

Automatic Classification of Kepler Planetary Transit Candidates Using Multilayered Neural Networks

Prabath Peiris

Capstone Project Proposal

Machine Learning Nanodegree

Udacity

Contents

1	Domain Background	3
1.1	Scientific Goals and Objectives of the Kepler Mission	4
1.2	Habitable Zone	5
1.3	The Transit Method of Detecting Extrasolar Planets	5
1.4	Project Proporsal and Motivation	7
2	Problem Statement	9
3	Datasets and Inputs	10
3.1	Kepler Pipeline	10
3.2	KOI and TCE Sttributes	12
3.3	Some Important Attributes	13
3.4	Missing Attributes	13
3.5	Classification Labels	14
4	Solution Statement	16
5	Benchmark Model	17
6	Evaluation Metrics	18

Chapter 1

Domain Background

Kepler is a space observatory launched by NASA in 2009 to discover Earth-sized exoplanets orbiting other stars [8]. Kepler mission developed several decades to answer the centuries-old question: How frequent are other Earth-like planets in Milkyway galaxy? In particular, what is the frequency of Earth-size planets in the Habitable-Zone of solar-like stars? There are three different types of exoplanets are common in our universe: gas giants, hot-super-Earths in short period orbits, and ice giants. The challenge is to find the terrestrial planets that are in the habitable zone of their stars where liquid water might exist on the surface on the planet.

The scientific objective of the Kepler mission is to explore the structure and diversity of planetary systems. This mission surveys a large sample of stars to determine the percentage of terrestrial and large planets that are in or near the habitable zone of a wide variety of stars and determine the distribution of size and shapes of the orbits of these planets. Kepler mission also estimates how many planets are in multiple-star systems. After collecting a large number of data points, using many techniques scientists determine the properties of those stars that harbor planetary systems including the planets itself.

1.1 Scientific Goals and Objectives of the Kepler Mission

The primary goal of the Kepler Mission is to survey our region of the Milky Way galaxy to discover hundreds of Earth-size or larger planets in or near the habitable zone of solar-like stars and determine the existence of solar system like planetary systems [2]. The scientific objective of the Kepler Mission is to explore the structure and diversity of extrasolar planetary systems. This is achieved by observing a large sample of stars to:

- Determine the frequency of terrestrial and larger planets in or near the habitable zone of a wide variety of stellar spectral types of stars.
- Determine the distribution of the size of planets and the size of the planets orbits.
- Estimate the frequency of planets in multiple star systems.
- Determine the distributions of orbital sizes, their light reflection properties (albedo), size, and density of short-period giant planets.
- Identify additional members of each photometrically discovered planetary system using complementary techniques.
- Determine the properties of those stars that harbor planetary systems. The spectral type, luminosity, and composition for each star showing transits are obtained from ground-based observations.
- Asteroseismology analysis of the data will be used to determine the mass, age, and size of the stars and astrometric analysis of the data will be used to

calculate their distances, which also is used to calculate the size of the stars.

1.2 Habitable Zone

Habitable Zone is the range in the distance from a star where liquid water could exist on the surface of a planet orbiting a star that possibly supports life. Liquid water is essential to all life on Earth, and so the definition of a habitable zone is based on the hypothesis that extraterrestrial life would share this requirement [10]. This is a very traditional definition, as a planet surface temperature may depend on other factors such as greenhouse gas abundance, its reflectivity, atmospheric and oceanic circulation, radioactive decay, and tidal heating within the planet. These energy sources can be easily allowed the planet to have subsurface liquid water reservoirs. Jupiter's moon Europa has liquid water ocean tens of kilometers below its surface that may well be habitable for some organisms. More than 20 planets, including the nearest extrasolar planet, Proxima Centauri b [1], have been found that are both roughly Earth-sized and orbiting within a habitable zone of their stars.

1.3 The Transit Method of Detecting Extrasolar Planets

The Kepler mission detects exoplanets using transit photometry [4]. When a planet passes in front of a star as a view from Earth, the event is called a "*transit*". For example, we can observe an occasional Venus or Mercury transit from Earth as a small black dot creeping across the Sun. During this transit, the transiting planet blocks the starlight and dips the brightness of the star. When this happens, we say the planet transits the star, and can be detected using transit photometry. During

transit, raw data is collected in the form of a sequence of stellar images, which are processed in to "light-curves" tracking the brightness of a star over time. Light-curves are graphs that show the brightness of an object over a period.

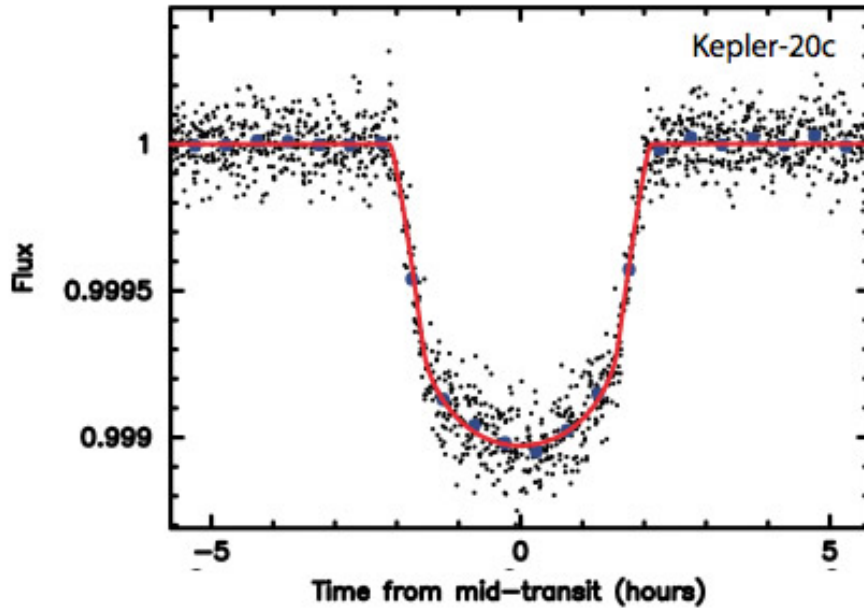


Figure 1.1: The light curve shown here was made from brightness data gathered by the Kepler Mission for discovery of planet Kepler-20c. Image Credit: NASA

Figure 1.1 shows the light curve made from the brightness data from Kepler Mission for the discovery of a planet named *Kepler-20c*. Data clearly show the flux of the host star drop notability when the planet is in transit. This is the primary signature we are looking in the transit photometry to discover planets that are in orbits around stars. These planets create this photometric signature periodically while they are orbiting around the star. To make things more complicated, the light curves of some other objects also create similar photometric signatures: eclipsing binaries stars, certain other variable stars. During the classification stage, we need

to classify these object as false positives.

1.4 Project Proporsal and Motivation

NASA's Kepler mission search for extrasolar planets has collected data from hundreds of thousands of star systems. Processing these large amounts of data, so far mission discovered 4,696 candidate exoplanets and 2,331 of them are confirmed as exoplanets leaving many more planet candidates to be confirmed in the dataset. These potential planetary candidates are searching using an algorithm that looks for periodic planetary transit in the light curves, but spurious intensity dips and other noise in the data due to non-planetary stellar variability has led to high false positive rates for detecting transits. As the initial planetary candidates found by this search method require extensive and costly subsequent validation, there is a need to reduce the error rate in exoplanet candidate identification. This exoplanet candidate data set refer to as Kepler Object of Interest (KOI). I present an algorithm (Multilayered Neural Network - MNN) to classify KOI as confirmed exoplanets or false positives using publically available data. MNN uses previously generated features extracted from the lightcurve time-series data, as well as newly generated autocorrelation features derived from the time-series specifically during the planet transits.

NASA launched the Kepler mission when I was an undergraduate student (Physics major) in 2009. Since the day the mission has been initiated, I closely follow the mission updates. I have moved to upstate NY to pursue my graduate studies in theoretical astrophysics (MS thesis in Numerical Relativity). Durig this time I have attended numerous scientific talks given by Kepler fellowship researchers at the University explaining and exploring the Kepler dataset. There are some groups are already using machine learning techniques to classify the Kepler data while others

are developing more manual and careful processes to extract planetary candidates. After taking the Udacity machine learning nano degree, it is natural for me to think of applying machine learning techniques to the Kepler dataset and see if that can improve the classification of planetary candidates.

Chapter 2

Problem Statement

Kepler is a single instrument spacecraft that collect most contiguous and long-running photometric time series possible. Kepler has the capability to observe approximately 170,000 stars simultaneously while it is operating. The fundamental objective of the Kepler mission is to detect a large number of transiting exoplanets. The ultimate goal of the primary mission was a characterize the frequency of exoplanets on diameter, orbital period and host star. The Manual classification of the findings of Kepler object has proven very time-consuming. The new space-based, transit photometry missions such as K2 [7], TESS [14], and PLATO 2.0 [13] also produce a large number of the dataset that demands some level of automation to do the classification.

Using machine learning classification techniques, we can speed up the process and provide a more continuous rating of planarity candidates. There are various of machine learning classification techniques has been applied to Kepler dataset including random forests, SVM, K-mean clustering [16, 12]. In this project, we are attempting to train a Multilayered Neural Network to identify the potential planetary candidates in the Kepler dataset.

Chapter 3

Datasets and Inputs

3.1 Kepler Pipeline

The Kepler Pipeline [8] is a data reduction pipeline used for translating the Kepler raw pixel data into possible transiting planet detections. Kepler mission perform photometric observations of carefully selected stars (around 156,000) using its 115 deg² field of view (FOV) as reviewed in Borucki et al. (2010) [3] and Koch et al. (2010) [9]. The Kepler Mission Science Operations Center (SOC) at NASA Ames Research Center performs major functions on these datasets including calibrate CCD array, download data (light curves) from the spacecraft periodically, remove systematic noise [15] and perform statistical tests to reject false positives and establish accurate statistical confidence in each detection.

Figure 3.1 show the major steps and modules of the pipeline. In particular the last two modules of the pipeline; those that identify as Threshold Crossing Events (TCEs) and their subsequent transit model fitting. TCE is a sequence of significant, periodic, planet transit-like features in the light curve of a target star. Transiting Planet Search module take takes as input the systematic error-correlated light curve for a star and searches a parameter space of possible transit signatures and outputs

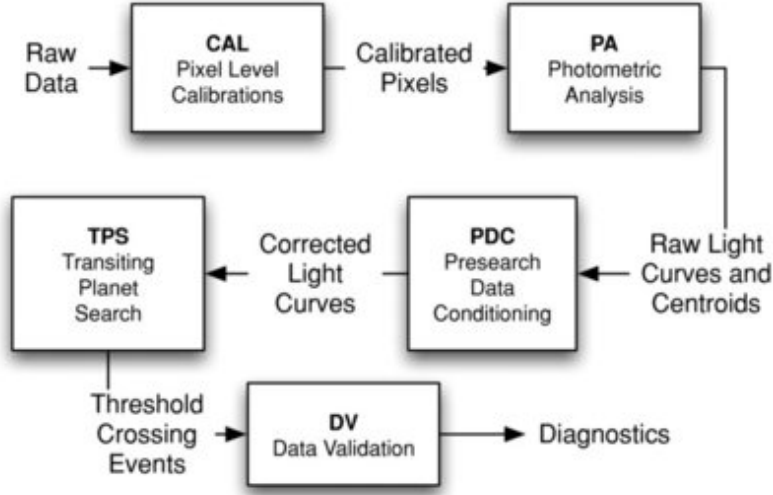


Figure 3.1: Data flow diagram for the SOC Science Pipeline. Image Credit: The American Astronomical Society

a TCE or say that does not exists TCE event on the target star. This produce smaller subset of target stars what is given to the Data Validation (DV) module. DV module takes initial TCE and gaps the transit signatures from the light curve and uses the Transiting Planet Search to find additional TCEs on the same target star. This process repeats until it finds all the TCEs on given star. More details on this process explained by Mandel and Agol (2002) [11] and Claret and Bloemen (2011) [5].

TPS algorithm detects transit-like features in light curves by applying noise compensating, wavelet-based matched filtering. TPS characterized the power spectral density (PSD) of the observation noise as a function of time to implement a whitening filter in the wavelet domain. The trial transit pulse is whitened and correlated against the whitened flux time series. Features with correlations above the threshold of 7.1σ are flagged as potential threshold crossing events and subjected

to additional tests in TPS to guard against false alarms.

Algorithm searches a parameter space with varying transit durations D and produce a Single Event Statistics (SES) time series that is the significance of the detection of the reference transmit pulse centered at that particular time for each D

$$SES(t) = N(t)/\sqrt{D(t)} \quad (3.1)$$

$\sqrt{D(t)}$, is the expected signal to noise ratio of a signal that exactly matches the template pulse and $N(t)$ is the correlated time series.

Multiple Event statistics (MES) is constructed that characterizes a significant detection in a search over varying orbital period p and epochs (phase) t_0 by folding $N(n)$ and $D(n)$. $MES > 7.1\sigma$ may produce a TCE if it also passes additional statistical tests. SES and MES are the basis of some of the attributes used in the training set.

3.2 KOI and TCE Sttributes

The Threshold Crossing Events (TCE) catalog contain a sequence of transit-like features in the flux time series of a given target star. These TCE data can download from NASA Exoplanet Archive databases ¹, and also the detail description of the table fields ² are listed as public data. Kepler Object of Interest (KOI) catalog contains object data including many attributes. The detail attributes are listed on NASA Exoplanet Archive website ³, and dataset can be download from the archive

¹http://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=q1_q17_dr24_tce

²http://exoplanetarchive.ipac.caltech.edu/docs/API_tce_columns.html

³http://exoplanetarchive.ipac.caltech.edu/docs/API_kepcandidate_columns.html

tables.⁴ All these data can be downloaded as bulk using data tools available in the archive website.

3.3 Some Important Attributes

There are a large number of attributes available in the KOI and TCE datasets. Out of these attributes, following attributes have significant importance in classification process according to the literature. We will be including these attributes alongside with other attributes that will be determined by using Independent Component Analysis.

- MES_{max}/MES_{min}
- SNR (for all-transit model fit)
- MES scaled by SES auto-correlation statistics
- χ^2 statistics for the all-transits model fit
- Ratio of the planet's semi-major axis to stellar radius
- The proportion of the light curve that was missing during this TCE's transit

3.4 Missing Attributes

In some circumstances, there are missing attributes in the dataset due to missing information in the stellar catalogs; the data validation fit fails to converge, or a processing timeout is reached. These missing attributes can be filled using reasonable methods such as mean, median, and in case of missing stellar attributes can be filled

⁴<http://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=cumulative>

using sun's parameters. These various substitutions will be tested when training the network.

3.5 Classification Labels

Each Threshold Crossing Event (TCE) is subject to a vetting process performed by the Kepler TCE Review Team (TCERT). During the triage (Initial) vetting stage, all TCEs are partitioned into two different sets: Problematic Light Curves that has instrumental noise and Kepler Object of Interest. KOI is a TCE that contains convincing transit-like features that do not present obvious evidence that the TCE was generated from non-transiting phenomena such as instrumental noise. These KOIs move to the next level of the vetting process performed by individuals manually inspecting light curves using detection statistics. Any indication they see the signal came from an eclipsing binary star or more complex forms of instrumental noise removed from the list. TCEs that survive this removal process are classified as Planet Candidates (PC).

We can identify three different types of classification labels in the processed dataset: Planetary Candidates (PC), Astrophysical False Positives (AFP) and non-transiting phenomena (NTP). PCs are confirmed as planets, statistically validated as planets or determined to be a planet candidate by the TCERT. AFPs are those TCEs that have been shown to be eclipsing binary stars or have shown evidence that the transiting object being detected is not located around the target star. NTP are those TCEs that failed the initial vetting process.

I am planning to use machine learning algorithm (Neural Network) to find a function that maps attributes produced by the Kepler Pipeline for each TCE to a classification label of PC, AFP or NTP. This classification function is purely based

on the statistical distributions of the attributes for each TCE and the algorithm does not attempt to physically model the process of the planet transit beyond what is already present in the TCE attribute catalogs.

Chapter 4

Solution Statement

Machine learning techniques contribute a way to automate some step of exoplanet discovery. The TCE vetting process is a tedious and time-consuming (mostly a manual) process. Modern astronomical observational instruments generate a large amount of data within a short period of observational time. These observations may contain such a crucial events that need future follow-up observations using other telescopes (using other wavelengths); hence, processing these time series data and extracting meaningful information is a time sensitive process. Solution to this is to process Kepler data using machine learning algorithms to express the classification while reducing human errors. I am planning to use trained Neural Network to process the TCE data to make the classification.

Chapter 5

Benchmark Model

I am planning to experiment with benchmarking the neural network by performing many training trials by varying number of training data sets. For each trail i , the training time to obtain a correct network t_i will be recorded. Trials which are not successful within a time limit T will consider as failures. Initial weight will be randomly chosen [6].

Chapter 6

Evaluation Metrics

Using Neural network classifier to make classification on Kepler TCEs is a discrete prediction machine learning problem. That is to say; the NN decide what category the given TCE belong to. There are many metrics available for us to check how well our NN is performing and I am planning to use couple of these metrics in this project,

- Accuracy Score
- Confusion Metrix
- Mean Squire Error

Chapter 7

Project Design

Theoretical workflow of the project

- Download data from NASA Exoplanet archives
- Clean up attribute values
 - Understanding correlated attributes and remove them from final attribute set
 - Produce additional Attributes
- Generate data into training, testing datasets.
- Develop a multilayered neural network using Python programming language
- Train the Neural Network
- Compare network results with the NASA Exoplanet Data

Bibliography

- [1] G. Anglada-Escudé, P. J. Amado, J. Barnes, Z. M. Berdiñas, R. P. Butler, G. A. L. Coleman, I. de La Cueva, S. Dreizler, M. Endl, B. Giesers, S. V. Jeffers, J. S. Jenkins, H. R. A. Jones, M. Kiraga, M. Kürster, M. J. López-González, C. J. Marvin, N. Morales, J. Morin, R. P. Nelson, J. L. Ortiz, A. Ofir, S.-J. Paardekooper, A. Reiners, E. Rodríguez, C. Rodríguez-López, L. F. Sarmiento, J. P. Strachan, Y. Tsapras, M. Tuomi, and M. Zechmeister. A terrestrial planet candidate in a temperate orbit around Proxima Centauri. *nat*, 536:437–440, August 2016.
- [2] N. M. Batalha. Exploring exoplanet populations with NASA’s Kepler Mission. *Proceedings of the National Academy of Science*, 111:12647–12654, September 2014.
- [3] William J. Borucki, David Koch, Gibor Basri, Natalie Batalha, Timothy Brown, Douglas Caldwell, John Caldwell, Jørgen Christensen-Dalsgaard, William D. Cochran, Edna DeVore, Edward W. Dunham, Andrea K. Dupree, Thomas N. Gautier, John C. Geary, Ronald Gilliland, Alan Gould, Steve B. Howell, Jon M. Jenkins, Yoji Kondo, David W. Latham, Geoffrey W. Marcy, Søren Meibom, Hans Kjeldsen, Jack J. Lissauer, David G. Monet, David Morrison, Dimitar Sasselov, Jill Tarter, Alan Boss, Don Brownlee, Toby Owen, Derek

Buzasi, David Charbonneau, Laurance Doyle, Jonathan Fortney, Eric B. Ford, Matthew J. Holman, Sara Seager, Jason H. Steffen, William F. Welsh, Jason Rowe, Howard Anderson, Lars Buchhave, David Ciardi, Lucianne Walkowicz, William Sherry, Elliott Horch, Howard Isaacson, Mark E. Everett, Debra Fischer, Guillermo Torres, John Asher Johnson, Michael Endl, Phillip MacQueen, Stephen T. Bryson, Jessie Dotson, Michael Haas, Jeffrey Kolodziejczak, Jeffrey Van Cleve, Hema Chandrasekaran, Joseph D. Twicken, Elisa V. Quintana, Bruce D. Clarke, Christopher Allen, Jie Li, Haley Wu, Peter Tenenbaum, Ekaterina Verner, Frederick Bruhweiler, Jason Barnes, and Andrej Prsa. Kepler planet-detection mission: Introduction and first results. *Science*, 327(5968):977–980, 2010.

- [4] D. Charbonneau, T. M. Brown, D. W. Latham, and M. Mayor. Detection of Planetary Transits Across a Sun-like Star. *apjl*, 529:L45–L48, January 2000.
- [5] A. Claret and S. Bloemen. VizieR Online Data Catalog: Limb-darkening coefficients (Claret+, 2011). *VizieR Online Data Catalog*, 352, March 2011.
- [6] Leonard GC Hamey. Benchmarking feed-forward neural networks: models and measures. In *NIPS*, pages 1167–1174. Citeseer, 1991.
- [7] S. B. Howell, C. Sobeck, M. Haas, M. Still, T. Barclay, F. Mullally, J. Troeltzsch, S. Aigrain, S. T. Bryson, D. Caldwell, W. J. Chaplin, W. D. Cochran, D. Huber, G. W. Marcy, A. Miglio, J. R. Najita, M. Smith, J. D. Twicken, and J. J. Fortney. The K2 Mission: Characterization and Early Results. *passp*, 126:398–408, April 2014.

- [8] J. M. Jenkins, D. A. Caldwell, H. Chandrasekaran, J. D. Twicken, S. T. Bryson, E. V. Quintana, B. D. Clarke, J. Li, C. Allen, P. Tenenbaum, H. Wu, T. C. Klaus, C. K. Middour, M. T. Cote, S. McCauliff, F. R. Girouard, J. P. Gunter, B. Wohler, J. Sommers, J. R. Hall, A. K. Uddin, M. S. Wu, P. A. Bhavsar, J. Van Cleve, D. L. Pletcher, J. A. Dotson, M. R. Haas, R. L. Gilliland, D. G. Koch, and W. J. Borucki. Overview of the Kepler Science Processing Pipeline. *apjl*, 713:L87–L91, April 2010.
- [9] D. G. Koch and et al. Kepler Mission Design, Realized Photometric Performance, and Early Science. *apjl*, 713:L79–L86, April 2010.
- [10] R. K. Kopparapu, R. M. Ramirez, J. SchottelKotte, J. F. Kasting, S. Domagal-Goldman, and V. Eymet. Habitable Zones around Main-sequence Stars: Dependence on Planetary Mass. *apjl*, 787:L29, June 2014.
- [11] K. Mandel and E. Agol. Analytic Light Curves for Planetary Transit Searches. *apjl*, 580:L171–L175, December 2002.
- [12] S. D. McCauliff, J. M. Jenkins, J. Catanzarite, C. J. Burke, J. L. Coughlin, J. D. Twicken, P. Tenenbaum, S. Seader, J. Li, and M. Cote. Automatic Classification of Kepler Planetary Transit Candidates. *apj*, 806:6, June 2015.
- [13] H. Rauer, C. Catala, C. Aerts, T. Appourchaux, W. Benz, A. Brandeker, J. Christensen-Dalsgaard, M. Deleuil, L. Gizon, M.-J. Goupil, M. Güdel, E. Janot-Pacheco, M. Mas-Hesse, I. Pagano, G. Piotto, D. Pollacco, C. Santos, A. Smith, J.-C. Suárez, R. Szabó, S. Udry, V. Adibekyan, Y. Alibert, J.-M. Almenara, P. Amaro-Seoane, M. A.-v. Eiff, M. Asplund, E. Antonello, S. Barnes, F. Baudin, K. Belkacem, M. Bergemann, G. Bihain, A. C.

Birch, X. Bonfils, I. Boisse, A. S. Bonomo, F. Borsa, I. M. Brandão, E. Brocato, S. Brun, M. Burleigh, R. Burston, J. Cabrera, S. Cassisi, W. Chaplin, S. Charpinet, C. Chiappini, R. P. Church, S. Csizmadia, M. Cunha, M. Damasso, M. B. Davies, H. J. Deeg, R. F. Díaz, S. Dreizler, C. Dreyer, P. Eggenberger, D. Ehrenreich, P. Eigmüller, A. Erikson, R. Farmer, S. Feltzing, F. de Oliveira Fialho, P. Figueira, T. Forveille, M. Fridlund, R. A. García, P. Giommi, G. Giuffrida, M. Godolt, J. Gomes da Silva, T. Granzer, J. L. Grenfell, A. Grottsch-Noels, E. Günther, C. A. Haswell, A. P. Hatzes, G. Hébrard, S. Hekker, R. Helled, K. Heng, J. M. Jenkins, A. Johansen, M. L. Khodachenko, K. G. Kislyakova, W. Kley, U. Kolb, N. Krivova, F. Kupka, H. Lammer, A. F. Lanza, Y. Lebreton, D. Magrin, P. Marcos-Arenal, P. M. Marrese, J. P. Marques, J. Martins, S. Mathis, S. Mathur, S. Messina, A. Miglio, J. Montalbán, M. Montalto, M. J. P. F. G. Monteiro, H. Moradi, E. Moravveji, C. Mordasini, T. Morel, A. Mortier, V. Nascimbeni, R. P. Nelson, M. B. Nielsen, L. Noack, A. J. Norton, A. Ofir, M. Oshagh, R.-M. Ouazzani, P. Pápics, V. C. Parro, P. Petit, B. Plez, E. Poretti, A. Quirrenbach, R. Ragazzoni, G. Raimondo, M. Rainer, D. R. Reese, R. Redmer, S. Reffert, B. Rojas-Ayala, I. W. Roxburgh, S. Salmon, A. Santerne, J. Schneider, J. Schou, S. Schuh, H. Schunker, A. Silva-Valio, R. Silvotti, I. Skillen, I. Snellen, F. Sohl, S. G. Sousa, A. Sozzetti, D. Stello, K. G. Strassmeier, M. v Svanda, G. M. Szab, A. Tkachenko, D. Valencia, V. Van Grootel, S. D. Vauclair, P. Ventura, F. W. Wagner, N. A. Walton, J. Weingrill, S. C. Werner, P. J. Wheatley, and K. Zwintz. The PLATO 2.0 mission. *Experimental Astronomy*, 38:249–330, November 2014.

- [14] G. R. Ricker, J. N. Winn, R. Vanderspek, D. W. Latham, G. Á. Bakos,

J. L. Bean, Z. K. Berta-Thompson, T. M. Brown, L. Buchhave, N. R. Butler, R. P. Butler, W. J. Chaplin, D. Charbonneau, J. Christensen-Dalsgaard, M. Clampin, D. Deming, J. Doty, N. De Lee, C. Dressing, E. W. Dunham, M. Endl, F. Fressin, J. Ge, T. Henning, M. J. Holman, A. W. Howard, S. Ida, J. Jenkins, G. Jernigan, J. A. Johnson, L. Kaltenegger, N. Kawai, H. Kjeldsen, G. Laughlin, A. M. Levine, D. Lin, J. J. Lissauer, P. MacQueen, G. Marcy, P. R. McCullough, T. D. Morton, N. Narita, M. Paegert, E. Palles, F. Pepe, J. Pepper, A. Quirrenbach, S. A. Rinehart, D. Sasselo, B. Sato, S. Seager, A. Sozzetti, K. G. Stassun, P. Sullivan, A. Szentgyorgyi, G. Torres, S. Udry, and J. Villaseñor. Transiting Exoplanet Survey Satellite (TESS). In *Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave*, volume 9143 of *procspie*, page 914320, August 2014.

- [15] J. C. Smith, M. C. Stumpe, J. E. Van Cleve, J. M. Jenkins, T. S. Barclay, M. N. Fanelli, F. R. Girouard, J. J. Kolodziejczak, S. D. McCauliff, R. L. Morris, and J. D. Twicken. Kepler Presearch Data Conditioning II - A Bayesian Approach to Systematic Error Correction. *pasp*, 124:1000–1014, September 2012.
- [16] G. Torres, D. M. Kipping, F. Fressin, D. A. Caldwell, J. D. Twicken, S. Ballard, N. M. Batalha, S. T. Bryson, D. R. Ciardi, C. E. Henze, S. B. Howell, H. T. Isaacson, J. M. Jenkins, P. S. Muirhead, E. R. Newton, E. A. Petigura, T. Barclay, W. J. Borucki, J. R. Crepp, M. E. Everett, E. P. Horch, A. W. Howard, R. Kolbl, G. W. Marcy, S. McCauliff, and E. V. Quintana. Validation of 12 Small Kepler Transiting Planets in the Habitable Zone. *apj*, 800:99, February 2015.