Science merit function for the Kepler mission

William J. Borucki,[a,*] Jon M. Jenkins,[b] and Riley M. Duren[c,d]
[a]NASA Ames Research Center, Moffett Field, California, United States
[b]NASA Ames Research Center, Moffett Field, California, United States
[c]University of Arizona, Tucson, Arizona, United States
[d]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, United States

Abstract. The Kepler mission was a National Aeronautics and Space Agency (NASA) Discovery-class mission designed to continuously monitor the brightness of at least 100,000 stars to determine the frequency of Earth-size and larger planets orbiting other stars. Once the Kepler proposal was chosen for a flight opportunity, it was necessary to optimize the design to accomplish the ambitious goals specified in the proposal and still stay within the available resources. To maximize the science return from the mission, a merit function (MF) was constructed that relates the science value (as determined by the PI and the Science Team) to the chosen mission characteristics and to models of the planetary and stellar systems. This MF served several purposes; predicting possible science results of the proposed mission, evaluating the effects of varying the values of the mission parameters to increase the science return or to reduce the mission costs, and supporting quantitative risk assessments. The MF was also valuable for the purposes of advocating the mission by illustrating its expected capability. During later stages of implementation, it was used to keep management informed of the changing mission capability and support rapid design tradeoffs when mission down-sizing was necessary. The MF consisted of models of the stellar environment, assumed exoplanet characteristics and distributions, detection sensitivity to key design parameters, and equations that related the science value to the predicted number and distributions of detected exoplanet. A description of the MF model and representative results are presented. Examples of sensitivity analyses that supported design decisions and risk assessments are provided to illustrate the potential broader utility of this approach to other complex science-driven space missions. [DOI: 10.1117/1.JATIS.6.4.044003]

Keywords: exoplanets; Kepler mission; planet detection; system engineering.

Paper 20047 received May 7, 2020; accepted for publication Oct. 19, 2020; published online Nov. 5, 2020.

1 Introduction

The Kepler mission[1–4] began as a Principal Investigator (PI)-led NASA Discovery-class mission (#10) that was designed to explore the structure and diversity of planetary systems. Specifically, the mission was to survey the extended solar neighborhood to detect and characterize hundreds of terrestrial and larger planets in or near the habitable zone (HZ); i.e., the region surrounding a star where its radiant flux is appropriate for liquid water to exist on the planet's surface.

After acceptances of the year-2000 Kepler proposal to the "NASA, Announcement of Opportunity Discovery Program 2000, A.O. 00-OSS-02, 2000.6" and the required Concept Study Report (CSR), authorization to proceed with mission development was received in 2001. Because no previous mission had been designed to do a high-precision [i.e., 10 parts-per-million (ppm) for bright, quiet stars] photometric search of thousands of stars to find planets, there was little history or experience to guide the development of such a mission. Fortunately, ground-based observatories that conducted searches for exoplanets by the photometric method

*Address all correspondence to William J. Borucki, William.J.Borucki@nasa.gov

had been constructed during the period between 1994 and 2000 and some experience was gained from operating the observatories.[5–9]

At the time of mission selection, the number and distribution of terrestrial planets expected in extra-solar planetary systems (exoplanets) were unknown. Theories of planetary system formation were not well enough developed to make useful predictions of their sizes, occurrence frequencies, or distributions. It was commonly assumed that other planetary systems would be similar to our Solar System with small rocky planets close to the star and giant planets at much larger distances. Early models of planetary system formation[10–13] predicted the formation of many planets smaller than that of the Earth and few larger than the Earth for values of the semi-major axes near 1 Astronomical Unit (AU). Contrary to such expectations, the early ground-based observations[5,14,15] showed the presence of giant planets in short period orbits. These observations were explained as planets that formed beyond the ice-line and later migrated inward toward their parent stars. During their inward spiral, it was expected that they would have swept away most of the small rocky planets. Because the mechanism to stop inward spiral before the planet fell into its star was unknown, it was possible that many stars had been swept clean of planets. Consequently, the Kepler mission was designed to characterize planetary systems for several possible situations; i.e., (1) like our own, (2) with only giant planets in inner orbits, (3) without any planets, and (4) with many small planets near 1 AU.

Before the Kepler mission proposal could be considered for implementation, quantitative predictions for the expected results and a discussion of the risks were required. Therefore, estimates of planetary sizes and distributions were necessary prior to mission development to provide estimates of the mission science return. Because theories predicted that the formation of stars implied the concurrent formation of planetary systems, the merit function (MF) model assumed that most stars would have planets and that a sufficiently large sample of stars would be observed such that a null result would be significant. In particular, it was assumed that if detections of 50 or more Earth-size planets in the HZ of solar-like stars could be expected, then the mission goals would be satisfied. Based on a geometric probability that the orbit of a planet near the HZ was aligned with our line of sight, and a working estimate that 10% of stars would have such planets, the mission would need to monitor at least 100,000 stars. Thus, the original proposal assumed 100,000 stars would be monitored with a sensitivity to detect Earth-size planets orbiting solar-like stars and that the mission duration would be four years. In the CSR, the number of target stars was increased to 170,000.

Space mission implementation is a high-risk endeavor, particularly for innovative investigations such as the Kepler mission involving a multidimensional science performance space with multiple (and often nonlinear) sensitivities to design parameters. The Kepler team employed traditional systems engineering practices of suballocating mission requirements for a given science reference case (e.g., detecting an Earth-size planet around a solar-type star) and developed comprehensive error budgets for key individual design parameters such as combined differential photometric precision.[2,16] However, an additional analytic framework was needed to manage the full mission complexity given additional and overlapping factors influencing science return such as the number and characteristics of observable stars, the diversity of planet sizes and orbital periods, and mission duration. Therefore, following selection for mission development, it was imperative to devise an MF to evaluate the science return for the nominal system design and a variety of situations involving reduced performance. The MF needed to connect the accomplishment of mission science goals to the mission design parameters and to astrophysical assumptions and risks. This MF served several purposes; it related the specific science goals with specific instrument and mission performances, related the mission performance with respect to the mission baseline (full science return) and "floor" (minimum science return) requirements during mission development, it enabled design trade-offs, and it supported quantitative risk assessment. When trade-offs were made, various options were explored and those that did most damage to the science value were identified and rejected. In a proactive approach, the derivative of the science results to changes in mission parameters were estimated and changes were made that enhanced the science when that could be done for minimal increases in resources. An especially valuable use of the MF was to predict the expected results as a function of assumptions about extrasolar planet frequencies, planet size and distributions, and associations to star characteristics. Risks were assessed by running "what if" scenarios to determine the quantitative impacts

of unknowns related to the target populations and to various levels of noise sources from the instrument and from stellar variability. The ability to estimate the effects of unknowns and mission-design changes on the expected science product[17] was particularly important for an exploratory mission where little was known about the situation to be investigated. Reference 17 provides an example of a prelaunch prediction for the Kepler mission.

The MF was an algorithm designed to produce numerical values for the science return given a set of inputs representing an instrument/spacecraft point design, estimates of the stellar structure of galaxy, exoplanet size and frequency distributions, and stellar and instrument noise. For example, increasing the mission duration affects the number of expected detections and the types of stars that can be searched for planets in their HZs. The MF included tables of stellar distributions of brightness, size, temperature, and mass. A model of the expected performance for the instrument point design was also included.

Calculations based on these inputs estimated the signal-to-noise-ratio (SNR) for each star and thus, the number of planets that could be detected in various orbits. A rating system was included that assigned values to the minimum size of a planet that could be detected as a function of the semi-major axis. In turn, the selection of the values for the semi-axes determined the number of transits observed for a given stellar mass. The model considered two separate aspects of the mission goals; assessing the capability of the mission to detect small planets in [i.e., planet radius $(Rp) \leq 2$ Earth-radius $(R_{\oplus})$] in the HZ, and the capability to determine the structure of planetary systems by finding small planets at a range of semi-major axes that were near but not in the HZ. Because the structures of other planetary systems were unknown, the model assumed that other planetary systems were similar to our Solar System in having three planets in or near the HZ.

Two complementary algorithms (labeled "model 1" and "model 2") were developed and their scores were combined to provide a single value for the MF. The first algorithm computed a score based entirely on the number and size of detectable planets in the HZ of each star. The second algorithm computed a score for Earth-size planets placed in several orbits near, but not in, the HZ. The nine values of the semi-axis ranged from 0.05 AU to the orbital distance of Mars (1.5 AU). This range covered positions in and near the HZ for star types from M- to G-dwarfs. Together, the two algorithms addressed the goal of determining the occurrence frequency of small planets in the HZ and the goal of exploring the diversity of planetary systems.

The results of the MF were found to be valuable for the purposes of illustrating the mission capability before it was selected for implementation, for understanding risk, and for keeping mission managers informed of the changing status of the Mission capability during its development.

In this paper, the science goals and mission overview are covered in Sec. 1. A short overview of the mission approach is presented in Sec. 2, while the structure of MF is discussed in Sec. 3. The expected results from mission are given in Sec. 4 and some factors controlling the results are presented in Sec. 5. Results from risk studies are considered in Sec. 6. A brief comparison of model predictions and mission results is presented in Sec. 7. A summary with a discussion of strengths and weaknesses of the MF approach is presented Sec. 8.

2 Science Goals and Overview of Mission Approach

The scientific goal of the Kepler mission (as stated in the CSR that was an update to the year-2000 mission proposal[4]) was to explore the structure and diversity of planetary systems with special emphasis on determining the frequency of Earth-size planets in the HZ of solar-like stars. This was achieved by surveying ~170,000 stars to:

1. Determine the frequency of 0.8 $R_{\oplus}$ and larger planets in or near the HZ of a wide variety of spectral types (SpT) of stars.
2. Determine the distributions of sizes and orbital semi-major axes of these planets.
3. Estimate the frequency of planets orbiting multiple-star systems.
4. Determine the distributions of semi-major axis, eccentricity, albedo, size, mass, and density of short period giant planets.

5. Identify additional members of each photometrically-discovered planetary system using complementary techniques.
6. Determine the properties of those stars that harbor planetary systems.

A photometer was designed to be placed in heliocentric orbit to observe the periodic dimming in starlight caused by planetary transits. The photometer consisted of a 0.95 m, Schmidt-type telescope with a 113 square-degree (sq-deg) field-of-view (FOV). During the 4 years of its operation, it monitored the individual brightness of over 170,000 stars imaged onto a large array of CCD detectors.[2,4] Several thousands of the targets were changed on a quarterly basis to accommodate asteroseismology and guest-investigator studies. Thus, over 190,000 stars were observed during the mission duration prior to the failure of two of its reaction wheels.

The time-series of brightness observations were used to measure the depth and time of the transits. These values were then used to estimate

- planet size from the brightness change and from independent measurement of the stellar radius.
- orbital period from the time between transits.
- semi-major axes based on the orbital period, the mass of the star, and Kepler's third law.
- whether the planet was in the HZ based on the heat flux incident on the planet calculated from the distance to the host star and on the stellar size and temperature.

The HZ was defined as the region around the host star where water on the surface of a rocky planet could be in liquid form and was considered the region most likely to be conducive to the evolution of life.

To help distinguish planetary transits from statistical fluctuations in the data and from astrophysical phenomena that mimic transits, at least three transits were required for a valid detection. Consequently, the mission design required a mission duration of at least three times the longest orbital period sought. Analysis using the MF indicated that the Mission duration should be four years to enhance the detectability of Earth-size planets transiting stars as large as the Sun.

Figure 1 shows the general layout of the instrument. The detector array was at prime focus and was cooled by heat pipes that carried the heat to a radiator in the shadow of the spacecraft. The low-level electronics were placed immediately behind the focal array. Because of the very large FOV, the focal surface was strongly curved. To focus the images on the flat charge-coupled detectors (CCDs), each CCD was covered with a field-flattening sapphire lens. A four-vane spider supported the focal plane and its electronics and contained the power and signal cables and the heat pipes.

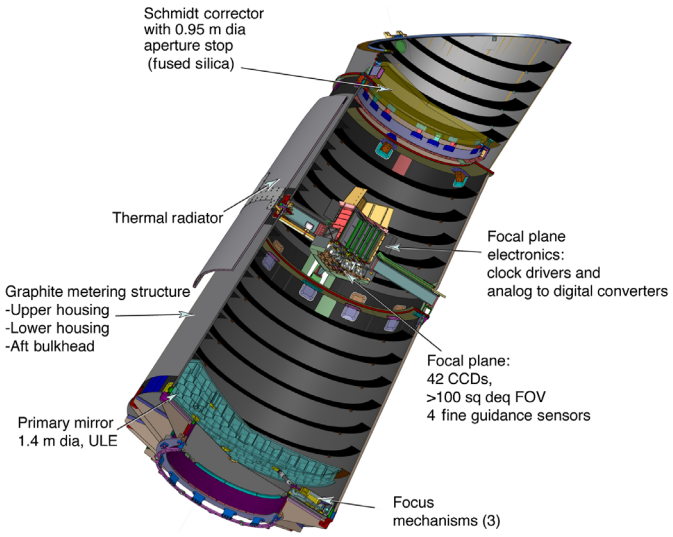

Fig. 1 Schematic diagram of the point design of the Kepler instrument.

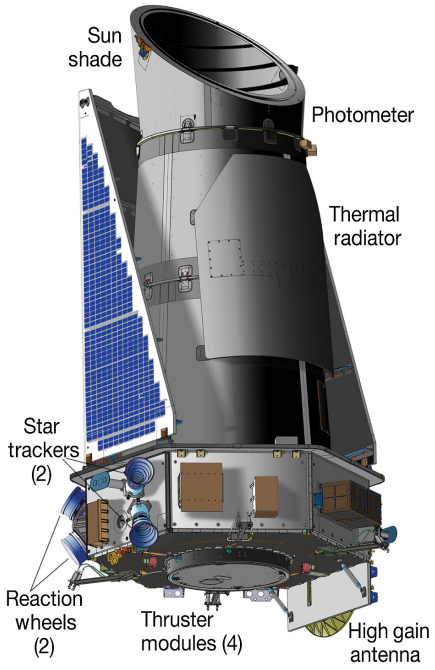

Fig. 2 Integrated spacecraft and photometer.

The spacecraft bus enclosed the base of the photometer and provided solar panels for power and contained the communication, navigation, and power equipment (Fig. 2). Several antennas with different frequency coverage and gain patterns were present for uplink commanding and for data downlink. Approximately 9 Gigabytes (GB) of science data were transferred to the ground every month when contact was made with the NASA Deep-Space Network. More extensive information and references can be found in Refs. 2 and 4.

The point design for the mission assumed

i. 113 sq-deg FOV.
ii. 0.95-m aperture with a 30% blockage due to the detectors situated at prime focus.
iii. 170,000 stars monitored simultaneously.
iv. 94% duty cycle for science observations after commissioning.
v. Mission duration of 4 years.

3 Merit Function Description

The MF had six parts;

i. Model of the stellar characteristics and distribution of apparent brightness.
ii. Models of the exoplanet characteristics and distribution.
iii. Spacecraft and instrument point design.
iv. Assignment of values for science products.
v. Algorithms to calculate the expected science products.
vi. Presentation and interpretation of the results.

3.1 *Model Values for Stellar Masses, Sizes, and Temperatures*

Tables of input values for stellar masses and sizes were obtained from Ref. 18. The distributions for areal density of stars as functions of spectral type and apparent magnitude were based on Besancon Model.[19] To maximize the number of stars observed, the FOV was chosen to be near the galactic plane. The stellar model included the stellar mass, size, and temperature for SpT

from A5 to M7. However, a case was considered where stars earlier than F5 were omitted because they would be too large to produce useful transit amplitudes or were too rare or had spectra that were inappropriate for planet confirmation by radial-velocity (RV) observations. Similarly, stars later than M7 were considered to be too dim to contribute meaningfully to the results.

Calculations showed that the power spectrum of stellar variability, even at a level of several ppm, influenced the minimum size and orbital characteristics of the planets that could be detected. The noise level contribution was a function of the transit duration. In turn, the transit duration depended on the star type and planetary semi-major axis. At the time of the mission development, only measurements of the Sun provided a sufficiently precise power spectrum that could be used as a model for other stars. Its power spectrum was known at a precision of ~ 10 ppm for periods appropriate for transit durations (i.e., several hours). The power spectrum[20] used in the model was based on the data[21] from the Differential Absolute Radiometer (DIARAD) instrument aboard the Solar and Heliospheric Observatory (SOHO) spacecraft appropriate for an average activity level.

The SNR was calculated as the square root of the integral of the ratio of the frequency response of a transit (of a given duration and depth) to the power spectrum of the expected noise: a combination of the nonwhite solar noise to which white noise associated with instrument and shot noise were added.[22] The SNR was then scaled by the square root of the number of transits, as determined from the orbital period. See Appendix A for the details of the calculation of the SNR.

3.2 *Models of the Exoplanet Characteristics and Distribution*

Two complementary models of planetary systems were used for the calculations. The first model (model 1) provided the science value for small planets in the HZ. This model assumed that there was only one planet in each planetary system and that it was in the HZ. However, for this model, the size of this planet was varied to estimate the mission capability for assumptions as to planet size. In particular, the expected number of detectable planets was tabulated for sizes of 0.53, 1, 1.3, and $2\,R_{\oplus}$; i.e., for Mars, Earth, and two super-Earth sizes. The approximate value of the semi-major axis and the period appropriate for the HZ for each star type was computed from the results.[23]

The second model (model 2) considered only non-HZ planets and assumed the size of each planet to be $1\,R_{\oplus}$. The semi-major axes of these planets were set to 0.05, 0.1, 0.2, 0.4, 0.8, 0.9, 1, 1.2, and 1.5 AU to explore the region near the HZ of a variety of stellar types. According to a study,[10] a peak near 1 AU should occur in the occurrence frequency of Earth-size planets for stellar masses between 0.5 and 1.5 times the mass of the Sun ($M_{\odot}$). To avoid double counting planets in the HZ, the results for a planet that overlapped the HZ position of a specific star-type were omitted for the calculations used in model 2. Although a distribution of planets as closely spaced as specified here might result in a dynamically unstable situation for some planets, the objective of the MF was to explore the statistical distribution of planets versus semi-major axis.

To better represent the Solar System, the model assumed that only three planets (representing Venus, Earth, and Mars) exist in this region. Therefore, the results from model 2 were multiplied by 2/9. This choice meant that the combination of models 1 and 2 considered three planets in and near the HZ.

3.3 *Assignment of Values for Science Products*

For the model calculations, planets with semi-major axes larger than 1.5 AU were ignored because planets in such orbits would not provide the minimum number of transits (three) during the mission life time of 4 years and because most such orbits would not be in the HZ for solar-like stars. Planets larger than twice the size of the Earth were not considered because they would be so large and/or massive that they could be detected by ground-based observatories. Based on the calculations in the literature,[24] Earth-size planets would be distributed at random positions; i.e., Earth and super-Earth size planets might be expected anywhere from the orbit of Mercury to the orbit of Mars.

The overall science value was very sensitive to the choice of the values assigned to various model outcomes. For example, if the value assigned to the detection of large-size planets in short-period orbits was set equal to that of a small Earth-size planet in the HZ, then maximization of the MF would bias the design to observe many different regions of the sky for a few months each. This result occurs because of the high probability of detecting large planets in short period-orbits versus the low probability of detecting the low-amplitude signals from the transits of small planets in long-period HZ orbits that provide few transits. Thus, the choice to detect the maximum number of planets would emphasize the detection of thousands of large planets in short period orbits, but would result in the complete loss of all Earth-size planets in the HZ of solar-like stars. Instead, the model factors that represent science values in the model were weighted to emphasize the detection of small planets in the HZ. The selection of the weights for the science products was made by the PI in consultation with the mission science team and the science community.

The values generated by the MF were based on the planet size and orbital position relative to the HZ. In particular, planets detected in the HZ were considered to be more valuable than those not in the HZ (non-HZ planets). The ability to detect smaller planets was considered to be much more valuable than the capability of finding larger planets because a positive result for a small size ensures that all larger sizes would also be detected. Therefore, the capabilities to detect Mars-size, Earth-size, 1.3 and 2 $R_{\oplus}$, planets were assigned weights of 40, 20, 5, and 1, respectively. Because ground-based observations can detect planets larger than twice the diameter of the Earth, planets larger than 2 $R_{\oplus}$ were not considered. Such planets were also expected to have extensive hydrogen atmospheres i.e., to be nonterrestrial planets incapable of supporting carbon-based life.

A weight of 100 was assigned to the results expected from the instrument design as proposed in the year-2000 proposal. Around 65% of the total score was allotted to planets detectable in the HZ and 35% to non-HZ planets. Modifications to the mission design or revisions to the mission assumptions were ranked relative to the CSR performance level (designated as case#0).

For model 1 (only HZ-planets), the algorithm calculated: (1) the probability of orbital alignment for each star, (2) the number of transits that occur for a planet in the HZ during duration of the mission, and (3) the SNR based on the size and brightness of the star, the transit duration, quadrature sum of noise sources, and the number of transits. Then the probability of recognition was calculated from the value of the SNR relative to the threshold SNR. The probability of recognition was based on a Gaussian distribution where the difference between the calculated SNR and the threshold SNR was taken to be the significance of the detection: i.e., for a threshold SNR $= 7\sigma$, and for calculated SNRs of 7σ and 8σ, the recognition rates would be 50% and 92%, respectively. A score was assigned based on the size of the smallest planet that can be detected in the HZ for each star. This process was repeated for all of the target stars in the FOV and then the scores were summed and normalized. To avoid over counting the score for stars for which smaller planets could be detected, the score for each selected size was calculated only for the stars that were in addition to those already accounted; i.e., those with smaller-size detectable planets.

The evaluation of model 2 (non-HZ planets) included the weighting of the results based on the rapidly decreasing geometrical probability of the orbital alignment with increasing semi-major axis. Planets very close to their star have a high probability of geometrical alignment (i.e., probability $= r* / R_{\text{orbit}}$, where $r*$ is the radius of the star and R_{orbit} is the orbital radius). For example, planets at 0.05 AU from their star are 20 times more likely to be correctly aligned than those at 1AU. Furthermore, the number of transits observed during the duration of the mission is much larger for planets with small semi-major axes than for planets with large semi-major axes. (Kepler's third law; i.e., the period squared is proportional to the semi-major axis cubed.) Thus, equal-size planets in the line-of-sight with a semi-major axis of 0.05 AU present about 89 times the number of transits of a planet with a 1 AU orbit. This results in the shorter-period planets having a total SNR ~ 9 times higher than long-period planets and are, therefore, much more easily recognized. Based on these two factors, most of the detections (and therefore, most of the MF scores) would be from planets orbiting too close to their star to be in or near the HZ.

However, planets near the HZ were considered to have a higher science value because such detections were more consistent with the goals of the mission. To reduce the biases introduced by small semi-major axes, the MF value of each recognized planet in Model 2 was weighted by multiplying its value by the reciprocal of the square of its semi-major axis (in AU).

The combination of weighting of the score for the position of each planet and limiting the contribution to the total score of 35% for non-HZ planets ensured that the MF scores would be dominated by the mission performance with respect to small planets in and near the 1 AU.

A final score was computed by adding the values for models 1 and 2. It is recognized that different weights and normalization factors could have been chosen for the calculations. The values and weights chosen here were considered appropriate by the PI and his science team. An outline of the MF approach is presented in Appendix B.

The orbital period and transit duration were then calculated for the selected values of the mass of each star type, the selected semi-major axis, and the semi-major axis for the HZ. The number of transits observed during the mission was calculated from the orbital period and mission duration. The calculated value of the transit duration was used to estimate the contribution of stellar variability noise and shot noise associated with the integrated photon flux. The value of stellar variability contribution to the total noise was derived from the power spectrum of the Sun for the calculated transit duration. Because the solar noise increases with period, the SNR of the transits depends on the transit duration.[22] A value for the instrument noise was based on "shot noise" and the effect of "jitter" associated with tracking error.

Once the properties of the transit and the number of transits that occur during the mission were computed, initial values for the noise level and SNR were computed for each star. Next, corrections for missing data and outages were made. To account for missing data when they were frequent but of short duration with respect to 24-h transits, the total number of photoelectrons counted over all transits was reduced by the expected loss; i.e., the duty cycle was reduced. This option lowered the total SNR of each model planet orbiting each star type at each model orbit and planet size. The lower SNR reduced the expected number of planetary detections because fewer events had sufficiently high SNR to exceed the detection threshold. This procedure for revising the SNR was labeled correcting the "completeness."

After correcting the expected number of detected planets for completeness, corrections were made for missing transits; i.e., a correction for "contiguity." For this calculation, a value was assumed for the expected number of outages that resulted in the loss of observations for periods of one or more days (i.e., a period sufficiently long to cause the complete loss of the longest expected transit).

Two separate multidimensional arrays were formed for calculations considering HZ and non-HZ planets. The coordinates of first array for the HZ planets were; stellar spectral type, stellar magnitude, and planet size. The coordinates of the array for non-HZ planets were; spectral type, stellar magnitude, and orbit size. For each point in the arrays the probability of an intersection of a transit with a data outage was computed from the binomial distribution for $N \leq N$ max transits (based on the orbital period, mission duration, and contiguity value). For each value of N, the SNR for the observed number of transits was calculated followed by the calculation of the detection rate. Summing these detection rates weighted by the probability of observing N transits for all possible values of N (N min $\leq N \leq N$ max) provided the total probability that a planet at this point in each array would be detected. Multiplying by the number of planets at this point in the array provided the total number of recognized planets at this location. The MF also accounts for the fact that the total number of observed transits is determined by the time to first transit as well as the orbital period and the mission duration. Repeating the procedure for all points in both arrays provided the number and distributions of expected detections given the input value of contiguity.

In the next section, the mission expectations are presented for the "point-design"; i.e., the expected results for the values of the mission parameters specified in the CSR. Section 5 will consider the effects of varying values of the mission parameters.

4 Expected Results from Mission Operations

4.1 *Mission Capability with Respect to Planet Size and Spectral Type*

The expected number of detectable planets increases with the assumed size of the planets until all the planets in the FOV are sufficiently large with respect to their parent star to produce transit SNRs well above the detection threshold.

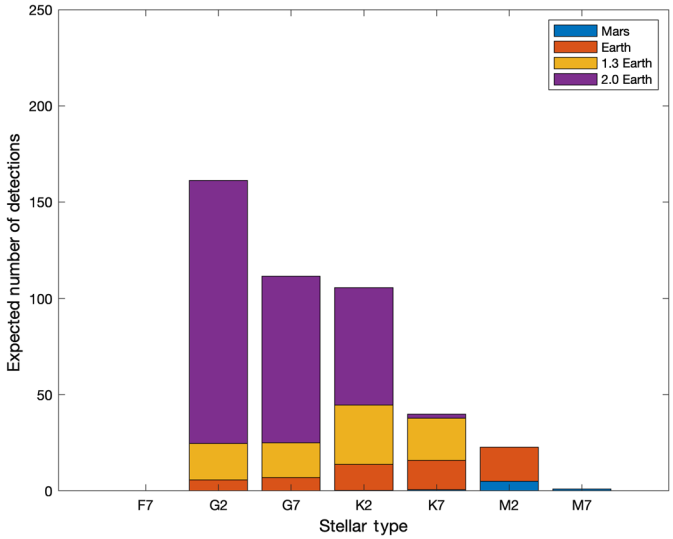

Fig. 3 Expected number of detections of planets in the HZ versus spectral type. The coloring of the bars is related to the size of the planets found at each star type.

Figure 3 shows the distribution of the expected number of detectable HZ planets with planet size and stellar type assuming that all the mission assumptions for case #1 are appropriate. The colored bars represent expectations for four different assumptions; (1) all stars have only Mars-size planets, (2) all stars have only Earth-size planets, (3) all stars have only 1.3 $R_{\oplus}$-size planets, and (4) all stars have only 2 $R_{\oplus}$ planets. In particular, for the assumption that all stars have HZ planets that are similar in size to Mars, then the results shown by the blue bars are applicable. Figure 3 shows that the mission has the capability of finding Mars-size planets mostly for M-type stars. However, if larger planets are common, then many more planets orbiting a wider range of stellar types are expected as shown by the sum of the brown-, yellow, and purple-colored bars. The predicted numbers of detectable planets for each stellar type and for any of the chosen planets' sizes (i.e., Mars-size, Earth-size, 1.3 and 2 $R_{\oplus}$) are indicated by the sum of the heights of the appropriate bars; e.g., the detection of 40 planets with Rp $\sim$ 1.3 $R_{\oplus}$ is expected for stars of spectral type K2. It is worth noting that for the M-type stars, the bars are saturated; considering a larger planet size does not lead to more detections because the amplitudes of the transits of small planets are already larger than the detection threshold and because there are no more M-type stars in the FOV that have magnitudes less than the limiting value.

Summing the values of the blue- and brown-colored bars indicates that 66 Earth-size planets could be detected in the HZ if all stars had such planets. If only 10% of the stars have Earth-size planets, a statistically useful number could be expected. If a null result was found instead, the mission results would indicate that such planets were rare; i.e., substantially less than 10% of stars had such planets in their HZ. The expected number of detectable planets which have a radius $\sim$1.3 $R_{\oplus}$ is more than double that for Earth-size planets. It is worth noting that planets with $Rp \leq 1.3\,R_{\oplus}$ are expected to have a rocky composition.[25] Consequently, it is likely there would be a sufficient number in each of the SpT G2 through M2 to provide a coarse estimate of the occurrence frequency of small/rocky planets vs spectral type. If most stars had planets twice Earth size, then 520 detections would be expected for planets in the HZ. Thus, the mission capability should be sufficient to address the mission goals for the planetary occurrence frequency for small planets, their occurrence frequency versus spectral type, and their size distribution.

Figure 4 shows the results for the expected occurrence frequency of planets that are not in the HZ (i.e., non-HZ planets). The calculations assume that at eight nonoverlapped values of the nine values of the semi-major axes (specified in the figure legend) there is the possibility of an Earth-size planet. The model assumes that each planetary system will have an average of three planets; one in the HZ and two in non-HZ orbits.

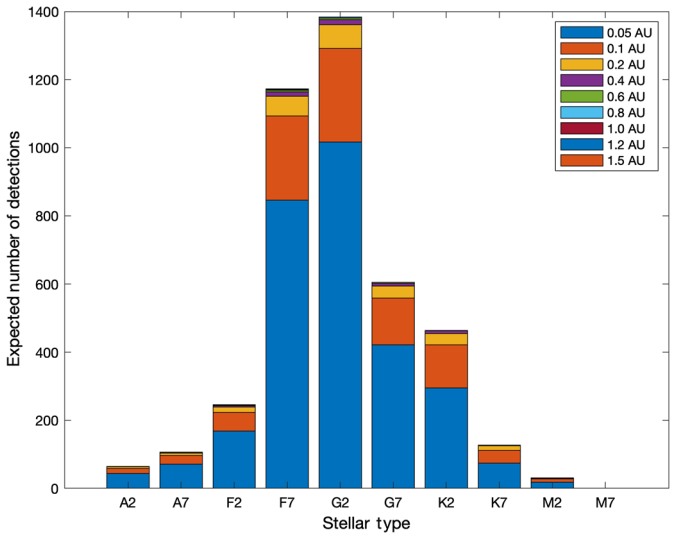

Fig. 4 Expected number of detections of planets not in the HZ versus spectral type. The coloring of the bars is related to the size of the semi-major axis for a given star type.

It is clear from Fig. 4 that most of the detections are expected at the smallest values of the semi-major axes (i.e., at the shortest orbital periods). This result is expected because the probability of transits is proportional to the ratio of the stellar size to the orbital size and because of the increased value of the SNR caused by the large number of transits for planets in short-period orbits. Summing the values in Fig. 4 indicates that over 4200 non-HZ planets should be found provided that planets at least as large as the Earth are common. Although this value might appear to be surprisingly large, it is actually similar to the values[26,27] found for the Kepler mission and is expected for the hypothesis that most of the 170,000 stars have planets.

Note that the predicted mission capability for non-HZ planets would increase if larger than Earth-size planets were considered. The choice that all non-HZ planets be Earth-size was made to ensure that the mission design emphasized Earth-size planets in and near the HZ as stated in the science goals (Sec. 2).

In contrast to the results for HZ-planets shown in Fig. 3, Fig. 4 indicates that the mission has the capability of detecting many non-HZ planets orbiting A- and F-type stars. This is the result of choosing the target stars based on brightness without regard for spectral type (i.e., Cases #0 and #1) and that HZ-planets orbiting such stars have orbital periods too long to show a minimum of three transits during a 4-year mission duration.

5 Analysis of Some Factors Controlling Science Results

5.1 *Number and Selection of Target Stars*

The number of expected detections and the accuracy of the estimates of the occurrence frequency and the distributions of planets with size and stellar spectral type depend on the number of useful stars that can be monitored with the required precision. Monitoring a large number of stars was required to get statistically meaningful results because the geometrical probability that a planet in a particular orbit will show transits is very low.[28] The probability is the diameter of the star divided by the diameter of the orbit: i.e., about 0.1 for planets with orbital periods of a few days and 0.005 for planets like that of the Earth orbiting in the HZ of a star like the Sun. Assuming that every star has a planet large enough to be detected and that the planet is in a properly aligned orbit in the HZ of a solar-like star, then several hundred planets should be detected based on monitoring 100,000 stars. During the development of the CSR, it was recognized that an increase in the instrument memory capability would allow the number to be increased to 170,000 stars. However, increasing the number of stars implies a reduction in

"time-on-target" due to a longer period required to orient the spacecraft toward the Earth to download the data. (Downlinks were performed once per month and took ~12 h to complete.)

Maximizing the number of detected planets requires maximizing the number of useful stars in a single FOV. Although the number of stars in a fixed FOV grows almost exponentially with the acceptable upper limit to the magnitude, only a fraction of the added stars is suitable for the search because some of these are too dim or too large to provide a SNR above the threshold value.

Prior to the approval of the mission for development, the choice of target stars for the Kepler mission was necessarily based only on apparent brightness because stellar sizes were unknown for most stars dimmer than ninth magnitude. However, after the start of mission development, members of the Kepler team made multi-band color measurements of 2 million of the stars in the selected FOV to determine their spectral type and surface gravity and thereby their size.[29] Consequently, it was possible to obtain a selection of target stars that maximized the detectability of small planets; i.e., to do "cherry-picking."

Cherry-picking consists of selecting those stars with the size and apparent brightness to provide the highest probability of showing a pattern of transits that are above the detection threshold for Earth-size planets in the HZ. Thus, stars with the lowest noise are not necessarily the best choice. For example, although two stars of equal apparent brightness but of dissimilar size have similar levels of shot noise, the smaller star (i.e., the one with later spectral type and lower mass) will show more transits for planets in the HZ during the duration of the mission and these transits will have larger amplitudes than those for the larger star. The probability of the geometrical alignment of the orbit with the line-of-sight and the duration of the transits also enter into the selection of the cherry-picked targets.

The MF contains an algorithm that could be selected to choose the most appropriate stars as a function of the number of stars monitored in a fixed FOV. (The optimized selection of stars for the actual mission was based on observations[30] of stars in the selected FOV.) In this study, the following factors were considered:

1. Stars considered that are too dim to give a detectable signal for an Earth-size planet can still produce detectable signals for transits of the larger planets.
2. For a given planet size, a small star will produce a larger relative signal than a larger star and thereby show detectable transit amplitudes even though it is dimmer than a somewhat larger and brighter star.
3. The probability of an orbital alignment is proportional to the size of a star.
4. The orbital periods for planets in the HZ of small (low-mass) stars are short, and therefore, provide many transits during the mission duration thereby increasing the SNR compared to the longer periods for HZ planets of larger, more massive stars.

The MF was run to estimate the expected results of observing for three selections of stars in a fixed FOV near the galactic plane. Table 1 shows results for the three cases: case #0: the selection of the 100,000 brightest stars between ninth and 15th magnitudes as specified in the original proposal; case #1: 170,000 brightest stars between ninth and 15.5 magnitudes; and case #2: 170,000 cherry-picked stars between ninth magnitude and 16.5 magnitude that were chosen to maximize detectability of small planets near the HZ.

Table 1 MF value, number of detectable planets in the HZ, and number of detectable planets not in the HZ.

Case	Maximum magnitude limit	Number of stars monitored	Calculated MF value	Number of detectable HZ planets ($Rp \leq 1.3\,R_{\oplus}$)	Number of $\sim 2\,R_{\oplus}$ planets in HZ	Number of detectable non-HZ planets
0	15	100,000 brightest	83	125	278	3205
1	15.5	170,000 brightest	100	156	442	4200
2	16.5	170,000 cherry-picked	140	255	788	3707

For cases #0 and #1, it was assumed that all stellar types were observed because the information on stellar type and size needed to choose the most appropriate stars was not available. For case #2, it was assumed that a ground-based survey to classify the stars was conducted prior to the launch of the mission. Thus, SpT earlier than F5 were rejected and replaced by later types. An examination of Table 1 shows that the MF value increased by 20% when the number of stars monitored are increased from 100,000 to 170,000 while the number of small planets ($Rp \leq 1.3\,R_{\oplus}$) in the HZ and planets not-in-the HZ increase by 25% and 31%, respectively. However, when cherry-picked target stars are chosen rather than the 170,000 brightest stars, the MF for small planets in the HZ increases by 40%, but the number of non-HZ planets decreases by 12%. The decrease in the number of non-HZ planets occurs because case#2 selects for slightly dimmer and smaller stars that can provide detections of $2\,R_{\oplus}$ planets in the HZ. This choice decreases the number of short-period planets that would have been detected around the brighter and larger stars. The mission capability to detect $2\,R_{\oplus}$ planets in the HZ increases by 280% from case #0 to case #2.

In Table 1, the number of $2\,R_{\oplus}$ planets are shown separately from the planets with radii $\leq 1.3\,R_{\oplus}$ because the larger planets might be too large to have a rocky composition.[25]

Table 1 can be used to examine the effects of changing the assumptions that scale linearly with the number of stars; e.g., the planetary occurrence rate and/or decreasing the size of the FOV. However, many of the expected mission results have nonlinear dependencies on the model assumptions. Therefore, model calculations are required to estimate their effects. Figures 3 and 4 show the results of some of these calculations.

Comparison of the numbers of detectable planets in columns 5 and 6 of the Table indicates that most of the increase in the number of HZ planets results from the inclusion of the "large" $2\,R_{\oplus}$ planets. These planets have 4 times the signal amplitude of the $1\,R_{\oplus}$ planets, and therefore, are much easier to detect among the many dim stars in the FOV. The capability of the mission to find several hundred planets in the HZ reduces the risk that the mission might not find any planets in the HZ if the occurrence frequency was as low as 1%.

The model results shown in Fig. 4 indicate that the number of detectable non-HZ Earth-size planets is much larger than those found in the HZ. (Note the greatly increased vertical scale.) Again, it is clear that a useful distribution of occurrence frequency versus semi-major axis and stellar type can be expected even if the frequency of small planets is 1% or less. Most of these are in short-period orbits that produce many transits, have a high probability of orbital alignment, and thereby have a high probability of detection. These calculations assume that all non-HZ planets are Earth size. Clearly many additional planets would be detectable if the assumption of larger size planets were allowed. The choice of only $1\,R_{\oplus}$ planet-size for non-HZ planets assures that the mission design emphasizes Earth-size planets in and near the HZ as stated in the science goals (Sec. 1). Although the expected number of non-HZ planet detections is quite large, the value is similar to that detected by the Kepler Mission; i.e., a total 4600 confirmed planets and candidates.[27] A comparison of the number of non-HZ planets presented in Table 1 shows that the selection of the dimmer, cherry-picked stars results in the detection of a slightly lower number of non-HZ planets.

Figures 5 and 6 show the results of model calculations for the value of the MF and numbers of exoplanets versus number of stars observed and for different selections of target stars.

The pair of dashed lines at the bottom of Fig. 5 imply that there are very few stars that are small and bright enough to provide an SNR sufficiently large to detect Mars-size planets in the HZ. However, when we consider Earth-size and larger planets, then much larger fractions of the stars show detectable transits because of the increased SNR. For these planets, increasing the number of observed stars significantly increases the number of expected detections regardless of the fact that many of the additional stars will necessarily be dimmer. The pair of solid lines in Fig. 5 shows that the mission capability to detect large terrestrial-size planets continues to increase well past the value of 100,000 stars that was assumed for the point design presented in the original proposal. The red curves show that a further gain in the number of HZ planets can be obtained by cherry-picking, i.e., determining the spectral type and luminosity class of each star in the FOV and then choosing only those that provide the largest transit signals. Much of the gain shown by the red curves relative to the blue curves is due to the replacement of early-type stars with later SpT.

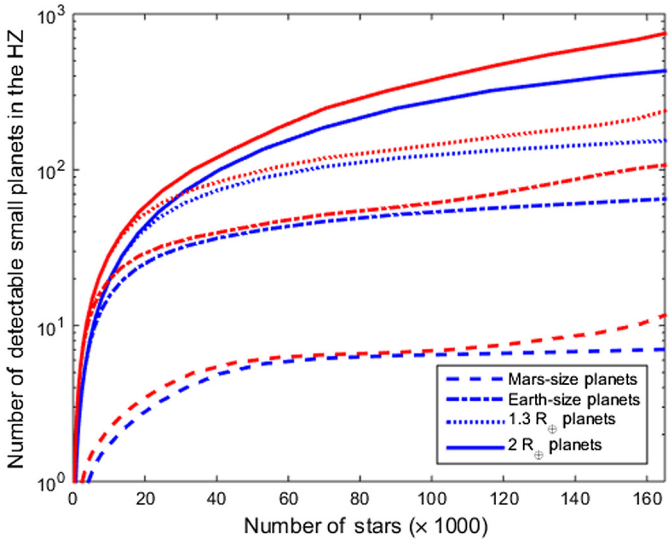

Fig. 5 Mission capability of detecting planets of various sizes in the HZ versus the number of stars observed for a fixed FOV. The blue lines of each pair show the results for the choice of the 170,000 brightest stars while the red upper line of each pair shows the results when 170,000 cherry-picked stars are chosen. The red line lies above the blue line because all early-spectral-type stars (i.e., large stars) have been removed from the cherry-picked selection.

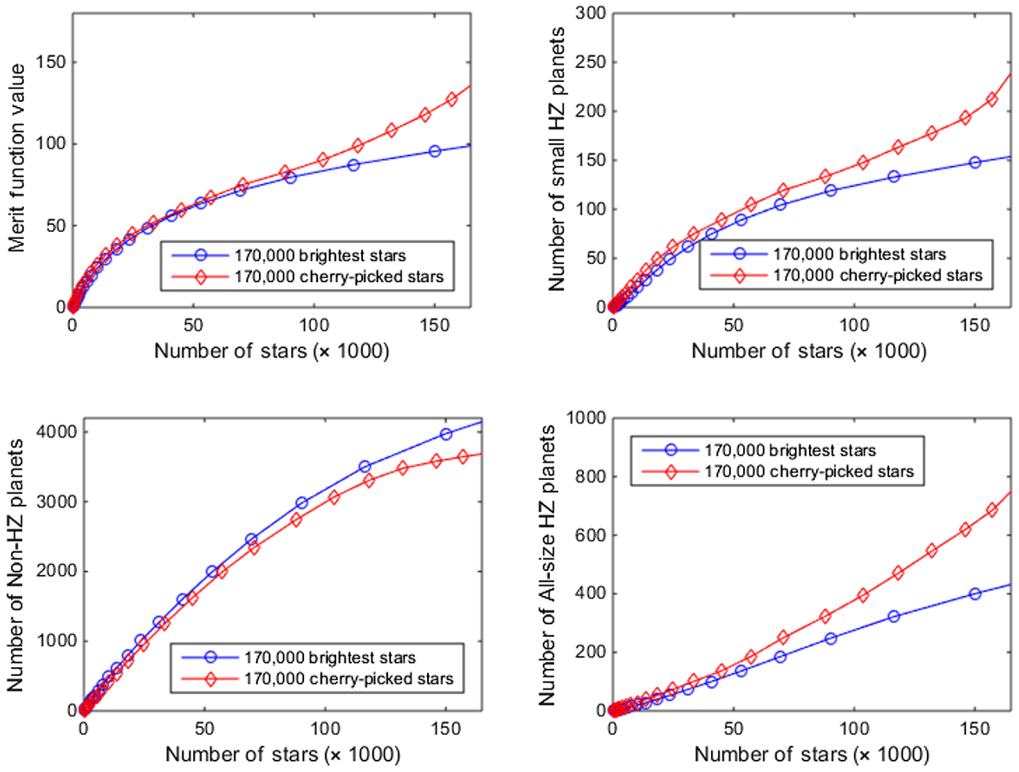

Fig. 6 Model results for the values of the MF, number of small planets ($Rp \leq 1.3\,R_{\oplus}$) in the HZ, total number of planets ($Rp \leq 2\,R_{\oplus}$) in the HZ, and the number of non-HZ planets versus the number of monitored stars. The results shown in Figs. 5 and 6 are based on a statistical table[16] for the number of stars versus size and magnitude.

Computing values for the MF for HZ planets at a particular value for the number of stars monitored consists of evaluating the number of stars that could show detections of the smallest planet size, multiplying that number by the science value for the appropriate planet size, and then summing the product of the additional number of planets at each size and value; $\mathrm{MF} = \sum(n_i - n_{i-1}) * v_i$, $i = 1$ to 4, where n_i is the number of detectable planets of each of the four planets sizes and $\mathrm{v_i}$ is the value of the detection for planets of that size (where $n_0 = 0$).

The curves in upper left-hand panel of Fig. 6 show the calculated MFs for three cases: case#0; monitor 100,000 brightest stars in the FOV, case#1: monitor 170,000 brightest stars in the FOV, and case#2: monitor 170,000 stars chosen that have the highest probability of showing detectable transits. (Note: to avoid saturating the detectors, only stars ninth magnitude and fainter were considered in all calculations.) Both the red and blue curves indicate that there is a substantial increase in the science product for increasing the number of monitored stars from the 100,000 in the original proposal to 170,000 in the CSR. The red curve indicates that a prelaunch program to classify all the stars in the FOV provides a substantial increase in the MF when the results are used to select the most promising stars.

Comparison of the curves showing the numbers of detectable planets in the upper- and lower-righthand panels of Fig. 6 shows that most of the increase (note the changed range of ordinate values) in the number of HZ planets results from the inclusion of the "large" 2 $R_\oplus$ planets. These planets have 4 times the signal amplitude of the 1 $R_\oplus$ planets, and therefore, are much easier to detect among the many dim stars in the FOV. The capability of the mission to find several hundred planets in the HZ reduces the risk that the mission might not find any small planets in the HZ if the occurrence frequency was as low as 1%.

In the lower left-hand panel of Fig. 6, the model results show that the number of detectable non-HZ Earth-size planets is much larger than those found in the HZ (it is worth noting the greatly increased vertical scale). It is clear that a useful distribution of occurrence frequency versus semi-major axis and stellar type can be expected even if the frequency of small planets is 1% or less. Most of these are in short-period orbits that produce many transits, have a high probability of orbital alignment, and thereby have a high probability of detection. These calculations assume that all non-HZ planets are Earth-size. Clearly many additional planets would be detectable if the assumption of larger size planets had been used. The choice of only 1 $R_\oplus$ planet-size for non-HZ planets assures that the mission design emphasizes Earth-size planets in and near the HZ as stated in the science goals (Sec. 1). Although the expected number of non-HZ planet detections is quite large, the value is similar to that detected by the Kepler Mission; i.e., a total of 4600 confirmed planets and candidates. The curves in the lower left panel also show that choosing the dimmer cherry-picked stars will result in the detection of a slightly lower number of non-HZ planets.

5.2 *Effects of Star Selection Criterion on Mission Capability*

Figures 7(a) and (b) compare the mission capabilities for two selections of target stars. Figures 7(a) shows the results for the selection of 170,000 brightest (case#1) stars including all SpT, while Fig. 7(b) shows the selection of 170,000 cherry-picked stars (case#2) that did not include early type stars. Selection of case#2 provides a dramatic increase in the mission capability to detect HZ planets for later spectral-type stars with a modest reduction of detections for G2-type stars. Choosing cherry-picked stars increased mission capability to detect both small planets and 2 $R_\oplus$ planets by 78%. A comparison of the results in Figs. 7(a) and 7(b) shows that a more uniform sampling of the number of planets versus spectral type is obtained when cherry-picked stars are selected. In fact, the Kepler mission did include a prelaunch activity[29] to classify the spectral type of stars in the FOV to provide the spectral classifications needed for case#2.

Figures 7(a) and 7(b) shows that no planets are detectable orbiting F7 stars. This result stems from the requirement that at least four transits be observed during a four-year mission and the fact that planets in the HZ of F7 and earlier (i.e., hotter) types have orbital periods too long to provide even 3 transits during the 4-year mission duration.

In contrast to what is expected for HZ planets, Fig. 8 shows that the detection of non-HZ planets orbiting stars hotter than G2-types can be expected. This occurs because planets in inner orbits will show a sufficient number of transits to satisfy the criteria of a minimum of four transits

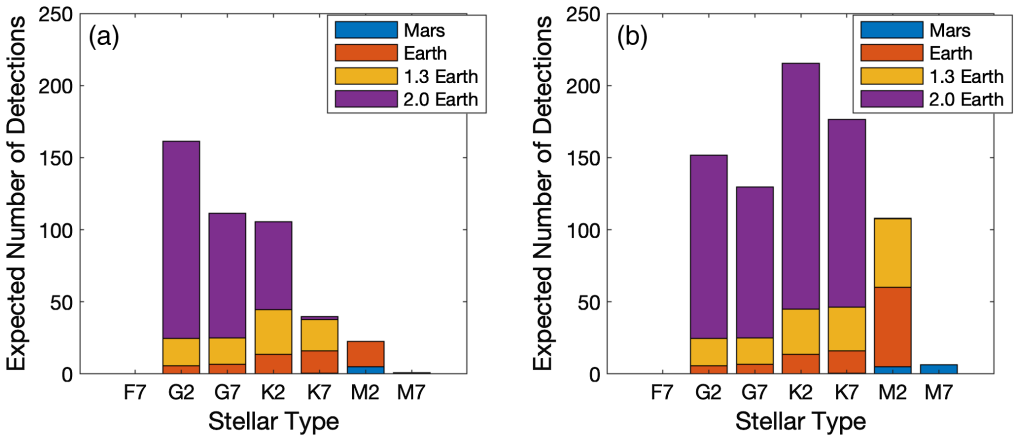

Fig. 7 Expected number of detections based on the instrument capability, characterization of the stars in the Kepler FOV, the assumptions of one terrestrial planet in each HZ, and a requirement of detecting a minimum of three transits. (a) Results for the selection of 170,000 brightest stars (case#1). (b) Results for 170,000 cherry-picked stars (case#2). The predicted numbers of detectable planets for each stellar type and for any of the chosen planet sizes (i.e., Mars-size, Earth-size, 1.3 Re, and 2 Re) are indicated by the sum of the heights of the appropriate bars.

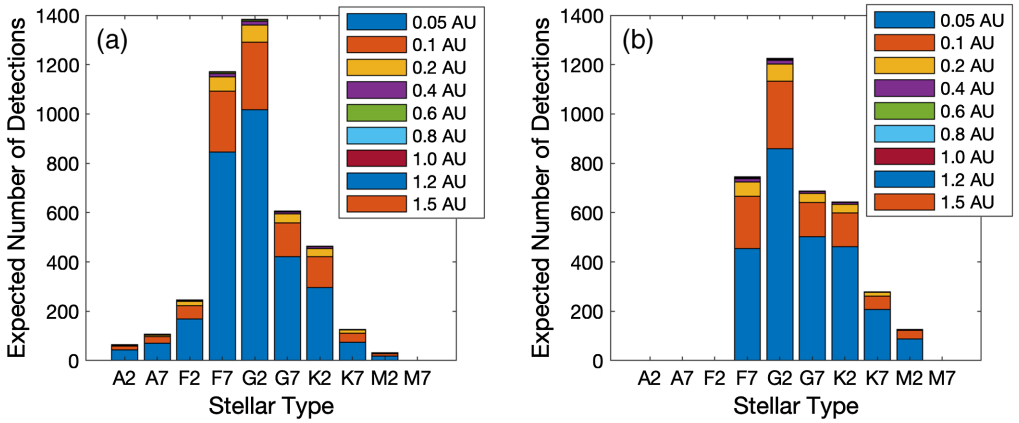

Fig. 8 Expected number of detections from non-HZ planets distributed over a range of semi-major axes for a mission duration of 4 years. It is worth noting the greatly enlarged values on the ordinate compared to those in Fig. 7. The coloring of the bars is related to the number of planets found at each semi-major axis for a given star type. (a) Case #1: 170,000 stars selected based on their brightness. (b) The expected detections for case #2: stars chosen based on the expected SNR of small-planet transits.

and provide an SNR greater than the detection threshold. It is clear from the length of the blue-colored segments as a fraction of the total height of each bar that the results are strongly biased toward the detection of innermost planets.

A comparison of Fig. 8(a) and 8(b) shows the effects of target star selection on the mission capability for defining the distribution of non-HZ planets. Similar to the results for HZ planets, smaller numbers of planets orbiting F7- and G2-type stars are expected for case#2, while the number of planet detections for later-type stars is enhanced.

The results shown in Figs. 7 and 8 show that the mission capability should be sufficient to detect several hundred terrestrial planets in the HZ and several thousand in the non-HZ regions if they are common. Even if only a fraction of the numbers calculated here were found, useful estimates of the occurrence frequency and distributions with planet size and semi-major axis would be obtained. Furthermore, if only null results were obtained, the results would still provide statistically meaningful upper bounds and would imply that Earth-sized planets in our galaxy are rare.

Although the discussion above demonstrates that the expected number of detections is large, the predictions depend not only on prevalence and characteristics of planetary systems, but on the actual performance of the instrument. Examples of some of the risks associated with instrument and mission characteristics are given next.

6 Results from Several Trade Studies and Risk Assessments

The MF proved invaluable both in guiding early design decisions and in later stages of program development in terms of managing cost and risk. For example, in response to cost growth during development, the MF supported evaluation of science impacts for two rounds of rapid (6 week) cost-cutting trade study efforts where the systems engineering team explored over 150 descope options. The existence of the MF enabled quantitative assessments of the individual and combined science impacts of multiple descope options on time scales that would have been impossible with traditional analysis tools. Additionally, the MF enabled quantitative risk assessments for late-breaking issues uncovered during ground testing of the Kepler observatory with profound implications for launch go/no-go decisions. Here, we summarize some representative cases of how the MF was used for these purposes.

6.1 *Effects of Reduced Mission Duration*

Reduction of the mission duration could result from major component failure or insufficient funding. Figure 9 shows the effects of reducing mission duration. A comparison of the upper panels shows that the reduction of the mission duration from 4 to 3 years would result in the loss of mission capability to detect Earth-size planets in the HZ of G2-dwarfs because too few transits would be observed. Because a major goal of the mission is the determination of small planets in the HZ of Sun-like stars (i.e., G-type stars), the loss of G2-type stars would cause a serious loss of science merit. In particular, the MF value would fall from 120 to 84 for case#1 and from 169 to

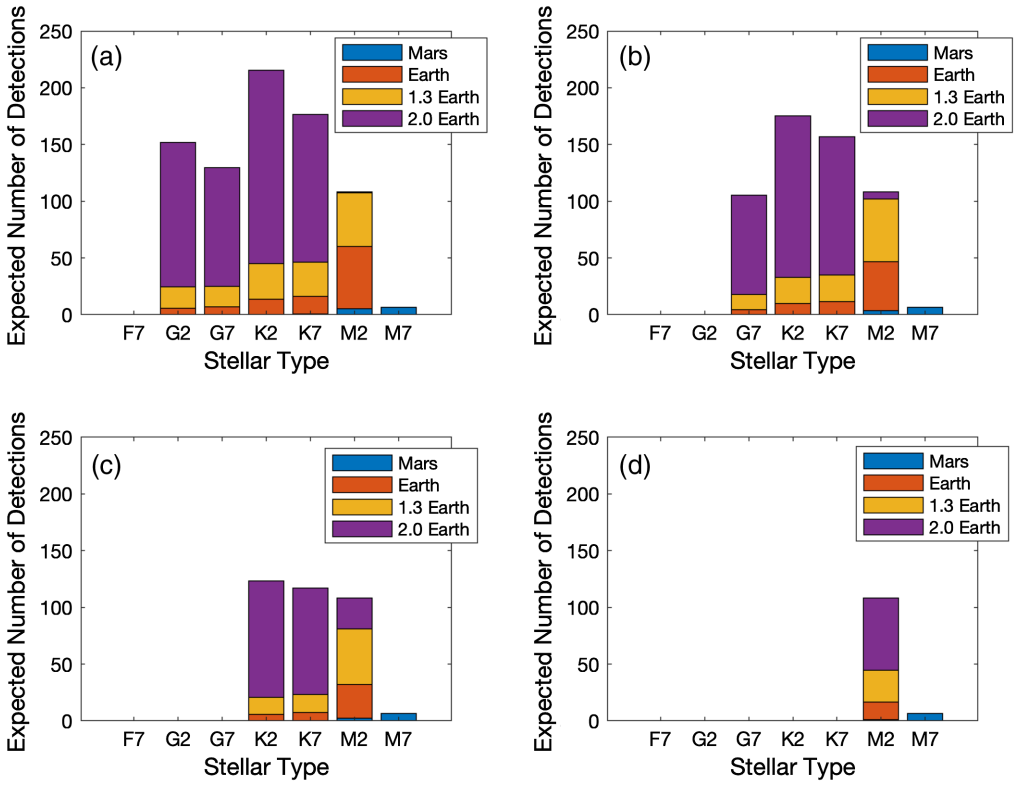

Fig. 9 Mission capability to detect HZ-planets when the mission duration is reduced from 4 years to 1 year for cherry-picked stars (case#2). (a) 4-, (b) 3-, (c) 2-, and (d) 1-year mission durations).

123 for case#2; i.e., the science value would be reduced by a third. Figure 9 shows that a two-year mission would be too short to detect planets orbiting any G dwarfs while a 1-year mission could be expected to be sensitive only to planets orbiting M-dwarfs.

Nevertheless, a mission duration of only 1 year would still be expected to produce nearly 100 detections of non-HZ planets. Thus, a null result from a severely curtailed mission duration would still be meaningful.

6.2 *Effects of Higher than Expected Instrument Noise*

As the value of the instrument noise increases, not only does the number of expected planetary detections decrease, but the relative occurrence frequency with respect to the stellar type changes; i.e., another bias is introduced. Figure 10 shows model results for G, K, and M SpT for values of the instrument noise level relative to the point design. Two sets of curves are shown. The curves shown in black represent case#1 and curves shown in red represent case #2. The vertical axis displays the fraction of all possible planet detections (i.e., the exoplanet population) for each star type.

The black curve marked by square symbols (case #1) shows that the mission has the capability of detecting all the small planets orbiting M-dwarfs until the instrument noise exceeds three times the point design value. However, the curve marked in red for case #2 indicates that the detected fraction of the total population decreases rapidly as the noise level exceeds the value for the point design. The difference in the results for cases #1 and #2 is caused by the selection of M-dwarfs as faint as 16.5 magnitude to maximize the number observed (~9400) for case #2 while observing only the ~2000 M-dwarfs that are brighter than 15.5 magnitude for case #1. As the instrument noise increases, it readily overwhelms the many faint stars associated with case #2. Nevertheless, an examination of an earlier figure (i.e., Fig. 7) shows that the number of small planets detectable in the HZ of M-dwarfs is several times larger for case# 2 compared to that for case #1.

The situation shown for K-dwarfs in Fig. 10 is analogous to that for the M-dwarfs in that a larger fraction of the population of planets orbiting K-dwarfs is detectable because case #1 selects only the brightest stars. Because case #2 selects for K-dwarfs as faint as 16th magnitude, the number of selected stars is larger (68,000 versus 23,000), but many of the planets transiting fainter stars for the case # 2 selection are not detectable at increased noise levels. The tripling of the detectable fraction for K-dwarfs shown in Fig. 7 is mostly due to the removal of A- and F-type dwarfs and their replacement by K-dwarfs when case#2 is selected.

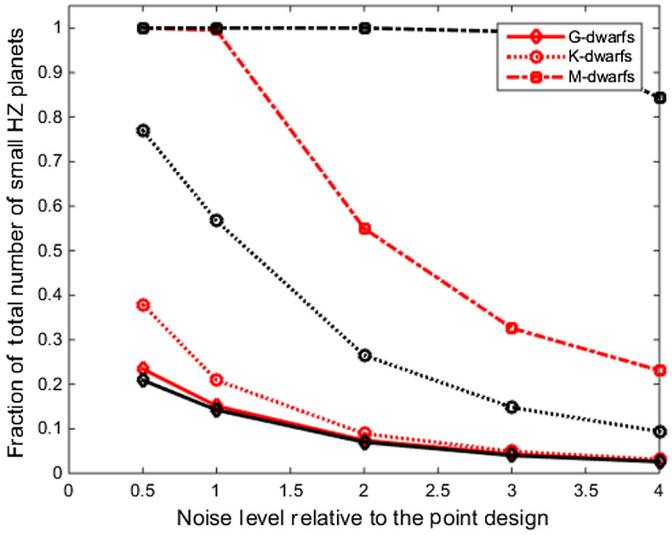

Fig. 10 Fraction of the population of small planets ($Rp \leq 1.3\,R_{\oplus}$) in the HZ that can be detected in the FOV. Black curves represent case#1, while red curves represent case#2.

The decrease in the fraction of the population of G-dwarfs with increasing instrument noise is nearly independent of the case number.

It should also be noted that changes in the instrument noise level cause changes in the relative number of planets with spectral type as seen in both Figs. 7 and 10.

6.3 *Effects of Raising the Detection Threshold*

During the mission design, investigations[17] using the MF were made regarding various system and stellar noise characteristics. However, the effects of nonwhite noise and astrophysical false-positive events were difficult to predict. The Kepler mission design assumed that a threshold of 7σ for detection would be satisfactory to reduce the number of statistical false alarms to ≤ 1. However, this value was not sufficient to avoid frequent false-positive (i.e., astrophysical events that mimicked a planetary transit) and false-alarm events (i.e., statistical noise). False-positive events[31] were often caused by images from eclipsing binary stars in or near the target star images. Both types of events often led to frequent false alarms that wasted valuable resources needed to confirm a planet.

Consequently, rather than following up all 7σ detections, often only those events with much higher values attracted the resources needed for follow up observations. Figure 11 shows the predicted effects on the detection rates as the required threshold level is increased from 7σ to 12σ. The results imply that many more planets remain to be detected if the false alarm and false-positive rates can be reduced sufficiently to allow 7σ detections to be confirmed.

6.4 *Effect of Reduced Field of View*

Kepler's wide (113 deg) FOV was key to observing the required population of stars. This necessitated the development of a large focal plane consisting of 42 science CCDs each with two output channels.[3,4] Additionally, managing the heat, power, mass, and volume of the focal plane array assembly required an innovative and complex set of electronics. Several issues were encountered during implementation that presented a risk of degraded FOV that was challenging to analyze with conventional tools. First, system testing revealed the presence of a design issue—a high frequency oscillation in signal chain amplifiers—that manifested as a "rolling band" and other time-varying artifacts[32] that impacted a significant number of CCD pixels with an associated reduction in useable FOV. The detection of this issue following focal plane and telescope integration required the team to evaluate the risk of accepting the performance as-is or pursue a

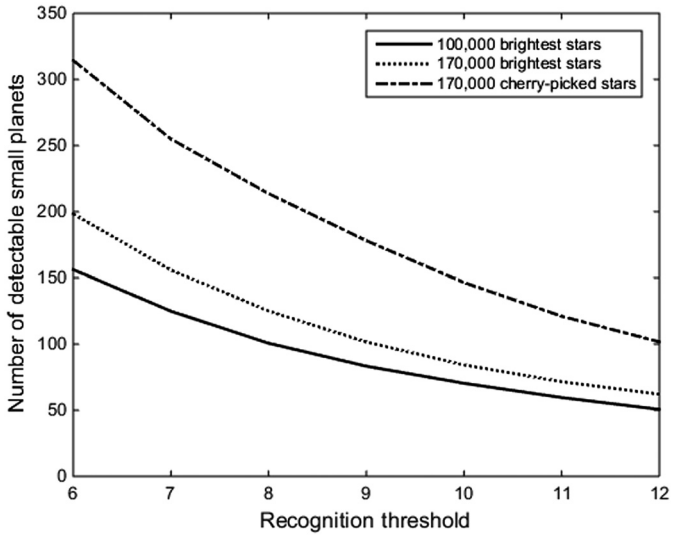

Fig. 11 Predicted number of detectable small planets ($Rp \leq 1.3\,R_{\oplus}$) versus the value of the recognition threshold. It is worth noting that a value of 7 was chosen for mission point design and the calculations in this paper.

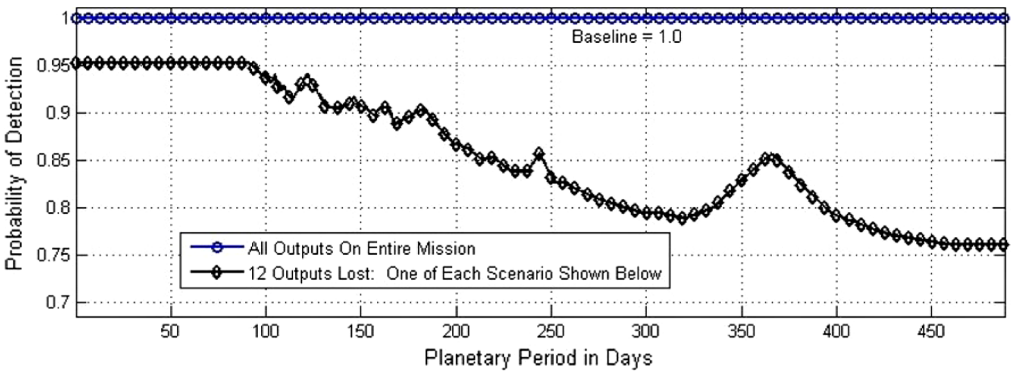

Fig. 12 Impact on detection probability at different planetary orbital periods for a scenario that assumed the loss of 12 CCD channels as follows: single channel, two adjacent channels, two opposite channels, three (asymmetric) channels, and all four channels in a signal chain group. On this plot, the mission performance floor for an Earth/Sun analog (365 day period) is 0.25 (off-scale here), indicating such a scenario would not have a devastating impact on science return.

design change and rework that would have required a costly launch delay. Second, other issues involving the reliability of electronic components were uncovered before launch, presenting a risk of losing one or more CCDs entirely. In both cases, the MF played a key role in assessing the potential impacts to science return. The focal plane CCDs were arranged in a symmetrical pattern to provide consistent viewing over the course of a year, during which the observatory completed four quarterly rolls around the telescope line of sight to main proper geometry with respect to the sun. Hence, the many permutations of FOV impacts from possible CCD artifacts or channel failures were additionally convolved with the quarterly rolls of the focal plane on the sky and the different orbital periods of transiting planets. It was essential to understand these combined impacts. To help bound the risk and guide decisions, the MF evaluated a series of scenarios such as the one shown in Fig. 12. Based on these assessments, the team decided to accept the systematic artifacts and the risk of in-flight failure of CCD channels (of which three ultimately failed during operations with minor impact on the FOV).

7 Brief Comparison of Model Predictions with Mission Results

Only a rough comparison with the Kepler mission results is possible because of the many differences between the MF that was designed to perform trade-off studies and risk assessments and the Kepler survey that made astrophysical measurements in the presence of the many confounding factors present.

It is clear from Table 1 and the mission results[26,27] that the model predictions for the occurrence rate of HZ-planets are much higher than that found by the Mission; i.e., 255 versus 23 for $Rp \leq 1.3\,R_{\oplus}$, and 778 versus 104 for $Rp \leq 2\,R_{\oplus}$. However, it is worth noting that the mission found an additional 184 planets larger than $2\,R_{\oplus}$ in the HZ. Such situations would exclude smaller planets from the HZ. A comparison of the number of the predicted number of small ($Rp \leq 1.3\,R_{\oplus}$) planets in the HZ with those actually discovered shows that the model assumption that every star has such a planet in its HZ overpredicts the observed number by a factor of 11. The prediction is a factor of 7 for the larger, more easily recognized $2\,R_{\oplus}$ planets. Assuming that half the target stars were binary and that these seldom provided detectable planets in the HZ, compensation for this factor would double the fraction of small planets detected (i.e., 23/255) to ~46/255. (The mission detected only 10 planets orbiting binary stars compared to total of 4700 detected planets.) This result implies that the ratio of stars that have small planets in their HZ is of the order of 20%. This value is midrange to the values in the literature; 1%,[33,34] ~10% to 40%,[35–37] and 100%.[38] The wide range of estimates is due to both the uncertainty in the statistics of the observations and to the varied assumptions with respect to ranges of planet size and orbital period chosen for their computations.

The MF predicted number (i.e., ~3700) for non-HZ planets is similar to that found by the Earth-size planets with semi-major axes ≤2 AU, where the similarity between the predicted and

observed values might be a fortuitous result caused by the assumption of 3 planets in each planetary system.

A major reason for the differences in the expectations from the MF and the results from the Kepler mission was that the model necessarily assumed distributions of planet sizes, orbital distance, and occurrence frequency before these were determined by the Kepler observations. The model could not have calculated the relative merit of mission trade-offs without some numerical values for these parameters. However, prior to the data from Kepler, the values of these parameters were unknown. Thus, the occurrence frequency assumed by the MF was based on the theory that the Solar System was typical; i.e., that all other planetary systems would also have planets and that each would have an Earth-size planet in its HZ. When the proposal was written, everyone was aware that this assumption was likely to be optimistic and that only the observations could give us realistic estimates of the actual values. However, any other assumptions would not have been consistent with the then-current knowledge.

Another important reason for the differences was the need to place many small planets throughout a large range of semi-major axes to delineate the occurrence frequency near the HZ of many stellar types. This assumption was based on the need to estimate the capability of the mission to estimate the structure of planetary systems; especially near the HZ.

Other reasons for the differences include:

1. Around 50% of the target stars in the Kepler survey were probably binary or multiple-star systems. The nonrepetitive transit patterns from planets orbiting binary stars were much more difficult to detect than those from single stars.[39] Furthermore, the secondary star diluted the transit depth. Both effects would cause transit patterns to be missed.
2. Stellar variability[40,41] and system noise in the Kepler survey were much higher than expected and than was assumed here.
3. The high false-positive rate due to many faint stars in or near target apertures required treating transit patterns with an SNR much $< \sim 12$ with caution.
4. The model results for small planets included many planets orbiting very faint stars that could not be detected by the follow-up program.
5. The presence of planets larger than $2\,R_{\oplus}$ in or near the HZ would prevent small planets considered by the MF from occupying the HZ.

It should be noted that the MF was not intended to serve as a model for science interpretation and hypothesis testing, but rather an analytic framework to support mission performance prediction, design trades, and risk assessments by quantifying the key sensitivities; and by that latter standard, it worked as intended.

8 Summary

The MF is an algorithm designed to produce numerical values for the science return given a set of inputs representing an instrument/spacecraft point design, estimates of the stellar structure of galaxy, exoplanet size and frequency distributions, and instrument noise. Examples of the consequences of mission risks were explored.

Several examples demonstrated the usefulness of the MF; in particular, the results show that:

1. Increasing the number of stars observed from the 100,000 brightest stars in the original mission proposal to 170,000 substantially increases the mission capability to detect both HZ- and non-HZ planets by 59% and 31%, respectively).
2. Additionally, a prelaunch stellar observation activity to select the most promising targets (cherry-picking) further increases the number of discoveries and better defines the occurrence frequency and the size distribution of exoplanets. In particular, the mission capability to detect small planets in the HZ of K-type dwarfs nearly triples when stars are selected based on the expected SNR. For planets not in the HZ, cherry-picking the selection of targets stars changes the observed distribution of planets versus stellar spectral type by decreasing the number of planets associated with early-type stars while increasing those orbiting later-type stars.

3. Calculations showed that the risk that is incurred for shortened mission durations does not negate the mission capability to place a useful estimate of the occurrence frequency of Earth-size planets even for a mission duration as short as one year. However. reduction of the mission duration below 3 years to shorter periods does cause a severe loss with respect to the determination of the occurrence frequency for stars most like the Sun. This strongly influenced the adoption of 3 years as the floor for mission duration. Conversely, calculations showed that operating 5 years or more significantly improves science return and that influenced engineering decisions to ensure propellant capacity sufficient 8+ years operation, which ultimately enabled the extended mission ($K2$) following the failure of two reaction wheels.
4. Evaluation of the risk that a noise level twice that stated in the CSR would prevent meeting mission goals indicated that the capability would be reduced by ~2, but mission capability would still provide a useful number of planetary detections.
5. Even when the false-alarm and false-positive rates become so large that the threshold for signal recognition must be raised to almost double the point design, the expected discovery rate is still sufficiently high to accomplish most mission goals.
6. The science return is robust to loss of effective FOV resulting from unusable pixels due to electronic-induced image artifacts and in-flight hardware failures.

The MF described here represents a tool to assess the change in the science capability of a mission when the assumptions about unknowns are varied. Risks associated with uncertainties in model parameters, instrument performance, and the consequences of mission trade-offs can be assessed.

The strengths of the MF approach include the ability to explore the effects of the many unknowns in first-of-its-kind missions and to assess the effects of mission changes during development and operation. Weaknesses include omission of the effects of false-positive events, of the effects of various approaches to data analysis, and of the limitations of the ground assets needed for confirmations. Effective use of the science MF on Kepler required an awareness of these factors when evaluating the outputs of this tool and applying them to decision making. The same would apply to the potential development and use of science MFs for other space missions.

9 Appendix A

This appendix describes how the Merit Function estimates the SNR of a transit given the white Gaussian noise models for shot noise and residual pointing noise, and the power spectral density of solar variability observed by SOHO.[21] The mathematical development is similar to that for Kepler's transiting planet search algorithm described in Ref. [22].

Consider the problem of detecting a transit pulse signal, $\boldsymbol{s} = s(n)$, $n = 1, \ldots, N$, in a set of observations, $\boldsymbol{x} = x(n)$, $n = 1, \ldots, N$, where N is the length of the time series, and boldface indicates vectors. Further, suppose that $x(n)$ is composed of three noise sources:

- Solar variability, characterized as a nonwhite Gaussian noise process by a correlation matrix, R;
- Shot noise, characterized as a zero-mean, white Gaussian noise process, σ^2_{shot};
- Residual pointing noise, characterized as a zero-mean, white Gaussian noise process, $\sigma^2_{\text{pointing}}$.

Under these conditions, the optimal detector is a noise-compensating matched filter of the form:

$$l = \frac{\boldsymbol{x}^T R^{-1} \boldsymbol{s}}{\sqrt{\boldsymbol{s}^T R^{-1} \boldsymbol{s}}}, \tag{1}$$

where l is known as the detection statistic, which is compared against a threshold, η, to decide whether the signal is present or not.[42] That is, the signal is considered present when $l \geq \eta$ and considered absent when $l < \eta$.

Reference 43 derived a version of Eq. (1) for the special case where the observation noise is wide-sense stationary:

$$l = \frac{\int_{-\pi}^{\pi} \frac{X(\omega)S^*(\omega)\mathrm{d}\omega}{P(\omega)}}{\sqrt{\int_{-\pi}^{\pi} \frac{S(\omega)S^*(\omega)\mathrm{d}\omega}{P(\omega)}}}, \tag{2}$$

where $X(\omega)$ is the discrete-time Fourier transform of $x(n)$, $S(\omega)$ is the Fourier transform of the transit signal, $P(\omega)$ is the power spectral density of the observation noise, and $(\cdot)^*$ denotes the conjugate operator.

Under the null hypothesis (that there is no signal present), l is a zero-mean, unit variance random deviate. Under the alternative hypothesis, l is a unit variance random deviate with an expected value given by

$$\langle l \rangle = \frac{\int_{-\pi}^{\pi} \frac{\langle X(\omega) \rangle S^*(\omega)\mathrm{d}\omega}{P(\omega)}}{\sqrt{\int_{-\pi}^{\pi} \frac{S(\omega)S^*(\omega)\mathrm{d}\omega}{P(\omega)}}} = \frac{\int_{-\pi}^{\pi} \frac{S(\omega)S^*(\omega)\mathrm{d}\omega}{P(\omega)}}{\sqrt{\int_{-\pi}^{\pi} \frac{S(\omega)S^*(\omega)\mathrm{d}\omega}{P(\omega)}}} = \sqrt{\int_{-\pi}^{\pi} \frac{S(\omega)S^*(\omega)\mathrm{d}\omega}{P(\omega)}}. \tag{3}$$

The term, $\langle l \rangle$, is interpreted as the expected SNR of the signal waveform against the observation noise.

In order to deal with non-stationary observation noise,[22] developed an adaptive, wavelet-based approach to jointly estimating the characteristics of the noise process and detecting exoplanet transit signatures. The Merit Function employs a simpler algorithm for evaluating the expected SNR [i.e., Eq. (3)] based on Kepler's transiting planet search algorithm. First, Debauchies' 12-tap wavelet filter bank is used to expand the observed noise time series into a dyadic octave-band wavelet decomposition.[44] For the Merit Function, the average power of the solar variability time series is calculated in each of M pass bands, $P_{sol} = [P_{sol}(m), m = 1, \ldots, M]^T$. The total noise power in each band is obtained by adding the variances of the shot and residual pointing noise to the solar variability noise power in each band: $P_{\mathrm{total}} = [P_{\mathrm{sol}}(m) + \sigma^2_{\mathrm{shot}} + \sigma^2_{\mathrm{pointing}}, m = 1, \ldots, M]^T$. The transit pulse (of duration τ) is subjected to a similar wavelet decomposition and the average energy in each passband is calculated: $E_{\mathrm{transit}}(\tau) = [E_{\mathrm{transit}}(m, \tau), m = 1, \ldots, M]^T$. To evaluate the SNR, the integral in Eq. (3) is evaluated as

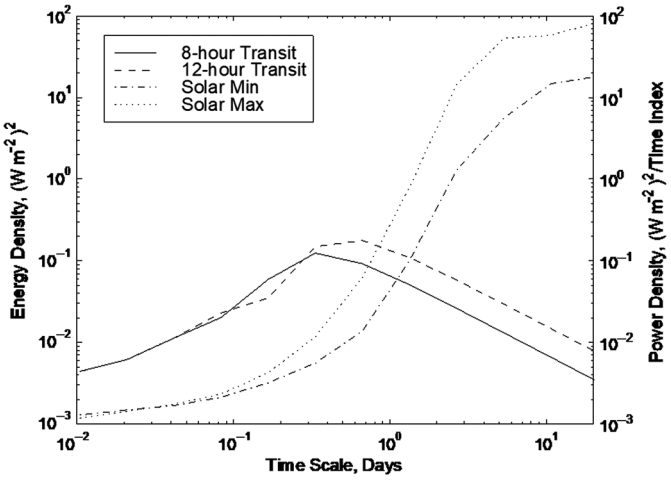

Fig. 13 Distribution of power as a function of timescale from a wavelet analysis of the time series of solar irradiance as measured by the DIARAD instrument aboard SOHO for the years of 1996, near solar minimum (dash-dotted curve), and for 2000, near solar maximum (dotted curve). The distribution of energy with timescale is also plotted for an Earth-size, 8-h transit (solid curve) and for a 12-h transit (dashed curve). The areas under the transit curves and above the solar variability curves indicate that the transits are detectable against the solar variations.[22]

$$\mathrm{SNR} = A_{\mathrm{transit}} \sqrt{\sum_{m=1}^{M} \frac{N_{\mathrm{transit}} E_{\mathrm{transit}}}{P_{\mathrm{sol}}(m) + \sigma_{\mathrm{shot}}^2 + \sigma_{\mathrm{pointing}}^2}}, \tag{4}$$

where N_{transit} is the number of observed transits, and A_{transit} is the transit depth. Note that Eq. (4) can be calculated efficiently by precalculating E_{transit} on a grid of transit durations to obtain a table of transit energy vectors which can be interpolated for any duration within the range of tabulated values.

Equation (3) indicates that the detectability of a signal against a background noise source is determined by timescales (or frequencies) at which the energy in the signal exceeds the noise power. Figure 13 illustrates this by providing a comparison of the frequency response of an 8-h and 12-h long Earth-size transits against the power spectral densities (PSDs) of the first year of DIARAD observations of the Sun for 1996 and the year 2000.

10 Appendix B. Outline of Merit Function Requirements and Design Requirements

1. Provide a numerical value proportional to the science value of the expected results.
2. Normalize the score such that the total value will be 100 points when all the requirements stated in the CSR are met.
3. Evaluate planets (HZ-planets) and planets not in the HZ ("non-HZ planets") separately. Assign more weight to the first group and less to the second in alignment with their importance to the goals of the mission. For the Kepler Mission, the score for HZ-planets received a weight of 0.65 while the non-HZ planets received a weight of 0.35.
4. Assign values for subsets of the goals and weight their importance. For the Kepler Mission, the weights for the capability to detect a small planets in the HZ were set much higher than for larger planets because that choice ensured that all larger planets could also be detected. The score was set higher for planets with large SNR transit patterns because they produce a higher recognition probability and a lower false-positive probability. For planets not in the HZ, the score for planets orbiting close to their star is less than those at larger semi-major axis to ensure a more uniform determination of the occurrence frequency versus semi-major axis.
5. The inputs should include instrument and mission parameters that affect the model results. For example: (a) photon flux (which is affected by the aperture, transmission and bandwidth of optics and filters, and the efficiency of the detectors.); (b) mission duration; and (c) size and sky location of the FOV.
6. The score for expected discoveries that do not make a substantial contribution to the accomplishment of mission goals or which are accomplished irrespective of mission choices should be set to zero to force the mission emphasis on the stated goals.

10.1 *Outline of Merit Function Design*

Set constants and evaluate input tables

1. Mass, size, and occurrence frequency of each star type.
2. Visual magnitudes to be considered: 9-16.5.
3. Frequency/amplitude of stellar/solar variability.
4. Photometer, read, and jitter noise values from CSR Table F5.

Two components are considered:

1. "Planets in HZ": Range of planet sizes in HZ of each star type.
2. Non-HZ planets: all planets are 1 $R_{\oplus}$ in size and distributed with logarithm of semi-major axis with values near 1 AU regardless of star type.

10.1.1 *Component #1. Planets in HZ*

Input:

- Planet sizes (relative to Earth). Mars, Earth, 1.3 Earth, and 2 Earth
- Score values for MF:

  40 for each recognized planet in the HZ for planet radius $Rp = 0.53\,R_{\oplus}$,
  20 for each recognized planet in the HZ for $Rp = 1\,R_{\oplus}$,
  5 for each additional planet with $Rp = 1.3\,R_{\oplus}$,
  1 for each additional planet with $Rp = 2\,R_{\oplus}$.

Calculations:

- Calculate the number of transits/star for 4- and 6-year mission durations as a function of stellar type and semi-major axis.
- Calculate the SNR for 4-year and 6-year mission durations for all magnitudes, SpT, and planet sizes.
- Calculate the probability of a transit for a planet in the HZ of each star type.
- Count the number of stars that have the correct orbital alignment versus spectral type and apparent magnitude and count the number of stars that have three or more transits and a detection threshold value of 7σ or greater. Include a correction for the recognition rate as function of threshold value.
- Compute the score for the chosen parameters.

10.1.2 *Component #2. Non-HZ planets*

Procedure is similar to component #1 with the following exceptions.

- The size of all planets $= 1\,R_{\oplus}$.
- There exists one planet at each of nine semi-major axes, but do not count planets that have a value of the semi-major axis that duplicates one in the HZ of that star type.
- Multiply the results by the expected number of non-HZ planets by 2/9 to get a total of three planets in each planetary system.

Acknowledgments

The authors would like to thank Gentry Lee of JPL for his suggestion that an algorithm be developed to quantify the sensitivities between system parameters and the expected science output. We would also like to acknowledge the help with program development from Jeff Garside and Santanu Das and for reviewers' suggestions from Sally Cahill and the anonymous reviewers. The Kepler Mission was the tenth Discovery Mission. The work was supported by the NASA Headquarter Science Mission Directorate. The NASA Ames Research Center and Jet Propulsion Laboratory managed the mission. The Ball Aerospace and Technology Corporation built the instrument and spacecraft. Many organizations contributed to Mission success, including the Laboratory for Space Physics, SETI Institute, Smithsonian Astrophysical Observatory, Lawrence Hall of Science, California Institute of Technology, University of California Lick and Keck Observatories, Space Telescope Science Institute, University of Texas McDonald Observatory, and Roque de los Muchachos Observatory in La Palma. Portions of this work were undertaken at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA.

References

1. W. J. Borucki et al., "The Kepler Mission: a mission to determine the frequency of inner planets near the habitable zone for a wide range of stars," in *Planets Beyond the Solar System and the Next Generation of Space Missions, Proc. Workshop held at Space*

Telescope Science Institute, Baltimore, Maryland, D. R. Soderblom, Ed., ASP Conference Series, Vol. 119, p. 153 (1997).
2. D. Koch et al., "Overview and status of the Kepler Mission," *Proc. SPIE* **5487**, 1491–1500 (2004).
3. D. G. Koch et al., "Kepler mission design, realized photometric performance, and early science," *Astrophys. J. Lett.* **713**, L79–L86 (2010).
4. W. J. Borucki, "Kepler mission: development and overview," *Rep. Prog. Phys.* **79**, 036901 (2016).
5. G. W. Henry et al., "Properties of sun-like stars with planets: 51 Pegasi, 47 Ursae Majoris, 70 Virginis, and HD 1147621," *Astrophys. J.* **474**, 503 (1997).
6. T. M. Brown and D. Charbonneau, "The STARE project: a transit search for hot Jupiters," in *Disks, Planetesimals, and Planets, ASP Conf. Proc.*, G. Garzon et al., Eds., Vol. 219, 584–589 (2000).
7. D. Charbonneau et al., "Detection of planetary transits across a Sun-like Star," *Astrophys. J.* **529**, L45–L48 (2000).
8. W. J. Borucki et al., "The Vulcan photometer: a dedicated photometer for extrasolar planet searches," *Publ. Astron. Soc. Pac.* **113**, 439–451 (2001).
9. T. M. Brown et al., "HST time-series photometry of the transiting planet of HD 209458," *Astrophys. J.* **552**, 699–709 (2001).
10. G. W. Wetherill, "Occurrence of Earth-like bodies in planetary systems," *Science* **253**, 535–538 (1991).
11. J. Lissauer, "Planet formation," *Annu. Rev. Astron. Astrophys.* **31**, 129–172 (1993).
12. G. W. Wetherill, "The formation and habitability of extra-solar planets," *Icarus* **119**, 219–238 (1996).
13. A. Morbidelli et al., "Building terrestrial planets," *Annu. Rev. Earth Planet. Sci.* **40**, 251–275 (2012).
14. M. Mayor and D. Queloz, "First discovery (RV) of exoplanet around a main-sequence star," *Nature* **378**, 355 (1995).
15. G. W. Marcy and R. P. Butler, "Detection of extrasolar giant planets," *Annu. Rev. Astron. Astrophys.* **36**, 57–97 (1998).
16. W. J. Borucki et al., "Finding Earth-size planets in the habitable zone: the Kepler Mission," *Proc. Int. Astron. Union* **3**(S249), 17–24 (2007).
17. W. Borucki et al., "Kepler: search for Earth-size planets in the habitable zone," *Proc. Int. Astron. Union* **4**(S253), 289–299 (2008).
18. T. Schmidt-Kaler, "Physical parameters of the stars," in *Astronomy and Astrophysics – Stars and Star Clusters*, K. Schaifers and H. H. Vogt, Eds., p. 1, Landolt-Bornstein New Series (2b), Springer-Verlag, New York (1982).
19. A. C. Robin et al., "A synthetic view on structure and evolution of the Milky Way," *Astron. Astrophys.* **409**, 523 (erratum: Astron. Astrophys., 416:157 (2004) (2003).
20. N. M. Batalha et al., "Stellar variability and its implications for photometric planet detection with Kepler," in *Proc. First Eddington Workshop Stellar Struct. and Habitable Planet Finding*, ESA SP-485, B. Pattrick, Ed., pp. 35–40, ESTEC, Noordwijk, The Netherlands (2002).
21. C. Fröhlich et al., "In-flight performance of the VIRGO solar irradiance instruments on SOHO," *Sol. Phys.* **175**(2), 267–286 (1997).
22. J. M. Jenkins, "The impact of stellar variability on the detectability of transiting terrestrial planets," *Astrophys. J.* **575**, 493–505 (2002).
23. J. E. Kasting, D. P. Whitmire, and R. T. Reynolds, "Habitable zones around main sequence stars," *Icarus* **101**, 108 (1993).
24. J. E. Chambers and G. W. Wetherill, "The terrestrial planets: N-body integrations of embryos in three dimensions," *Icarus* **136**, 304–327 (1998).
25. G. W. Marcy et al., "Masses, radii, and orbits of small Kepler planets: the transition from gaseous to rocky planets," *Astrophys. J. Suppl. Ser.* **210**, 20 (2014).
26. J. L. Coughlin et al., "Planetary candidates observed by Kepler. VII. The first fully uniform catalog based on the entire 48-month data set (Q1–Q17) DR 24," *Astrophys. J. Suppl. Ser.* **224**, 12 (2016).

27. S. E. Thompson et al., "Planetary candidates observed by Kepler. VIII. A fully automated catalog with measured completeness and reliability based on data release 25," *Astrophys. J. Suppl. Ser.* **235**, 38T (2018).
28. W. J. Borucki and A. L. Summers, "The photometric method of detecting other planetary systems," *Icarus* **58**, 121–134 (1984).
29. T. M. Brown et al., "Kepler input catalog: photometric calibration and stellar classification," *Astron. J.* **142**,112 (2011).
30. N. M. Batalha et al., "Selection, prioritization, and characteristics of Kepler target stars," *Astrophys. J. Lett.* **713**, L109 (2010).
31. T. M. Brown, "Expected detection and false-alarm rates for transiting Jovian planets," *Astrophys. J.* **593**, L125–L128 (2003).
32. D. A. Caldwell et al., "Instrument performance in Kepler's first months," *Astrophys. J. Lett.* **713**, L92–L96 (2010).
33. J. Catanzarite and M. Shao, "The occurrence rate of Earth analog planets orbiting Sun-like stars," *Astrophys. J.* **738**, 151 (2011).
34. A. Silburt, E. Gaidos, and Y. Wu, "A statistical reconstruction of the planet population around Kepler solar-type stars," *Astrophys. J.* **799**, 180 (2015).
35. E. A. Petigura, A. W. Howard, and G. W. Marcy, "Prevalence of Earth-size planets orbiting Sun-like stars," *Proc. Natl. Acad. Sci. U.S.A.* **110**(48), 19273–19278 (2013).
36. C. D. Dressing and D. Charbonneau, "The occurrence of potentially habitable planets orbiting M dwarfs estimated from the full Kepler dataset and an empirical measurement of the detection sensitivity," *Astrophys. J.* **807**, 45 (2015).
37. G. D. Mulders et al., "The exoplanet population simulator. I. The inner edges of planetary systems," *Astron. J.* **156**, 24 (2018).
38. W. A. Traub, "Kepler exoplanets: a new method of population analysis," arXiv:1605.02255 (2016).
39. D. Windemuth et al., "An automated method to detect transiting circumbinary planets," *Mon. Not. R. Astron. Soc.* **490**, 1313–1324 (2019).
40. R. L. Gilliland et al., "Kepler mission stellar and instrument noise properties," *Astrophys. J. Suppl. Ser.* **197**, 6 (2011).
41. G. Basri, L. M. Walkowicz, and A. Reiners, "Comparison of Kepler photometric variability with the Sun on different timescales," *Astrophys. J* **769**, 37 (2013).
42. H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*, pp. 19–155, 239–442, John Wiley and Sons, New York (1968).
43. S. Kay, "Adaptive detection for unknown noise power spectral densities," *IEEE Trans. Signal Process.* **47**, 10 (1999).
44. I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.* **41**, 909 (1988).

William J. Borucki retired from NASA Ames in 2015 but continues as a NASA Ames research associate. He received BS and MS degrees in physics from the University of Wisconsin (Madison) in 1960 and 1962, respectively, and an MS degree in meteorology from California State University (San Jose, California) in 1982. He is the author of more than 225 papers. During his 58 years at Ames, he contributed to the design of the Apollo heat shield, modeling of the effects of nitrogen and chlorine pollutants on the Earth's ozone layer, lightning activity in planetary atmospheres, and modeling the electrical properties of the Titan atmosphere. He is a co-investigator on the Huygens Entry Probe and is the principal investigator for the Kepler mission.

Jon M. Jenkins is a data scientist in the supercomputing division at NASA Ames Research Center. He received his bachelor's degree in electrical engineering (EE), a bachelor's degree in applied mathematics, an MS in EE, and a PhD in EE from the Georgia Institute of Technology in 1987, 1988, 1988, and 1992, respectively. He developed the science pipelines as the co-investigator for data processing on NASA's Kepler and Transiting Exoplanet Survey Satellite missions.

Riley M. Duren is a research scientist at the University of Arizona's Institutes for Resilience and Engineering fellow at NASA's Jet Propulsion Laboratory. His has worked on nine space missions (with the Shuttle Radar Topography Mission and Kepler as personal favorites). From 2008 to 2019, he served as chief systems engineer for JPL's Earth Science Directorate. His primary research today involves multi-scale carbon monitoring systems to support climate change mitigation.