Utah State University

From the SelectedWorks of Mario Yuuji Harper

April 2, 2013

Simulating Exoplanet Detection of Kepler via Monte-Carlo Techniques

Manuel Marcano, Utah State University
Mario Yuuji Harper, Utah State University

Available at: https://works.bepress.com/mario_harper/5/
Simulating Exoplanet Detection of Kepler via Monte-Carlo Techniques

Manuel Pichardo Marcano,1,a) Mario Harper,1,b) and Margaret Jensen1,c)

Department of Physics, Utah State University, Logan, UT 84322

(Dated: 2 April 2013)

The Kepler mission was launched in 2009 with the intent to find earth-like planets around other star systems. Since it’s launch, the number of known (and suspected) exoplanets have increased substantially from 200 to 3000. With this increase in data now available, we have chosen to model through a monte-carlo simulation, the possibility of being able to see exoplanets in a random part of the sky (using the parameters of the Kepler spacecraft). We show that there is a $0.18 \pm 0.01$ percent chance of being able to detect exoplanets.

Keywords: Exoplanet, Kepler Mission, Monte Carlo Method

I. INTRODUCTION

It is apparent from the monte-carlo simulations that, given a spacecraft of Kepler’s parameters, $0.18 \pm 0.01$ percent of the exoplanets that actually exist in kepler’s field of view are detectable. The Kepler mission boasts that $5 \times 10^5$ stars (give or take $10^4$) are in its field of view, this was the sample size in running the simulations. We also find that most planets will be undetectable because the planet does not transit between our telescope and it’s host star (we will define this as angle). Some planets are not large enough to cause a luminosity change significant enough for detection (defined as luminosity change). Other planets have too long of a orbitting period and cannot be confirmed if it is indeed an actual planet (we assume a mission deadline at five years).

The monte-carlo in this project was designed to accomplish two tasks. First, to generate a field of data points (exoplanet systems). Second, to show the percentage of detectable planets out of our generated field. We pulled data from the publically available NASA Exoplanet Archive (Joint with the California Institute of Technology), and the Hipparcos Catalogue of exoplanets (European Space Agency). The relevant data was pulled into histograms as can be seen in Fig. 1, and fitted to a distribution function. The monte carlo we used has six distribution functions that determined type and detectability.

It is of some interest to note that the types of detectable planets, by nature of the technological and time constraints, are very similar. Gas giants orbitting close to the sun are the easiest to detect due to their relative mass and radius. The Kepler mission was commisioned to find earth-like planets, which we show, is a very difficult thing to do.

II. PROCEDURE

The monte carlo simulation was written in Mathematica. The first step was to import the data to the catalog. For this project two different databases were used to get the different planetary and stellar parameters (respectively) which were necessary to do the monte carlo simulation. The parameters that we were interested in was the inclination angle, stellar radius, stellar effective temperature, distance, planetary radius and the semi-major axis. For a detailed description of the catalogs accessed see Sec II.A. The second step (Section II B) after importing the data was the generation of different distribution functions from the data. These fitted distribution functions were then used to generate random numbers to do the monte carlo simulation (Section II C).

A. Data Aquisition

This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program. The NASA Exoplanet Archive is an online astronomical exoplanet and stellar catalog and data service that collates and cross-correlates astronomical data and information on exoplanets and their host stars and provides tools to work with these data. The data include stellar parameters (such as positions, magnitudes, and temperatures), exoplanet parameters (such as masses and orbital parameters) and discovery/characterization data (such as published radial velocity curves, photometric light curves, images, and spectra).

We also used the Hipparcos Catalogue. The Hipparcos Catalogue is the primary products of the European Space Agency’s astrometric mission, Hipparcos. The satellite, which operated for four years, returned high quality scientific data from November 1989 to March 1993. This space-astrometry mission measured the parallax angle of 2.5 million stars. From the parallax we can get the distance to the stars, as explain in Appendix A. Further descriptions of the catalogue may be found in Ref. 2.

---

a)Electronic mail: manuel.pichardo_marcano@aggiemail.usu.edu
b)Electronic mail: mario.harper@aggiemail.com
c)Electronic mail: maggiejensen@aggiemail.usu.edu
B. Data Analysis

From the 6 parameters obtained in the catalogues (Inclination angle, Stellar Radius, Stellar Effective Temperature, Distance, Planetary Radius and the Orbit Semi-Major Axis) relevant information that determines the detectability of an exoplanet were obtained. A probability distribution function was then fitted to each parameter. From these 6 parameters, three of them (planet period, star luminosity and stellar mass) were calculated using Kepler’s law and the Mass-Luminosity relation for main sequence stars. For this step we used the Statistical Analysis package from Mathematica. To visualize the data we use the built-in function "Histogram". To get the distribution function that best fits the data we used the function "EstimatedDistribution". This build-in function in Mathematica uses the maximum likelihood method attempts to maximize the log-likelihood function \[ \sum_{i=1}^{n} \ln(f(x_i|\theta)) \], where \( \theta \) are the distribution parameters and \( f(x|\theta) \) is the Probability Density Function (PDF) of the symbolic distribution. With the fitted distribution we then would calculate the moment of the distribution to see how good was the fitted distribution found. Figure 1 show an example of the Probability Density Function of the fitted distribution plotted along with the PDF of the data set.

![Figure 1](image1.png)

FIG. 1. Histogram of the Probability Density Function (PDF) for the star distances (in parsecs) plotted along the PDF of the fitted Distribution. The fitted distribution (plotted in red) was the Weibull Distribution.

C. Monte Carlo Simulation

The Mathematica function fitting was tested against the Cramer-von Mises test and a Jaque-Bera test for goodness of fit (the Jaque-Bera test is for normal distributions). We found that many of the data parameters made histograms such as shown in 1, these tended to fit well to a Weibull Distribution which is designed to have high partial kurtosis and minimize skew. This test reflects the data as currently available, although evidences hint that if all planets around stars are visible, we might be able to expect normal distributions of planets, perhaps even similar to our solar system. From the distribution functions fitted to each of the parameters we needed to pick random numbers to do the Monte Carlo Simulation. For this step we use the "RandomVariate" Mathematica function.

The Monte-Carlo simulation began by assuming that there exists one planet to every star, and knowing that (as a high estimate) Kepler can potentially see \( 5 \times 10^5 \) stars, thus a star/planet count of \( 5 \times 10^5 \) makes the input of our test. The simulation was streamlined by testing the most stringent parameters first to narrow our detectable planets. First was inclination, followed by the radii, temperature, distance, semimajor axis, magnitude and the calculated values of planet period, star luminosity and stellar mass.

III. RESULTS

In the monte carlo simulation run with the estimated star field of kepler. This was iterated 100 times, giving us a total of \( 5 \times 10^7 \) simulated exoplanet systems. We chose this to show a potential 100 star field that kepler could look at to show the possible maximum number of detectable exoplanets. In the 100 runs of the simulation, the lowest number of detectable planets was 860 and the maximum was 1008. The mean value of the 100 runs was at 929.83 planets with an standard deviation of 30.6957, which allows for a fair range of variability.

![Figure 2](image2.png)

FIG. 2. This graph shows the mean result of detectable exoplanets found after 100 iterations. The lines show the individual results, we emphasize the apparent variability.

We can see from Fig. 3 that these planets start from 5 earth radii peaking at ten earth radii, fairly large planets. Also, these planets are clustered closer to their host stars, making it easier to detect transits. Thus, we show that there is a 0.18 \( \pm \) 0.01 percent chance of being able to detect exoplanets given our assumptions and parameters. The results are comparable to what the Kepler mission is currently seeing.
FIG. 3. Histogram of detectable exoplanet radii (measured in earth radii), generated from $5 \times 10^7$ simulated exoplanets. Minimum radius can be seen as 5 earth radii.

IV. IMPLICATIONS AND FUTURE WORK

Due to time constraints in the making of the experiment we had to make many assumptions to simplify some of the numerous calculations involve in the simulation. Future work can be done in correcting for some of the simplifying assumptions to get a more realistic model of what we expect future findings of the kepler mission to be. Some of the assumptions that we can change is the following: 1- We assumed that every planet would have at least one planet. We justified this assumption with a recent paper published in Nature where the authors concluded, using gravitational microlesing to detect planets, that every star of the Milky Way hosts (on average) one planet or more in an orbital-distance range of $0.5 - 10 \cdot$ AU. It appears that having planets around stars in our Galaxy is a rule rather than the exception, however, more than one planet around stars appear to be likely. 2- The second assumption made was pertaining to the relations of different parameters of a planet and to its host star. We assumed that the star and planet are not related to each other by size and mass, and both are independent. We know that this is not the case, as there are slight correlations among certain metallicity, mass and radius of stars and exoplanets.

We also found that most of the known extrasolar planets have been discovered using the radial velocity or transit methods. An important limitation of this project was that we only accounted for planets that could be detected using the transit method. Kepler is capable of detecting exoplanets using the radial velocity method, in fact, the first planet discovered via this method was HD 209458. This method confirms that a observed radial velocity variation is due to an orbiting companion and to some other form of stellar variability. The project can be extended to try to encompass the planets made detectable by using the two methods and even include other not less effective methods of detection like the gravity microlensing technique.

ACKNOWLEDGMENTS

In addition to those listed on the title page, many other people have contributed directly or indirectly through conversations and consultations. We would like to express our special thanks to Dr. Shane Larson from the Physics department at Utah State University. Without his guidance the termination of this project would not have been possible.

Appendix A: Getting distances from Parallax

For the experiment we had to get a distance distribution for the stars. This information was available online in the Hipparcos Catalog. We got the distances to the stars (in parsecs) from the stellar parallax of each star. The Hipparcos mission measured parallax angles in milli-arcseconds over a baseline of 1 AU. The definition of parsec is the distance $d$ at which 1 AU subtends an angle of 1. Since the parallax angle from our query is reported in milli-arcseconds, the distance in parsecs is given by:

$$d(\text{pc}) = 1000/\text{parallax}$$

Appendix B: Fractional decrease in the brightness of the star

First, we assume that the star is a luminous disk with luminosity:

$$L_\star = A_\star \sigma T^4$$

$$= \pi R_\star^2 \sigma T^4$$

When the planet is in front of it, the luminosity is reduced by the area of the planet. During this transit:

$$L_T = (A_\star - A_p) \sigma T^4$$

$$= \pi \sigma T^4(R_\star^2 - R_p^2)$$

The luminosity during transit compared to the star by itself is:

$$\frac{L_T}{L_\star} = \frac{\pi \sigma T^4(R_\star^2 - R_p^2)}{\pi \sigma T^4 R_\star^2} = \frac{R_\star^2 - R_p^2}{R_\star^2} = 1 - \left(\frac{R_p}{R_\star}\right)^2$$

Appendix C: Minimum inclination for eclipse

We assume that the star and planet are in circular orbits about a mutual center of mass and are separated
by a distance $a$. We assume that the angle of inclination (the angle between your line of sight and the angular momentum vector of the orbit) is $i$ and the stellar radii are $r_1$ and $r_2$.

The observed separation of the two bodies is $A = S + r_1 + r_2$. We get an eclipse when $S = 0$. For a given inclination angle $i$ the observed separation between the star and planet is:

$$A = s + r_1 + r_2 = a \cdot \sin\left(\frac{\pi}{2} - i\right) = a \cdot \cos i$$

When $S = 0$ we get:

$$A = r_1 + r_2 = a \cos i$$

$$\cos i = \frac{r_1 + r_2}{a}$$


