VESPA WORKSHOP GROUP PROJECTS

Andrew Vanderburg, Tim Morton, and Juliette Becker

1 Adding observations:

Goal: Learn how the inclusion of different observational constraints affects the calculated false positive probability for a given planet candidate. Experiment with including different constraints and determine how the false positive probability changes.

Once a planet candidate is identified in data from a transit survey, follow-up observations and archival observations from other (often ground-based) telescopes are a crucial step in the process of confirming the existence of the candidate. This group activity is aimed at quantifying how each of these supplemental observations contribute to eliminating false positive scenarios.

This case study will focus on the planetary system orbiting GJ 9827. This red dwarf star, located about 30 parsecs from Earth, was observed in 2017 by *Kepler* during Campaign 12 of its K2 mission. Analysis of the K2 data showed the signals from three transiting planet candidates, and two teams independently began studying the system and collecting follow-up observations to validate the planetary candidates (Niraula et al., 2017; Rodriguez et al., 2018).

Between the two teams, a wealth of archival and follow-up data were amassed, including:

- The K2 light curve, publicly available at `https://www.cfa.harvard.edu/~avanderb/k2c12/ep246389858.html`
- An archival spectrum from the CfA Digital Speedometer spectrograph at the 1.5m Wyeth reflector at Oak Ridge Observatory in the town of Harvard, MA (Rodriguez et al., 2018)
- An archival spectrum from the CORAVEL-type spectrograph on the 165 cm Richey-Chrétien telescope at Vilnius University Observatory (Sperauskas et al., 2016)
- An archival spectrum from the HARPS spectrograph on the 3.6m ESO telescope at La Silla, Chile (Houdebine et al., 2016)
- Archival images from the Palomar Observatory Sky Survey (POSS) taken in 1953, available at `https://archive.stsci.edu/cgi-bin/dss_form?target=gj+9827&resolver=SIMBAD`
- An adaptive optics image from Keck observatory collected after the identification of transits around GJ 9827 (Rodriguez et al., 2018), available online at `https://exofop.ipac.caltech.edu/k2/edit_target.php?id=246389858`.
- Seven spectra from the FIES spectrograph on the Nordic Optical Telescope on La Palma in the Canary Islands

Some additional observations are available online at ExoFOP-K2: `https://exofop.ipac.caltech.edu/k2/edit_target.php?id=246389858`.

The goal of this case study will be to determine which supplementary observations and constraints (in addition to the K2 light curve) are most important for ruling out false positive scenarios. In particular, the supplementary observations allow constraints to be imposed on `vespa`'s FPP calculation include:

- Stellar parameters (effective temperature, surface gravity, metallicity) from spectroscopy
- Limits on the presence of nearby companions from adaptive optics imaging (in the form of a "contrast curve") available on ExoFOP-K2. These constraints can be incorporated by including a contrast curve constraint in `fpp.ini`.
- Limits on the masses of the putative transiting objects from a lack of large radial velocity variations in the archival Digital Speedometer, Vilnius, and FIES spectra. This can be implemented by setting the probability of the Eclipsing Binary (EB) and EB at twice the period scenarios to zero and re-calculating the FPP with only the BEB and HEB scenarios.
- Limits on the presence of background stars at the present day position of GJ 9827 based on archival "patient imaging". Since GJ 9827 has high proper motion, its apparent position in the sky has changed by nearly 30 arcseconds since it was imaged by POSS. This can be used to rule out background stars down the limiting magnitude of the POSS survey. (See Figure 2 of Rodriguez et al. 2018).
- Because GJ 9827 has three transiting planet signals, the *a priori* likelihood that each of them are indeed genuine exoplanets is considerably higher than if there was only one transiting signal around GJ 9827.

Experiment with multiple `vespa` runs for GJ 9827 while including or excluding various supplementary observations and constraints. Which constraints are the most important? Which ones are superfluous? Keep in mind that not all planetary systems will benefit equally from the same follow-up observations. Can you imagine a situation in which a different set of observations are helpful than those most useful for validating GJ 9827?

2 Unvalidated planets Part 1:

Goal: Learn about tricky scenarios which can lead to mis-validation of false positive planet candidates. Investigate how the photometric aperture used to extract a light curve can change the appearance of a candidate signal and the false positive probability inferred.

Searching for and validating planets is hard and there are lots of pitfalls which can cause false positives to appear to be very planet-like. In this case study, we will focus on ways to identify and avoid cases where false positive signals have been validated. In particular, we will focus on the cases of K2-78b, K2-82b, and K2-92b. These three objects were identified and statistically validated by Crossfield et al. (2016), who used `vespa` to calculate false positive probabilities and found FPPs below 1%, which was deemed sufficient to call these objects validated planets. Subsequently, Cabrera et al. (2017) used pixel-level diagnostics to identify these three objects as false positives.

The first goal of this case study should be to replicate the false positive probability calculation performed by Crossfield et al. (2016). This can be done by downloading the data used by Crossfield et al. (2016) from ExoFOP-K2 from the following pages:

- K2-92b: `https://exofop.ipac.caltech.edu/k2/edit_target.php?id=211152484`
- K2-78b: `https://exofop.ipac.caltech.edu/k2/edit_target.php?id=210400751`
- K2-82b: `https://exofop.ipac.caltech.edu/k2/edit_target.php?id=210483889`

From these pages, download the light curves used by Crossfield et al. (2016) (labeled "Photometry from k2phot v1.0"), high-resolution imaging contrast curves, retrieve spectroscopic parameters from Crossfield et al. (2016) (available at `https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/apjsaa33cet7_mrt.txt`), calculate `maxrad` from the aperture size shown in photometry diagnostics uploaded to ExoFOP-K2 (labeled "Photometry diagnostic plots from k2phot v1.0"), and include all of the photometric bands in the EPIC catalog as measurements in `star.ini`. Do the calculated false positive probabilities match the literature FPPs from Crossfield et al. (2016)?

These three signals were incorrectly validated because the scenarios presented to `vespa` violate one of its primary assumptions – that the transit signal has been localized to within a radius of `maxrad` of the target star. One way to determine where the transit signals are coming from, which is highly effective for K2 data in particular, is to extract light curves from different regions of the *Kepler* images. Download the following *.ps* files and examine the plots in the right-hand column of the second page:

- K2-92b: `https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/ep211152484.01.ps`
- K2-78b: `https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/ep210400751.01.ps`
- K2-82b: `https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/ep210483889.01.ps`

These plots show how the planet candidate transit signals change when the light curves are extracted from different photometric apertures. It is important to know that the red transit model plotted over each light curve shows transits of the same depth (the model is the best-fit for the default aperture shown in the left-hand column). In the plots for both EPIC 211152484 and EPIC 210400751, the depth of the transit signals increase in the largest photometric apertures – an indication that the transit signal is likely not originating from the target star. The case of EPIC 210483889 is more difficult to interpret because there is a significantly brighter star just to the north of the target, which contaminates the light curve of the target star. However, in this case, the flux from the nearby brighter star dominates the flux from the fainter target, so any signals detected on this star are suspect.

You can download K2 light curves from these different photometric apertures from the following links:

- `https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/ep210400751-9pixaperture.csv`
- `https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/ep210400751-156pixaperture.csv`
- `https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/ep211152484-9pixaperture.csv`
- `https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/ep211152484-195pixaperture.csv`

For EPIC 211152484 and EPIC 210400751, re-calculate the FPPs for these three planet candidates using smaller photometric apertures (9 pixels) than used by Crossfield et al. (2016) which exclude the contaminating signals. Also, recalculate them with larger photometric apertures, which make the v-shaped nature of the signals more obvious. Does `vespa` correctly identify the false positives in these larger apertures?

3 Unvalidated planets Part 2:

Goal: Learn about tricky scenarios where eclipsing binary stars can be mistakenly validated as exoplanets. Investigate how underestimated uncertainties in the inputs to `vespa` can affect the calculated false positive probabilities.

Like Group Project 2, this case study will focus on signals which were validated as planets, but which later were shown to be false positives. However, the mechanism by which these candidates were mistakenly assigned the label of validated planet is different than the mechanism discussed in the previous case study. In particular, we will focus on three other planet candidates which were initially detected and validated by Crossfield et al. (2016): K2-51b, K2-67b, and K2-76b. After their validation, Shporer et al. (2017) performed radial velocity observations of the host stars and found that the masses of the companions were quite large – too large to be transiting planets. All three objects have masses greater than the brown dwarf hydrogen-burning limit at 0.08 $M_{\odot}$, and therefore are all stars, not planets.

First, replicate the `vespa` analysis done by Crossfield et al. (2016). You can download the light curves used by Crossfield et al. (2016) and the contrast curves from adaptive optics images from ExoFOP-K2:

- K2-51b: `https://exofop.ipac.caltech.edu/k2/edit_target.php?id=202900527`
- K2-67b: `https://exofop.ipac.caltech.edu/k2/edit_target.php?id=206155547`
- K2-76b: `https://exofop.ipac.caltech.edu/k2/edit_target.php?id=206432863`

Read in the column of the *.csv* files labeled `fdt_t_roll_2D` for a light curve with low-frequency variability removed.

The stellar parameters used by Crossfield et al. (2016) are available in their Table 7 (available at `https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/apjsaa33cet7_mrt.txt`), and photometric measurements from the EPIC catalog are available at `http://archive.stsci.edu/k2/epic/search.php`. How well do the results of this `vespa` analysis match what Crossfield et al. (2016) calculated?

Shporer et al. (2017) revised the parameters of the host stars of these three objects, and found that the sizes of the primary stars K2-51 and K2-67 were underestimated by Crossfield et al. (2016). Replace the stellar parameters for these two objects with the revised stellar parameters from Shporer et al. (2017), and re-run the `vespa` analysis. Does changing the stellar parameters cause `vespa` to return a different FPP?

Shporer et al. (2017) also identified erroneous measurements from broad-band photometric surveys as a potential cause of `vespa`'s overconfidence in the planetary nature of these three objects. While the *Kepler* Input Catalog (KIC) was the result of a homogenous set of observations, the Ecliptic Plane Input Catalog (EPIC) used for K2 was compiled together from many different archival photometric surveys. This causes the potential for unexpected offsets and errors between measurements from different surveys. One way around this problem is to only use photometry from a single survey from the EPIC, such as the 2MASS survey.

Repeat the `vespa` analysis, but instead of including broad-band photometric measurements from all bands included in the EPIC, only use measurements from the 2MASS survey, namely J, H, and K-band. Does running `vespa` with these inputs give a significantly different answer than before? Also, repeat the `vespa` analysis from before, but artificially inflate the uncertainties on the broadband photometry by a factor of 3. Does this lead to a significantly different FPP?

An in-depth analysis describing what led to the mis-validations of these three targets is available online: `http://nbviewer.jupyter.org/github/timothydmorton/VESPA/blob/cautions/notebooks/shporer-analysis.ipynb`.

4 Additional Challenges/Investigations:

4.1 Ephemeris Matching:

In general, `vespa` only considers false positive scenarios from stars other than the target within a small radius (`maxrad`) of the target star. However, in rare cases, there can be signals introduced into the light curves of stars from large distances away in *Kepler* data. Coughlin et al. (2014) have identified several mechanisms which can cause signals from stars in *Kepler* data to contaminate stars well outside the photometric aperture (with size `maxrad`) used. In particular, Coughlin et al. (2014) identified the following mechanisms:

1. Direct PRF contamination: scattered light from the point spread function of a nearby star which overlaps the position of the contaminated star.

2. Column Anomaly or Charge Transfer Inefficiency: Signal can couple between stars on the same CCD column and be transferred from a “parent” star, the source of the signal, to a “child” star, the contaminated target.

3. Electronic Crosstalk: Signals can be electromagnetically be coupled between wires connecting different CCDs on the *Kepler* array.

4. Antipodal Reflection: An optical ghost image from a star on the other side of *Kepler*’s focal plane can reflect onto the target star and contribute a signal.

Coughlin et al. (2014) identified a set of nearly 700 false positive signals in *Kepler* data due to contamination from one of those four mechanisms.

This investigation will be to see what happens when `vespa` encounters one of these unusual false positives. While these distant false positives violate one of the key assumptions of `vespa` (namely that the contaminating star is within `maxrad` of the target star), `vespa` can often identify them as false positives by the shape of the transit signals.

You can download the light curve for KOI 2600 at `https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/koi2600lc.csv`. The parameters are listed on CFOP: `https://exofop.ipac.caltech.edu/kepler/edit_target.php?id=2600`. This particular false positive is an example of contamination from column anomaly.

Often, column anomaly worsens over the course of the *Kepler* mission, and the strength of signals caused by this phenomenon grows over time. An example of this was investigated by Gaidos et al. (2016) around the star KOI 6705, who found that a candidate transit signal grew in depth over the course of the mission and concluded that it was caused by column anomaly.

Download the light curve for KOI 6705 from `https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/koi6705lc.csv`. How does the transit signal change between the first and second half of the mission?

4.2 Circumstances beyond your control:

No matter what follow-up observations you manage to collect, some planetary systems are just plain hard to validate and confirm. This is because the *a priori* false positive probabilities of some candidates is higher than others, depending on the properties of the planet candidate and host star. These priors are beyond the control of follow-up observers, but are important and worth understanding. Here, we will focus on how some properties of the planet candidate host star can make validation easier or harder.

In particular, this group project investigation will focus on the impacts of parameters like:

- The host star's *Kepler* or TESS magnitude. Is it harder to validate a planet candidate around a bright star or a faint star? Equivalently, are there more potential false positive scenarios for a planet candidate around a bright star or a faint star?
- The host star's position in the sky (importantly, its galactic latitude). How does the distance from the galactic planet, a proxy for the density of stars in the background, affect the calculated FPP?
- The host star's stellar type (dwarf vs. giant), or in particular, the star's surface gravity. At a given transit depth, how does the FPP depend on whether the star is a giant or a dwarf?
- The precision to which the source of the transits can be isolated in RA/Dec. Some transit surveys are able to localize the position of transit signals to high precision, while it is more difficult for others. For example, transit signals from *Kepler*, with its ultra-stable pointing, relatively small pixels, and long observing baselines/high signal-to-noise transit signals are often localized to better than one arcsecond. On the other hand, the increased pointing jitter in K2 makes it difficult to localize signals this precisely. It will be even more difficult to localize transit signals from TESS, which has pixels 5 times larger than *Kepler*. How does the localization of the transit signal, as described by the `maxrad` keyword in `fpp.ini` affect the FPP?
- The precision with which secondary eclipses can be excluded in the light curve. For weak signals, it is often difficult to rule out even fairly significant secondary eclipses in the light curve. How much does the calculated false positive probability depend on this parameter?

Take a generic planet candidate (for example, EPIC 248463350, the star focused on in the first session), and run `vespa` while tweaking these parameters in the `fpp.ini` file and `star.ini` file. Investigate the FPP as a function of these parameters and figure out which of these parameters have the greatest influence.

4.3 Exotic false positive scenarios

The heart of `vespa` is a Bayesian model comparison, where the likelihoods of different scenarios (exoplanet, eclipsing binary, background eclipsing binary, etc), are calculated and compared to one another. A planet is said to be validated when the planet model is favored over all others considered with some highly confident statistical threshold (often 99% to 99.9%).

However, there are exotic false positive scenarios which do not fall into these broad categories, and which are not considered by `vespa`. These scenarios, therefore, can trip up `vespa`, and lead to inaccurate false positive probabilities. Some examples of such scenarios include:

1. White dwarf eclipsing binaries. White dwarfs are about the same size as small planets (ranging in size from a bit smaller than the Earth to super-Earths). Therefore, the primary eclipses of white dwarf eclipsing binaries can sometimes perfectly mimic the transit of a super-Earth (see Figure 7 of Muirhead et al. 2013). The secondary eclipses of white dwarf eclipsing binaries can also be mistaken for transiting planets (see Figure 2 of Kruse & Agol 2014). Since `vespa` does not explicitly account for this scenario, it could erroneously validate these objects as planets.

2. Some exoplanets are known to be in the process of disintegrating (Rappaport et al., 2012). These exoplanets show asymmetric transit light curves with longer durations than for solid-body exoplanets, due to a cloud of ablated material orbiting near the planet. Since `vespa` does not explicitly consider the scenario of disintegrating planets, it can yield erroneous FPPs for objects like KIC 12557548 (Rappaport et al., 2012), KOI 2700 (Rappaport et al., 2014), K2-22 (Sanchis-Ojeda et al., 2015), and WD 1145+017 (Vanderburg et al., 2015).

3. Sometimes, planet candidates can arise from instrumental artifacts, as opposed to astrophysical phenomena. While `vespa` has a rudimentary ability to model some of these scenarios (via scenarios labeled "long" and "boxy", which can be switched on in the FPP calculation) , in many cases, instrumental signals are not well represented among the models considered by `vespa`. Recently, the *Kepler* team has identified a set of known instrumental signals, by searching for upwards "transits" (brightenings of the star) in *Kepler* light curves. One particularly vexing signal towards the star KIC 11961208 was pointed out by Mullally et al. (2018) to caution observers from mistakenly validating instrumental false positives. This signal has a period of 425.39839 days, with the first transit taking place at BJD - 2454833 = 225.70856. The inverted light curve can be found at `https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/kic11961208lc.csv`.

Investigate what happens in some of these unusual cases, where `vespa` does not explicitly consider some of the more exotic false positive scenarios.

4.4 Bad detrending on the light curve:

A potential source of error in the FPP calculation process is that a given light curve may not had long-term signals effectively removed. Before `vespa` analyzes a light curve, these low-frequency signals must be filtered away either using Fourier methods, polynomial or spline fitting, or use of a Gaussian Process. What happens, though, when this process is not executed perfectly. Poor detrending can both distort the transit signal (if the flattening method is too aggressive), or leave artifacts in the baseline flux (if it is not aggressive enough).

Test what happens to the FPP calculation when detrending is too aggressive or not aggressive enough. Use the example given in the `vespa` tutorial covered on the first day of the workshop: `https://github.com/timothydmorton/vespa-tutorial/blob/master/Notebooks/VESPA_Tutorial.ipynb`. The following line of code in Appendix 1 controls the aggression of the detrending filter:

```
trend = scipy.signal.savgol_filter(epic_target['flux'],101,polyorder=3)
```

Changing the second argument of the function, called the "window length", (given in the example as 101) changes the aggression of the filter. Increasing the number decreases the aggression. Experiment with different values for the window length and choose two values: one significantly smaller than 101 and one significantly larger than 101, and calculate the FPP of the planet candidate using these poorly detrended light curves. How do the calculated FPPs change?

5 Download links for referenced papers:

- Cabrera et al. (2017): https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/aa31233-17.pdf
- Coughlin et al. (2014): https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/Coughlin_2014_AJ_147_119.pdf
- Crossfield et al. (2016): https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/Crossfield_2016_ApJS_226_7.pdf
- Gaidos et al. (2016): https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/Eric_Gaidos_2016_ApJ_817_50.pdf
- Houdebine et al. (2016): https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/Houdebine_2016_ApJ_822_97.pdf
- Kruse & Agol (2014): https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/275.full.pdf and supplementary material at https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/Kruse.SM.pdf
- Muirhead et al. (2013): https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/Muirhead_2013_ApJ_767_111.pdf
- Mullally et al. (2018): https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/Mullally_2018_AJ_155_210.pdf
- Niraula et al. (2017): https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/Niraula_2017_AJ_154_266.pdf
- Rappaport et al. (2012): https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/Rappaport_2012_ApJ_752_1.pdf
- Rappaport et al. (2014): https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/Rappaport_2014_ApJ_784_40.pdf
- Rodriguez et al. (2018): https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/Rodriguez_2018_AJ_155_72.pdf
- Sanchis-Ojeda et al. (2015): https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/Sanchis-Ojeda_2015_ApJ_812_112.pdf
- Shporer et al. (2017): https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/Shporer_2017_ApJL_847_L18.pdf
- Sperauskas et al. (2016): https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/aa27850-15.pdf
- Vanderburg et al. (2015): https://www.cfa.harvard.edu/~avanderb/saganworkshop2018/wd1145_017.pdf

References

Cabrera, J., Barros, S. C. C., Armstrong, D., et al. 2017, A&A, 606, A75

Coughlin, J. L., Thompson, S. E., Bryson, S. T., et al. 2014, AJ, 147, 119

Crossfield, I. J. M., Ciardi, D. R., Petigura, E. A., et al. 2016, ApJS, 226, 7

Gaidos, E., Mann, A. W., & Ansdell, M. 2016, ApJ, 817, 50

Houdebine, E. R., Mullan, D. J., Paletou, F., & Gebran, M. 2016, ApJ, 822, 97

Kruse, E., & Agol, E. 2014, Science, 344, 275

Muirhead, P. S., Vanderburg, A., Shporer, A., et al. 2013, ApJ, 767, 111

Mullally, F., Thompson, S. E., Coughlin, J. L., Burke, C. J., & Rowe, J. F. 2018, AJ, 155, 210

Niraula, P., Redfield, S., Dai, F., et al. 2017, AJ, 154, 266

Rappaport, S., Barclay, T., DeVore, J., et al. 2014, ApJ, 784, 40

Rappaport, S., Levine, A., Chiang, E., et al. 2012, ApJ, 752, 1

Rodriguez, J. E., Vanderburg, A., Eastman, J. D., et al. 2018, AJ, 155, 72

Sanchis-Ojeda, R., Rappaport, S., Pallè, E., et al. 2015, ApJ, 812, 112

Shporer, A., Zhou, G., Vanderburg, A., et al. 2017, ApJ, 847, L18

Sperauskas, J., Bartašiūtė, S., Boyle, R. P., et al. 2016, A&A, 596, A116

Vanderburg, A., Johnson, J. A., Rappaport, S., et al. 2015, Nature, 526, 546