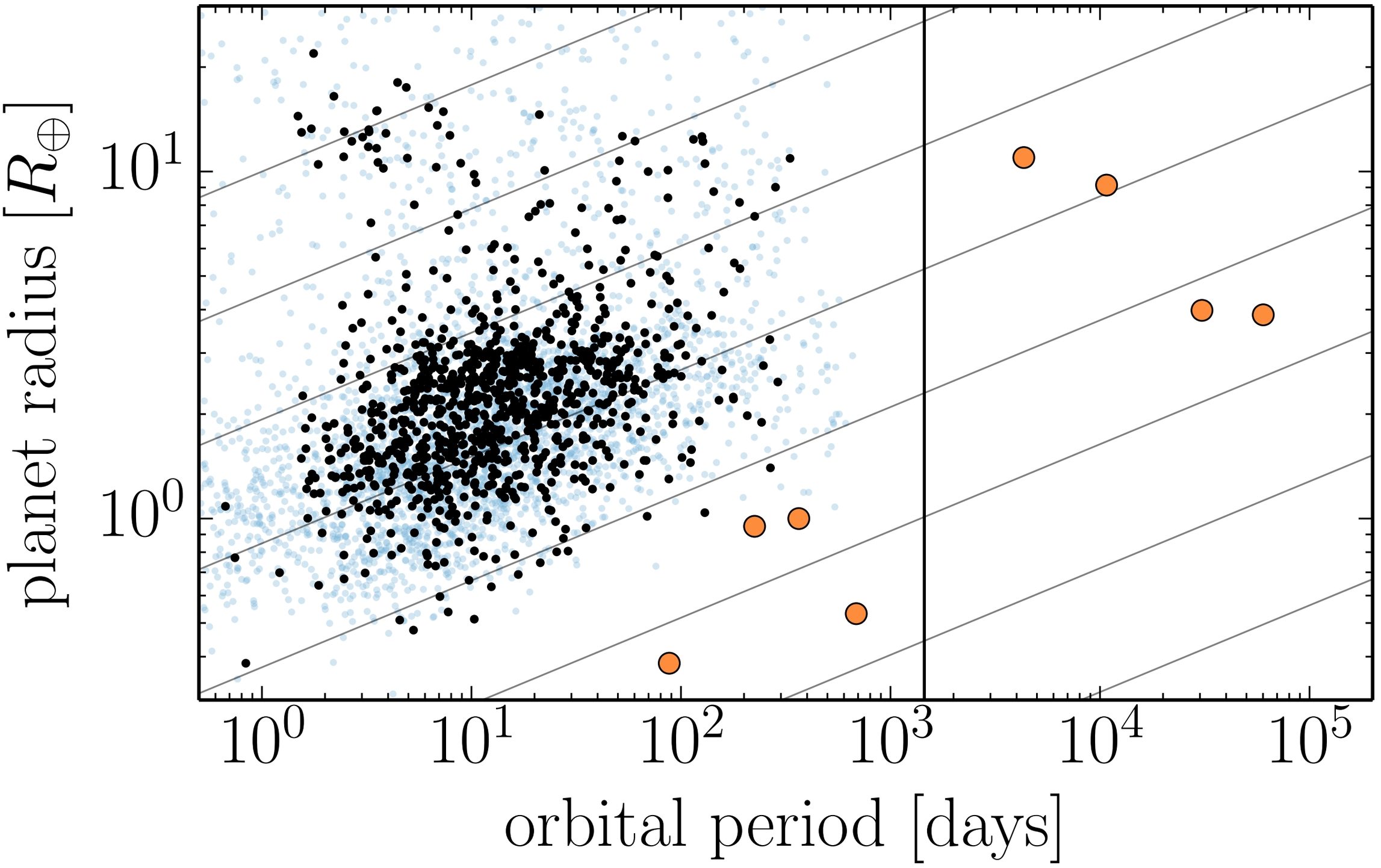
1. catalog of planet (candidates)
2. measurement of completeness
3. measurement of precision



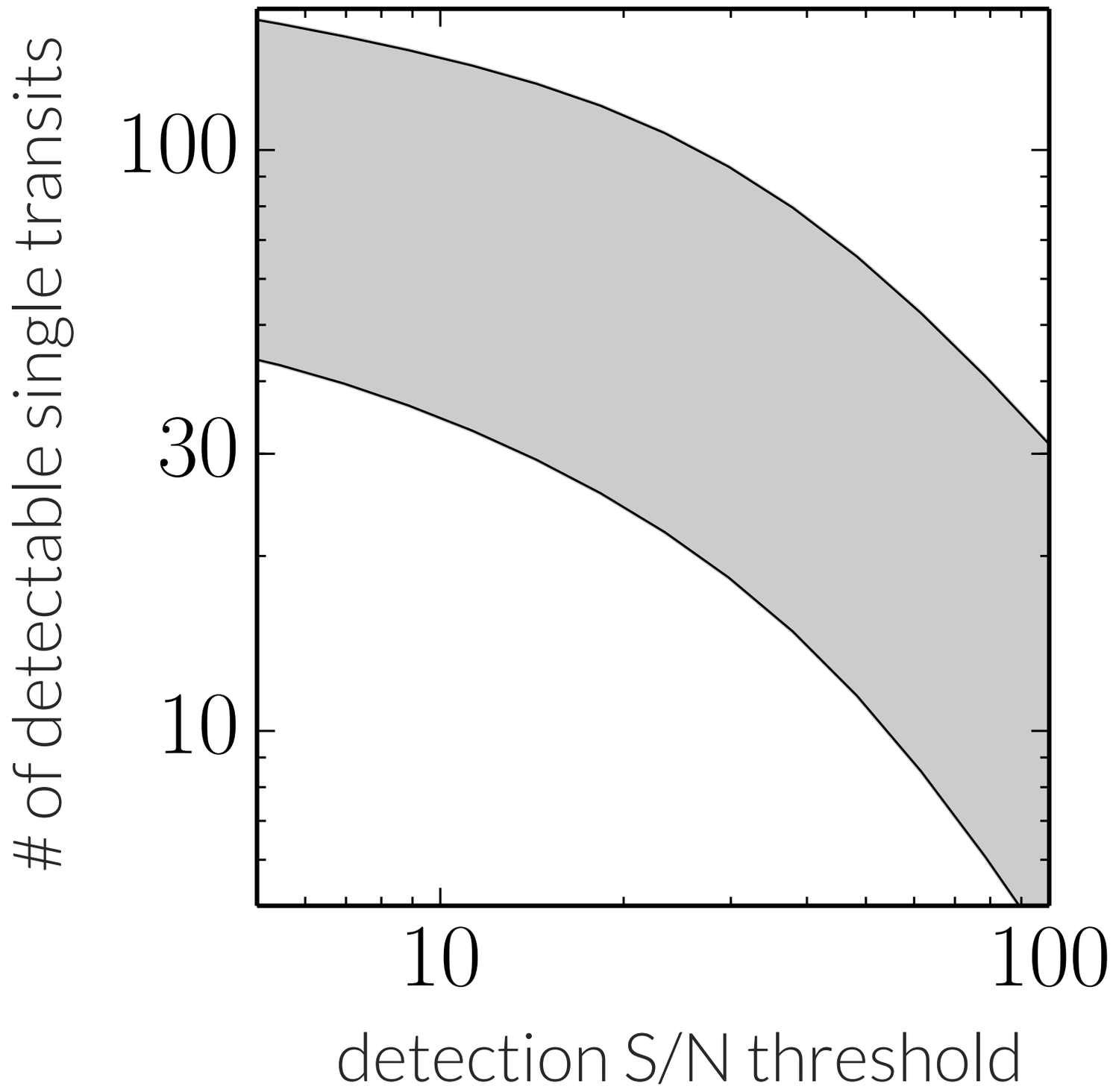
Data from **NASA Exoplanet Archive**



Data from **NASA Exoplanet Archive**



Data from **NASA Exoplanet Archive**



Extrapolated from **Dong & Zhu (2013)**

How to find a Transiting Planet

the traditional way…

How to find a (periodic) transit signal

1. de-trending
2. grid search in period, phase, and duration
3. vetting of candidates

False Alarms
&
False Positives

How to find a Transiting Planet

the Planet Hunters way…

Planet Hunters

Dan

www.planethunters.org/#/classify

Apps | new arxiv | Recent ArXiv | Betterize (with a bett | NYU Proxy | keplerscience.arc.na | Kepler Science Cente | Using convolutional

PLANET HUNTERS

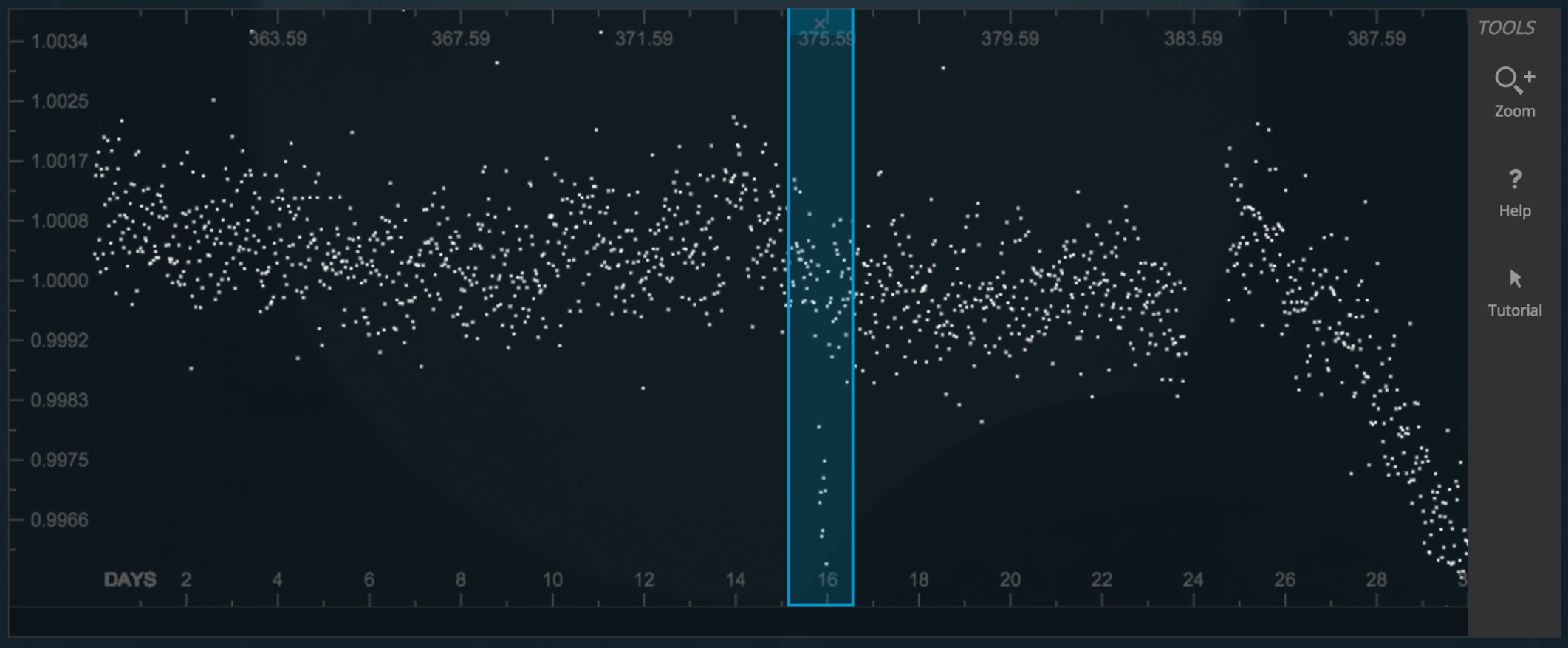
Classify | Science | About | Education | Profile | Talk | Blog

SIGN UP | SIGN IN | English

Do you see a transit?

If so, highlight it on the light curve below!

We have new K2 data! Help us classify it all.



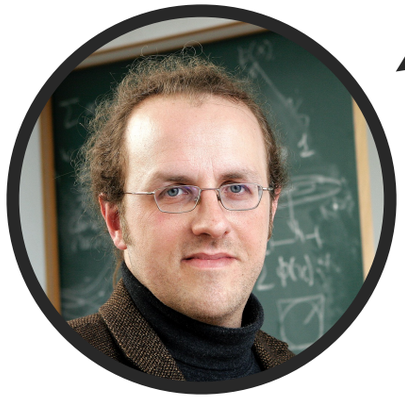
Star Information

Magnitude N/A Type N/A Temp N/A Radius N/A

Finished

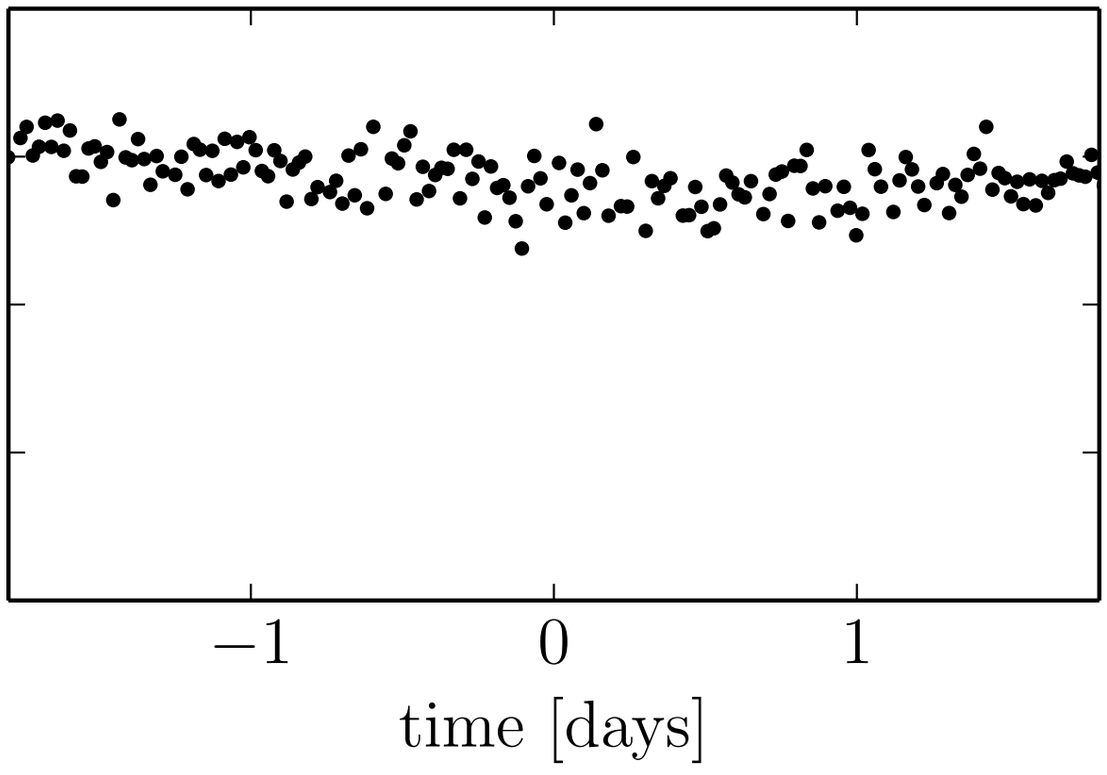
Can we Teach the Machine to Learn™?

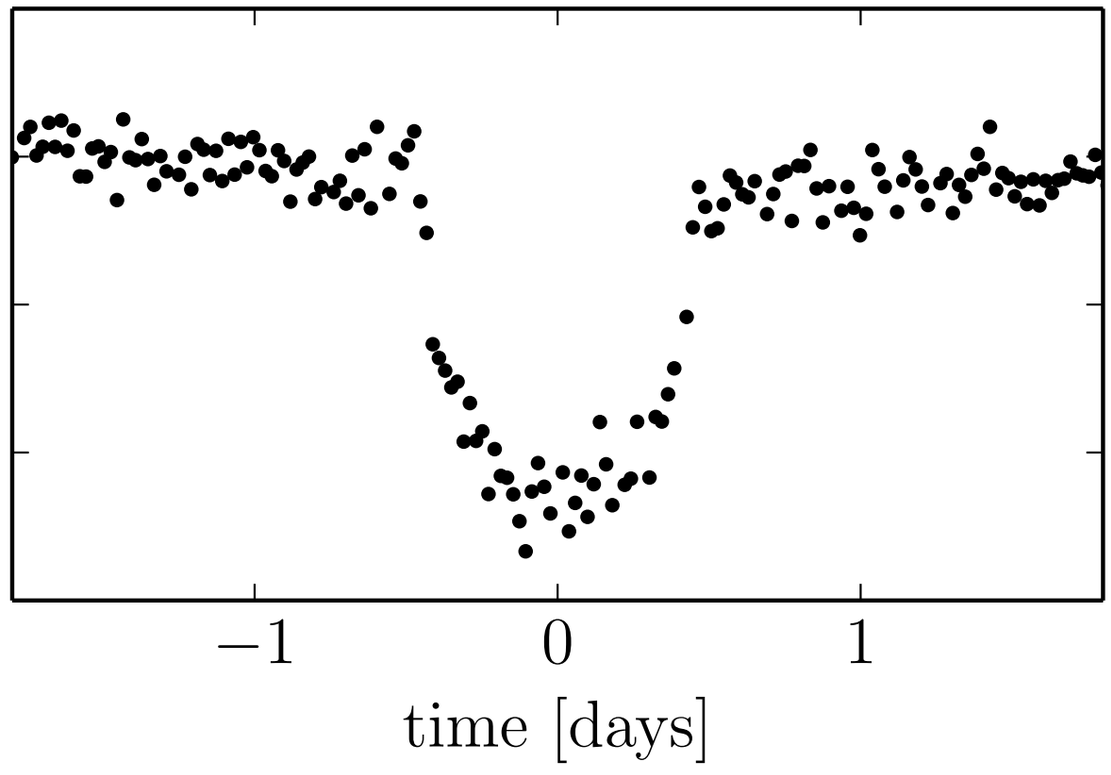
Get rid of the pipeline!



Bernhard Schölkopf
MPI-IS

Supervised Classification

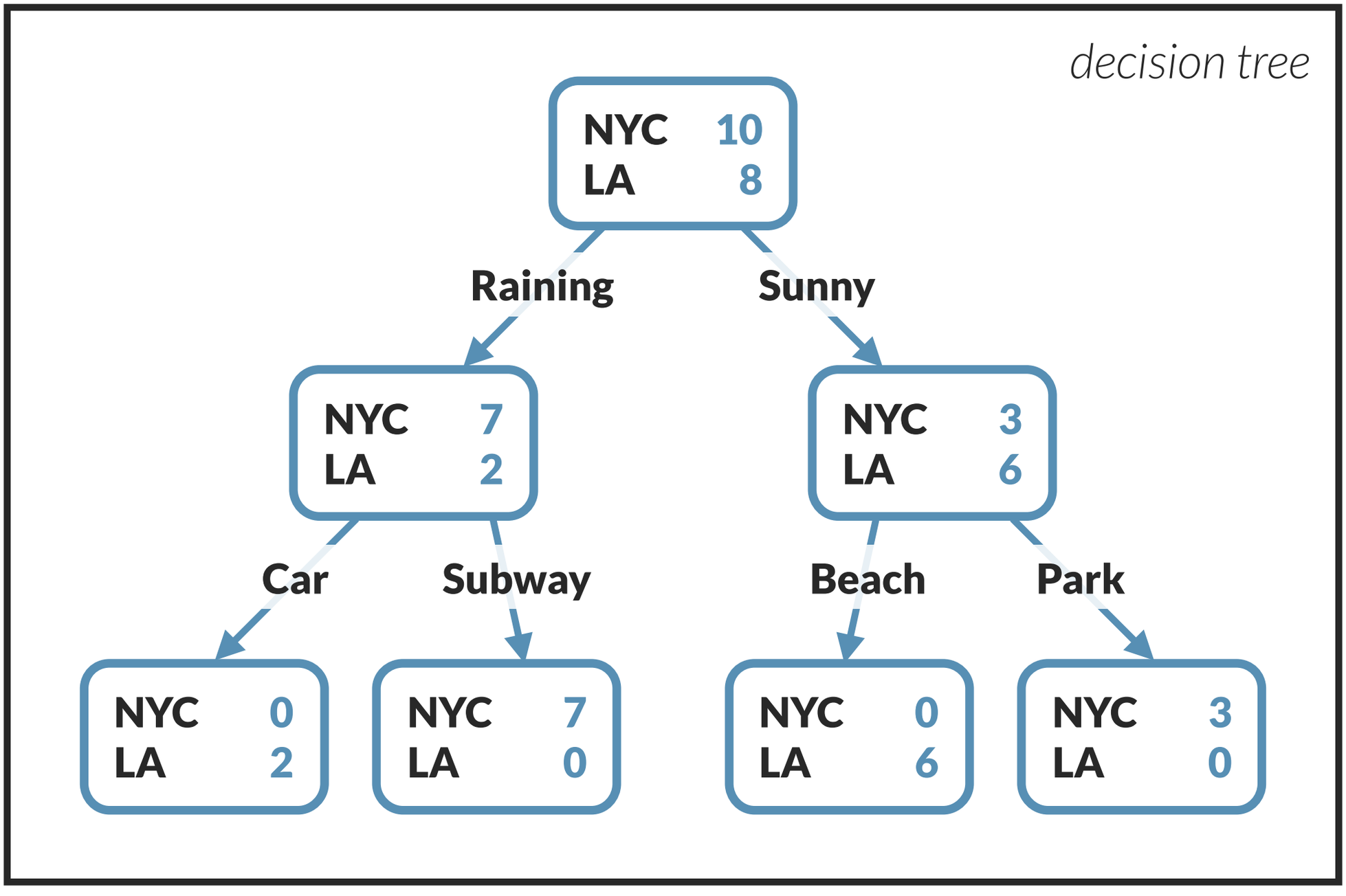




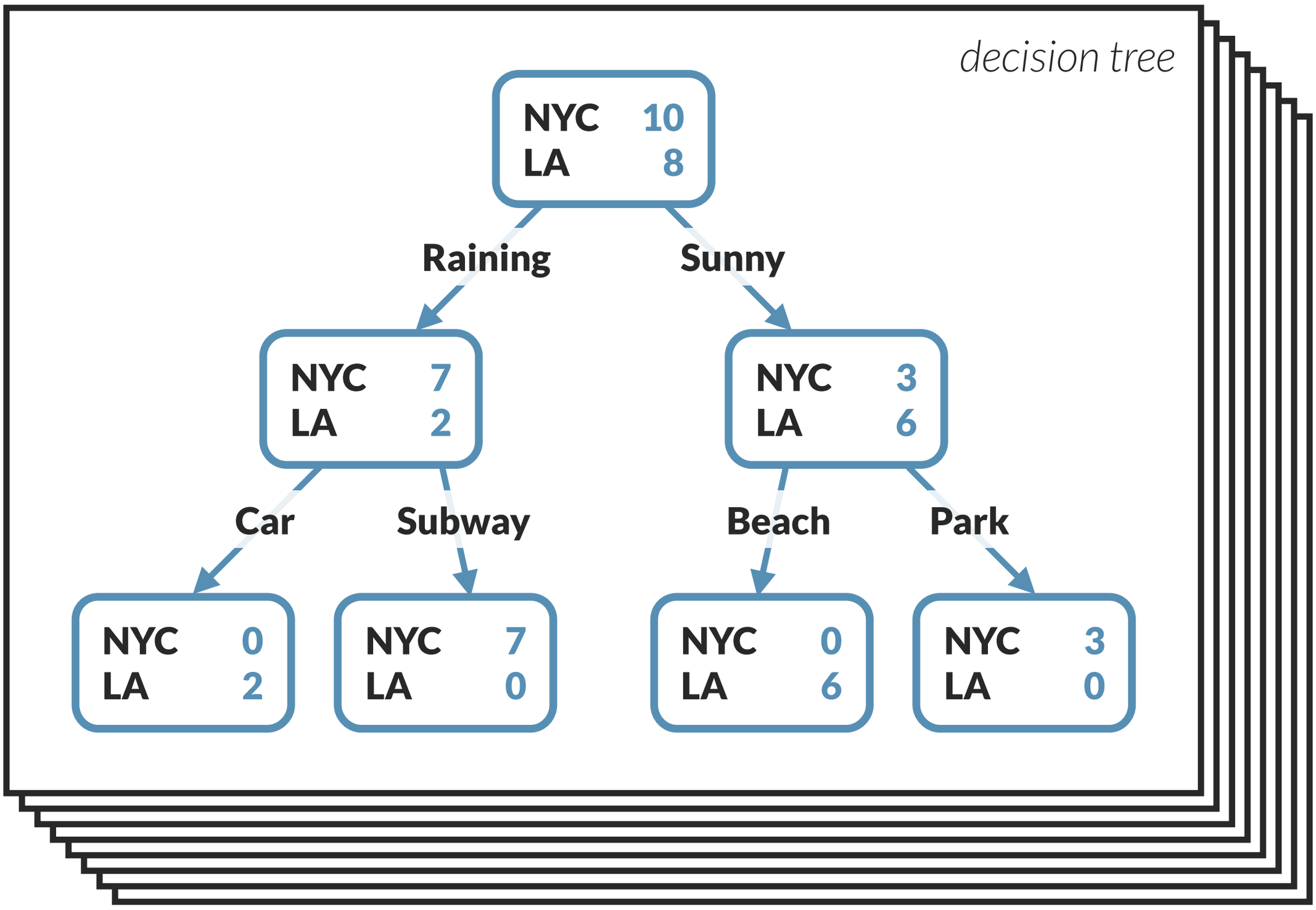
`no_transit` vs. `transit`

Supervised Classification

Random Forest™ Classification



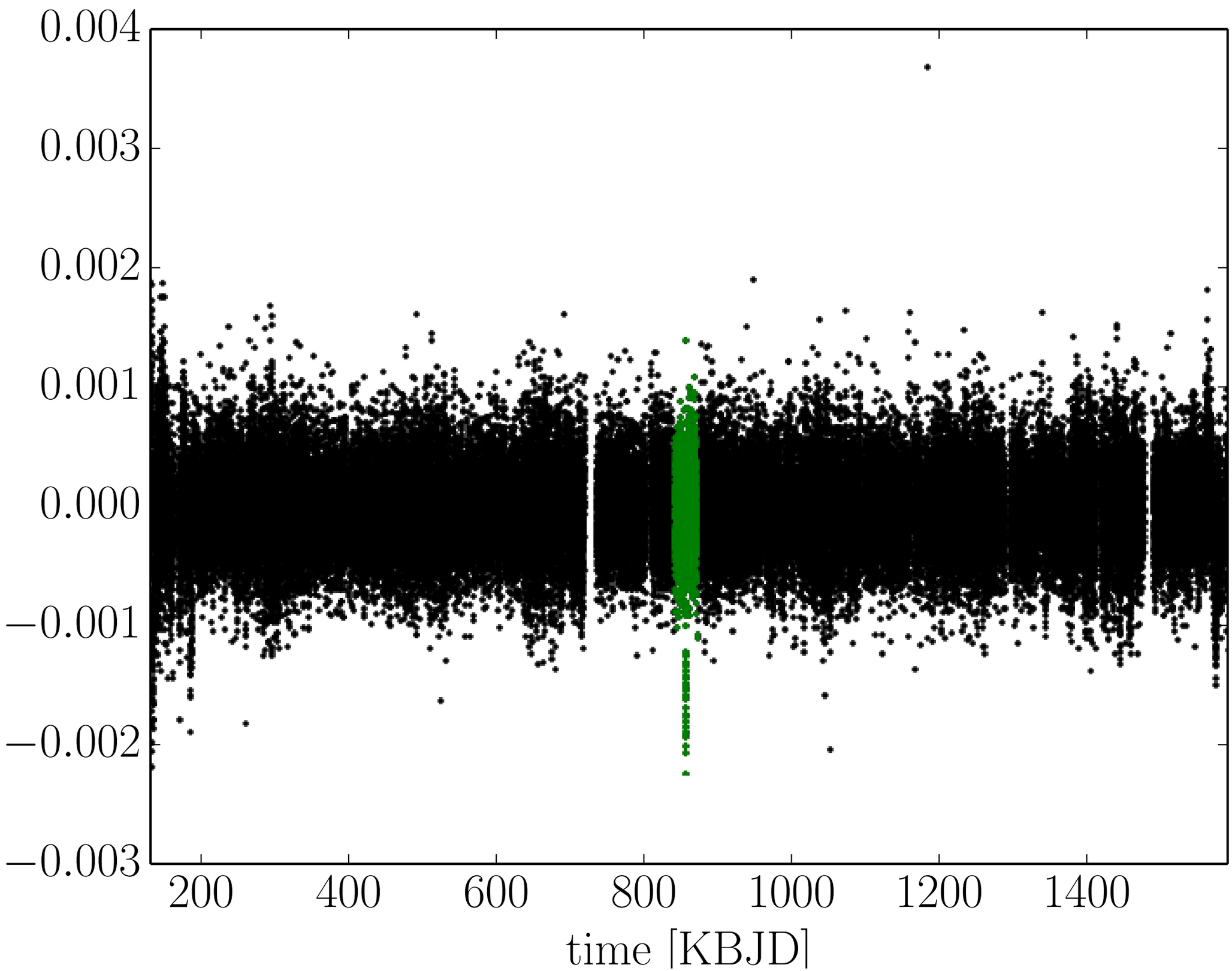
Random Forest™ Classification

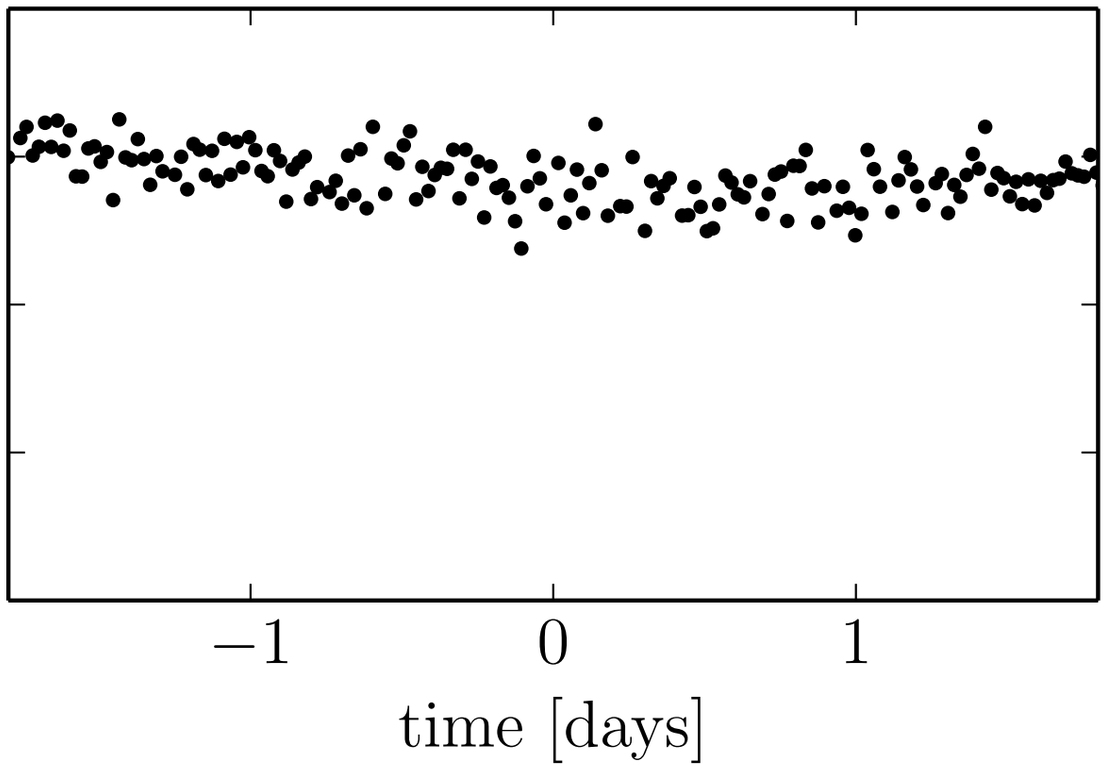


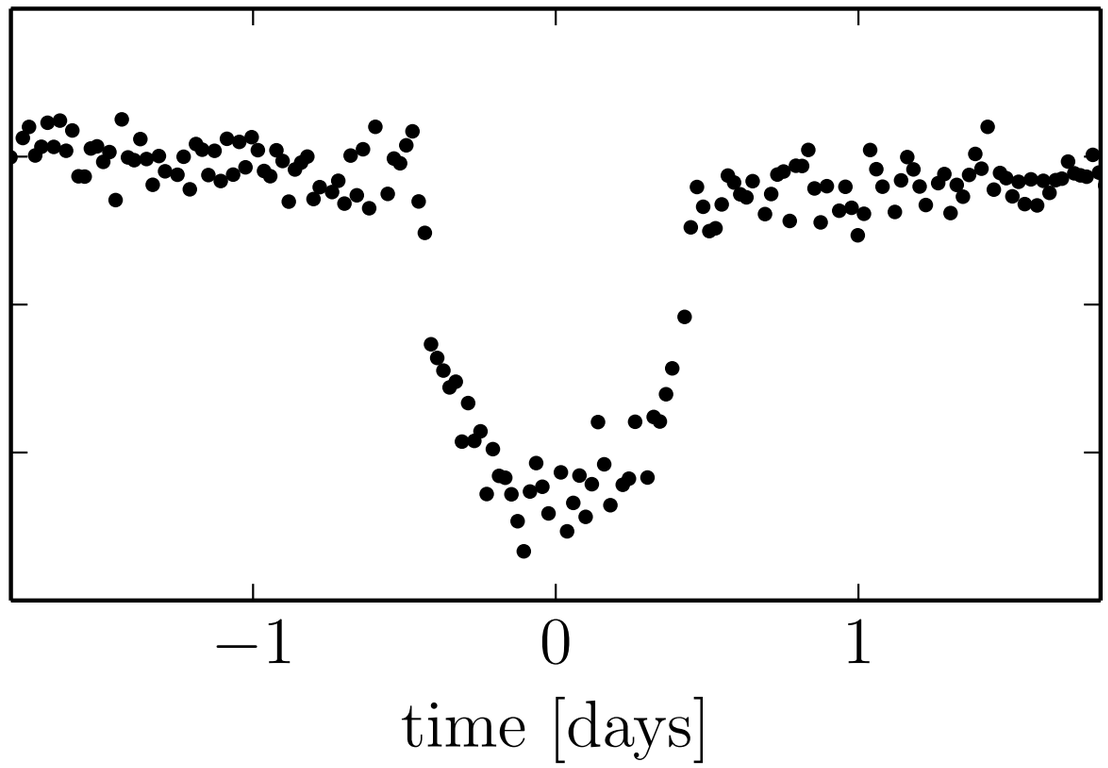
features light curve sections

training set simulated transits

test set held-out light curve


0.004
0.003
0.002
0.001
0.000
−0.001
−0.002
−0.003
200
400
600
800
1000
1200
1400
time [KBJD]





no_transit vs. transit

scikit-learn.org

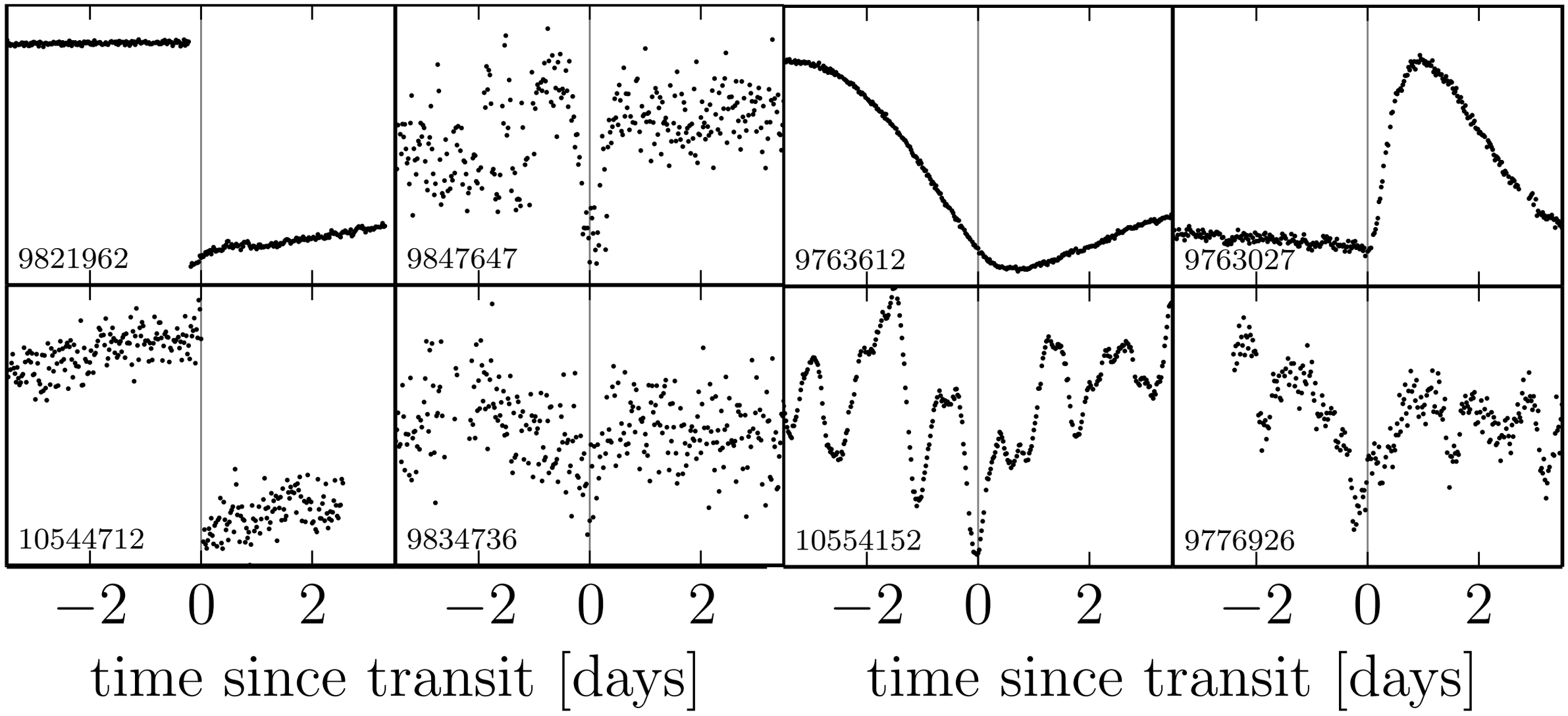
Preliminary Results

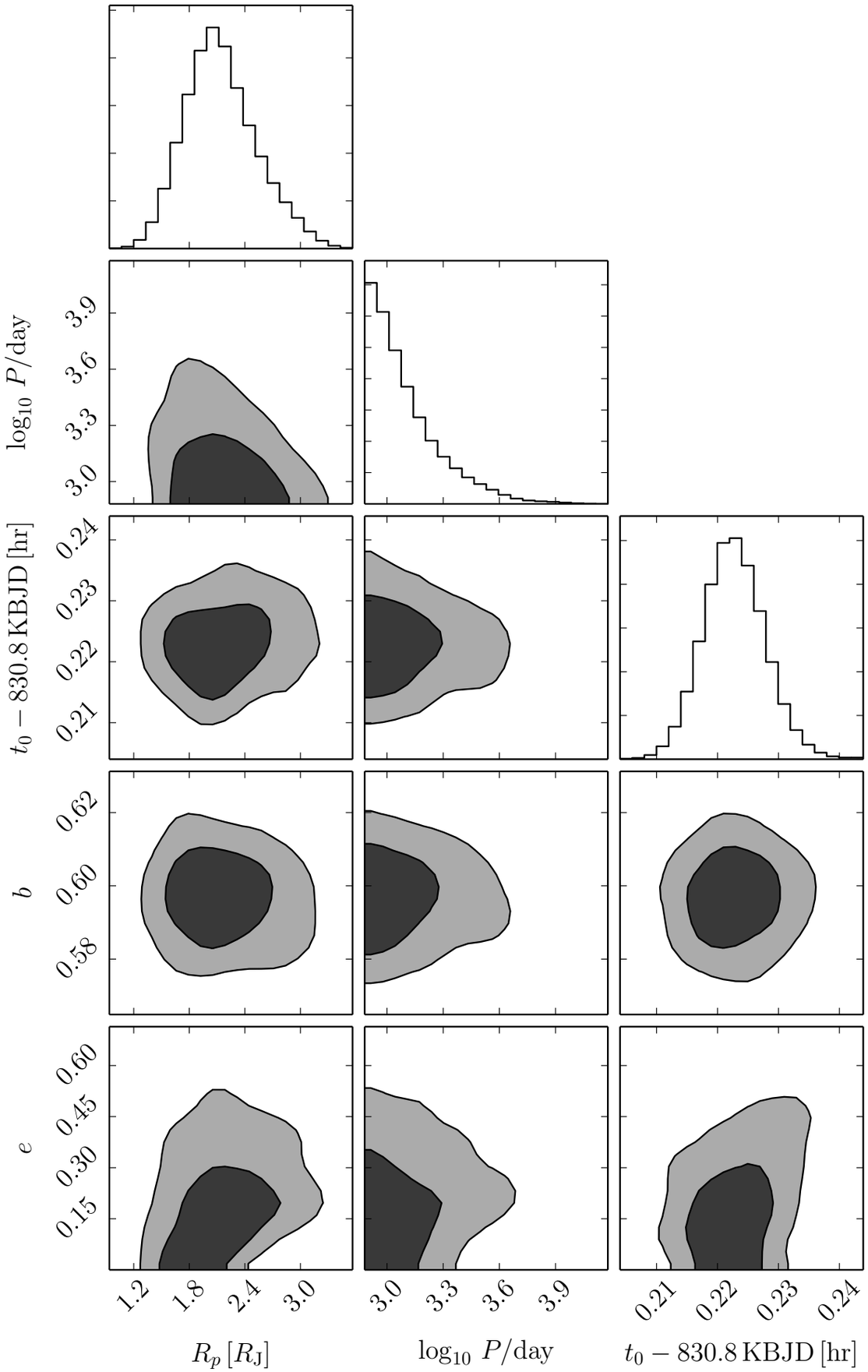
3,000 light curves

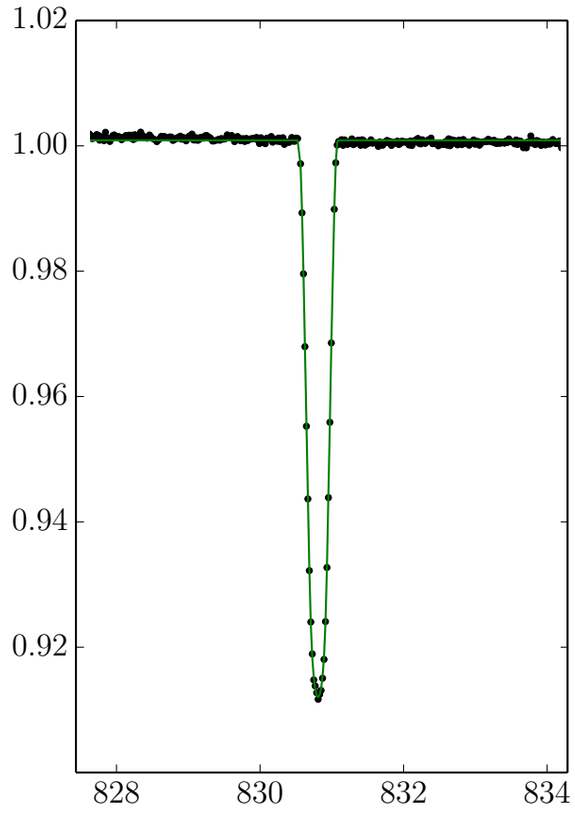
273 false positives

1 transit candidate

False Positives

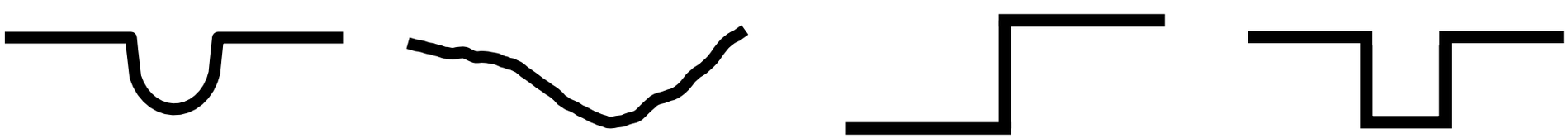



$R_p\,[R_\mathrm{J}]$
$\log_{10}\,P/\mathrm{day}$
$t_0 - 830.8\,\mathrm{KBJD}\,[\mathrm{hr}]$
b
e


1.02
1.00
0.98
0.96
0.94
0.92
828
830
832
834

No good model of the non-transits…

Temporary solution: Template likelihoods



Conclusions

1. can discover single transits using supervised classification
2. false positives are still a problem (but maybe less)
3. would like to combine method with realistic noise model