In search of Earth analogues

Detecting exoplanets amid stellar noise

The radial velocity method of exoplanet detection has led to some fantastic discoveries of planets outside our solar system. But the hunt for true Earth analogues is hampered by noisy data from stars. Jessi Cisewski outlines the challenge, and how statisticians and statistically savvy astronomers are trying to help.
It started with a teaser: the promise of a discovery "beyond our solar system". Forty-eight hours later, NASA announced that it had found seven Earth-sized planets orbiting a single star, 235 trillion miles away. The news was met with excitement around the world, as was the discovery six months before of Proxima Centauri b, a planet a little more massive than Earth, orbiting the star Proxima Centauri, some 25 trillion miles from home.

To date, more than 3000 extrasolar planets – or "exoplanets" – have been detected. But these recent discoveries are special: they have been found orbiting within the so-called Goldilocks zone, where a planet is at a sufficient distance from its host star for its surface to be neither too hot nor too cold to support liquid water. This raises the prospect that the planets may be home to organic life. Three of the seven TRAPPIST-1 planets are firmly within this habitable zone, says NASA – but further observations will be required to confirm whether any are rich in water. Alas, recent research suggests Proxima Centauri b may be a desert world, having been scorched by its star during an earlier, brighter phase (bit.ly/2k6PSHP).

The discovery of these planets is an impressive feat of modern science, but they relied on two very different methods. The TRAPPIST-1 system was detected using the transit method, whereby planets crossing the face of a star cause visible dips in the amount of stellar light (see "A different approach", page 25). For Proxima Centauri b, astronomers inferred its presence from the slight gravitational pull it exerts on its host star, which causes small shifts in the measured spectrum of stellar light because of the Doppler effect, and it is this method of detection – known as the radial velocity (RV) method – that is the subject of this article.

The RV method has led to many exoplanet discoveries, but there are a number of challenges that need to be overcome to detect true Earth analogues: that is, exoplanets that are similar to Earth, orbiting stars that are similar to our Sun. Indeed, one could argue that the discovery of Proxima Centauri b was relatively easy in comparison: that planet has an estimated minimum mass of about 1.27 times that of Earth, while the Proxima Centauri star is less massive than our Sun, and the planet orbits at a much closer distance than does Earth – all of which results in a more pronounced RV signal.

To detect smaller RV effects of the type created by Earth-like planets orbiting a solar-type star, astronomers will need more precise instrumentation. But at ever smaller scales, stellar noise becomes a bigger problem – which is where statisticians can lend a hand.

Back and forth

The RV method of exoplanet detection relies on a time series of stellar spectra. A stellar spectrum measures the amount or intensity of light across a wavelength range (the range depends on the spectrograph used). An example of a stellar spectrum, from the European Southern Observatory’s HARPS spectrograph, is displayed in Figure 1 (page 24). The parabola shape of the spectrum is due to the spectrograph, not the distribution of light coming from the star, while the various dips
in the spectrum are absorption lines, which are caused by the light passing through the outer layers of the star.

Now, let us imagine a hypothetical exoplanet system such as the one in Figure 2. As the exoplanet orbits its host star, the star will move around the centre of mass of the system. The best-case scenario for an astronomer, as an observer, is if the orbital plane of the exoplanet is perpendicular to the plane of the sky; this results in the maximum RV measurement. The RV is the forward and backward velocity of the star with respect to the observer.

This motion of the star means that, depending on the orbital phase, the light released will be from a source that is either moving towards or away from the observer, resulting in a shift of the stellar spectrum due to the Doppler effect. When the star is moving towards the observer, the spectrum is blueshifted, which means that absorption lines are shifted to shorter wavelengths, and when the star is moving away from the observer the spectrum is redshifted to longer wavelengths. Given a time series of spectra for a star, the astronomer seeks to detect the change in forward and backward motion to determine if an object is orbiting.

Once a collection of spectra is observed at different times, the RV for each time needs to be estimated. To infer the RV of a star at a particular time using an observed spectrum, one technique that astronomers employ is the cross-correlation function. Here, the observed spectrum is cross-correlated with a template spectrum that is similar to the observed star’s spectrum. The template is Doppler-shifted by a series of RVs, and then cross-correlated with the observed spectrum; the Doppler shift of the shifted template that correlates most strongly with the observed spectrum is used. The result is an estimate of the forward or backward velocity of the star due to the gravitational effect of an unseen object.

One of the current cutting-edge spectrographs is the aforementioned HARPS, which was engineered to have a sensitivity of the order of 1 metre per second (which is the lower bound on the detectable RV). HARPS drew a lot of public attention when it detected an RV signal of 1.38 m/s for the star Proxima Centauri, which led to the discovery of Proxima Centauri b. In 2011, the HARPS team measured RV signals between 0.5 and 1 m/s around stars HD 20794, HD 85512, and HD 192310. Though impressive, this level of sensitivity is still not enough to detect Earth analogues. The estimated forward and backward velocity of the Sun due to the gravitational effect of the Earth is only 10 cm/s – a pace at which it would take a person over 16 minutes to cover 100 m.

Stellar noise and statistics
Next-generation spectrographs are expected to get below the 1 m/s threshold, possibly down to the 10 cm/s level. If achieved, this would truly be a marvel of engineering. But solving this one problem creates others – problems to which statisticians may be able to contribute solutions.

Even for current spectrographs with a precision of 1 m/s, noise created by stellar activity becomes an issue. The type and degree of stellar activity depends on several factors such as the sunspot cycle and other solar phenomena. These can mimic Doppler shifts in the RV curve, making them a major roadblock in finding Earth analogues. This image was created by Kelvin Ma and is displayed in its original form under the Creative Commons Attribution-Share Alike 3.0 Unported license (bit.ly/2IS5Xbn).
as the age of the star, how fast it rotates, and its temperature (to name but a few). Some of the types of stellar activity are: convection, faculae, flares, granulation, magnetic cycles, oscillations, and spots.

Granulation, for example, appears to segment the star into a bunch of cells of rising hot fluid that cools and drops back down towards the interior of the star (see top left of Figure 3). These stellar granules last for a short period and then disappear as other granules form. The problem is that the rising hot fluid is brighter than the cooler, falling edges. Overall, granulation results in an apparent net blueshift of the order of metres per second.

Furthermore, as a star rotates, half the star’s light is moving towards the observer resulting in a blueshift, and the other half of the star’s light is moving away from the observer resulting in a redshift. If a spot is present, for example, this dark and cooler region may lead to a reduction in the blueshifted or redshifted light, which can result in a planet-like signal in the RV curve.

These various sources of stellar activity can result in time-dependent, non-uniform, asymmetric effects on parts of the stellar spectra, and therefore need to be accounted for when searching the RV curve for exoplanets. Various statistical methods have been proposed to address this.

Keeping pace
In a recent paper by astronomer Xavier Dumusque, an RV fitting challenge was presented to the astronomy community to assess the performance of various RV modelling methods of different planetary signals in the presence of stellar activity. Using realistic stellar spectra (with activity and instrumental effects), 48 planetary signals were injected into 15 systems, with RV semi-amplitudes of the planets ranging from 0.16 to 5.85 m/s; three of the systems did not have any planets.

Eight teams participated in the challenge, each developing their own method of detection which can be roughly categorised as either Bayesian or classical. Five of the eight teams took Bayesian approaches, of which two used Gaussian process regression to account for the stellar activity before fitting multiple models (with differing numbers of planets or planet properties), and then applying some flavour of Bayesian model comparison. Of the three classical approaches, one example involved the use of a discrete Fourier transform (DFT) to find a dominant signal, then modelling the signal with least-squares regression, and looking at a DFT of the residuals to see if other signals were present; this process was repeated until all signals seemed to have been removed (see Dumusque et al. for details).

The performance of the proposed methods was evaluated in several ways: the distribution of correct detections (scaled by level of confidence in the detection); correct detections but with incorrect estimated semi-amplitudes or periods; false positives (detecting a planet that was not present in the system); and false negatives (missing a planet that was present). Not all teams analysed all the data, which made formal comparisons difficult. However, when focusing on correct detections, it was found that groups which used Bayesian methods tended to make more correct detections than those using other methods. More data are required to fully assess the performance of the proposed methods, but this RV challenge was a step in the right direction.

As the engineering and design of spectrographs continues to improve, the question remains whether statistical methods can keep up. Yale astronomer Debra Fischer and her team are building a spectrograph called EXPRES that will surpass the 1 m/s barrier, moving it down near the 10 cm/s level. It is expected to be operating in autumn 2017. While EXPRES is under construction, there is a flurry of activity to get statistical methods ready for improved data. The teams who participated in the RV fitting challenge, along with astronomers and statisticians at Yale, Penn State, and Geneva, are among those working to develop statistical methods to address issues with stellar activity for the RV method. Additionally, the Statistical and Applied Mathematical Sciences Institute in Research Triangle Park, North Carolina has a working group that is considering these problems as a part of its 2016–2017 Program on Statistical, Mathematical and Computational Methods for Astronomy. Meanwhile, the Third Annual Workshop on Extreme Precision Radial Velocities will be held in August 2017 at Penn State. The first two were at Penn State and Yale, respectively; the conferences focus on all aspects of the RV method – from instrumentation to inference.

These collective efforts create an expectation that the challenges will be met and that statistical methods will ably support astronomers in their continuing search for exoplanets. Proxima Centauri b may not be as Earth-like on the surface as it is in size; the TRAPPIST-1 planets may or may not be all we hope them to be. But in the vastness of space, there remain plenty of opportunities to find Earth analogues among the stars.

References
1. Dumusque, X. (2016) Radial velocity fitting challenge. I. Simulating the data set including realistic stellar radial-velocity signals. Astronomy & Astrophysics, 593, A5.
2. Dumusque, X., Borsa, F., Damasso, M., Diaz, R. and Gregory, P. C. (2017) The radial velocity fitting challenge. II. First results of the analysis of the data set. Astronomy & Astrophysics, 598, A113.
3. Fischer, D. A., Anglada-Escude, G., Arriagada, P. et al. (2017). State of the field: Extreme precision radial velocities. Publications of the Astronomical Society of the Pacific 128(964), 066001.