PREDICTING THE YIELDS OF PHOTOMETRIC SURVEYS FOR TRANSITING EXTRASOLAR PLANETS

Thomas G. Beatty[1] and B. Scott Gaudi[2]

ABSTRACT

We develop a method for predicting the yield of transiting planets from a photometric survey given the parameters of the survey (nights observed, bandpass, exposure time, telescope aperture, locations of the target fields, observational conditions, and detector characteristics), as well as the underlying planet properties (frequency, period and radius distributions). Using our updated understanding of transit surveys provided by the experiences of the survey teams, we account for those factors that have proven to have the greatest effect on the survey yields. Specifically, we include the effects of the surveys' window functions, adopt revised estimates of the giant planet frequency, account for the number and distribution of main-sequence stars in the survey fields, and include the effects of Galactic structure and interstellar extinction. We approximate the detectability of a planetary transit using a signal-to-noise ratio (S/N) formulation. We argue that our choice of detection criterion is the most uncertain input to our predictions, and has the largest effect on the resulting planet yield. Thus, drawing robust inferences about the frequency of planets from transit surveys will require that the survey teams impose and report objective, systematic, and quantifiable detection criteria. Nevertheless, with reasonable choices for the minimum S/N, we calculate yields that are generally lower, more accurate, and more realistic than previous predictions. As examples, we apply our method to the Trans-Atlantic Exoplanet Survey, the XO survey, and the *Kepler* mission. We discuss red noise and its possible effects on planetary detections. We conclude with estimates of the expected detection rates for future wide-angle synoptic surveys.

Subject headings: methods: numerical — planetary systems — surveys — techniques: photometric

Online material: color figures

1. INTRODUCTION

There are four ways by which extrasolar planets have been detected. The first method to unambiguously detect an extrasolar planet was pulsar timing, which relies on detecting periodic variations in the timing of the received radio signal that occur as the pulsar orbits about the system's barycenter. The first system of three planets was found around PSR B1257+12 in 1992 (Wolszczan & Frail 1992), followed by a single planet around PSR B1620−26 (Backer et al. 1993). Although rare, the pulsar planets are some of the lowest mass extrasolar planets known: PSR B1257+12a is about twice the mass of the Moon.

The second way to find extrasolar planets is through radial velocities (RV), which uses the Doppler shift of observed stellar spectra to look for periodic variations in the target star's radial velocity. By estimating the mass of primary star, the observed radial velocity curve and velocity semiamplitude can then be used to directly calculate the inclination-dependent mass ($M_p \sin i$) of the companion object. To date, RV surveys have detected more than 230 planets around other stars, making it the most successful method of extrasolar planet detection. While the large number of detected systems having an unseen companion with a mass on the order of 1 $M_J \sin i$ statistically ensures that the majority of these are planetary bodies, the RV surveys are unable to provide more information than the minimum masses, periods, eccentricities, and the semimajor axes of the planets.

RV surveys also have a limited ability to detect planets much smaller than a few Earth masses. The state of the art in RV surveys is the High Accuracy Radial Velocity Planet Searcher (HARPS) spectrometer at the La Silla Observatory in Chile, which is capable of radial velocity measurements with precisions better than 1 ms^{-1} for extended periods of time (Lovis et al. 2006). HARPS is therefore able to detect planets with masses on the order of 3−4 $M_\oplus$ in relatively short period orbits. Unfortunately, planets closer to an Earth mass will be increasingly difficult to detect since intrinsic stellar variability, in the form of acoustic oscillation modes and granulations on the photosphere, makes more precise spectroscopic radial velocity measurements harder to acquire. However, it may be possible to surmount this obstacle (as in the case of Lovis et al. 2006) through the selection of stars with "quiet" photospheres and long integration times which serve to average out the stellar variability.

Gravitational microlensing is another technique for detecting extrasolar planets. Microlensing of a star occurs when a star passes near the line of sight of the observer to another background star. The gravity of the foreground star acts as a lens on the light emitted by the background star, which causes the star in the background to become momentarily brighter as more light is directed toward the observer. Planetary companions to the lens star can further magnify the background star, and create short-term perturbations to the microlensing light curve. To date, six planets have been detected using microlensing (Bond et al. 2004; Udalski et al. 2005; Beaulieu et al. 2006; Gould et al. 2006b; Gaudi et al. 2008). Unfortunately, the one-shot nature of microlensing observations means that information about systems discovered this way is generally sparser than that available for RV systems. Therefore, microlensing is most useful in determining the general statistical properties of extrasolar planets, such as their frequency and distribution, and not the detailed properties of the planetary systems.

Planetary transits are a fourth method by which extrasolar planets have been discovered, and the one that provides the most complete set of information about the planetary system. Only planets with very specific orbital characteristics have a transit visible from Earth, because the orbital plane has to be aligned to within a few degrees of the line of sight. Therefore transiting

[1] Department of Physics and MIT Kavli Institute, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139; tbeatty@space .mit.edu.

[2] Department of Astronomy, The Ohio State University, 140 W. 18th Avenue, Columbus, OH 43210.

planets are rare. Nevertheless, a transiting extrasolar planet offers the opportunity to determine the mass of the planet (when combined with RV measurements), since the inclination is now measurable, as well as the planetary radius, the density, the composition of the planetary atmosphere, the thermal emission from the planet, and many other properties (see Charbonneau et al. [2007], for a review). In addition, and unlike RV surveys, transiting planets should be readily detectable down to 1 $R_\oplus$ and beyond, even for relatively long periods.

Having accurate predictions of the number of detectable transiting planets is immediately important for the evaluation and design of current and future transit surveys. For the current surveys, predictions allow the operators to judge how efficient are their data-reduction and transit-detection algorithms. Future surveys can use the general prediction method that we describe here to optimize their observing setups and strategies. More generally, such predictions allow us to test different statistical models of extrasolar planet distributions. Specifically, as more transiting planets are discovered and characterized, predictions relying on incorrect statistics of extrasolar planet properties will increasingly diverge from the observed set of distributions.

Using straightforward estimates it appears that observing a planetary transit should not be too difficult, presuming that one observes a sufficient number of stars with the requisite precision during a given photometric survey. Specifically, if we assume that the probability of a short-period giant planet (as an example) transiting the disk of its parent star is 10%, and take the results of RV surveys which indicate the frequency of such planets is about 1% (Cumming et al. 2008), together with the assumption that typical transit depths are also about 1%, the number of detections should be $\approx 10^{-3} N_{\leq 1\%}$, where $N_{\leq 1\%}$ is the number of surveyed stars with a photometric precision better than 1%.

Unfortunately, this simple and appealing calculation fails. Using this estimate, we would expect that the TrES survey, which has examined approximately 35,000 stars with better than 1% precision, to have discovered 35 transiting short period planets. But, at the date of this writing, they have found four. Indeed, overall only 51 transiting planets have been found at this time by photometric surveys specifically designed to find planets around bright stars.[3] This is almost 100 times less than what was originally predicted by somewhat more sophisticated estimates (Horne 2003).

Clearly then, there is something amiss with this method of estimating transiting planet detections. Several other authors have developed more complex models to predict the expected yields of transit surveys. Pepper et al. (2003) examined the potential of all-sky surveys, which was expanded on and generalized for photometric searches in clusters Pepper & Gaudi (2005). Gould et al. (2006a) and Fressin et al. (2007) tested whether the OGLE planet detections are statistically consistent with radial velocity planet distributions. Brown (2003) was the first to make published estimates of the rate of false positives in transit surveys, and Gillon et al. (2005) model transit detections to estimate and compare the potential of several ground- and space-based surveys. As has been recognized by these and other authors, there are four primary reasons why the simple way outlined above of estimating surveys yields fails.

First, the frequency of planets in close orbits about their parent stars (the planets most likely to show transits) is likely lower than RV surveys would indicate. Recent examinations of the results from the OGLE-III field by Gould et al. (2006a) and Fressin et al. (2007) indicate that the frequency of short-period Jovian worlds is on the order of 0.45%, not 1.2% as is often assumed by

extrapolating from RV surveys (Marcy et al. 2005a). Gould et al. (2006a) point out that most spectroscopic planet searches are usually magnitude limited, which biases the surveys toward more metal-rich stars, which are brighter at fixed color. These high metallicity stars are expected to have more planets than solar-metallicity stars (Santos et al. 2004; Fischer & Valenti 2005).

Second, a substantial fraction of the stars within a survey field that show better than 1% photometric precision are either giants or early main-sequence stars that are too large to enable detectable transit dips from a Jupiter-sized planet (Gould & Morgan 2003; Brown 2003).

Third, robust transit detections usually require more than one transit in the data. This fact, coupled with the small fraction of the orbit a planet actually spends in transit, and the typical observing losses at single-site locations due to factors such as weather, create low window probabilities for the typical transit survey in the majority of orbital period ranges of interest (von Braun & Ciardi 2007).

Lastly, requiring better than 1% photometric precision in the data is not a sufficient condition for the successful detection of transits: identifiable transits need to surpass some kind of a detection threshold, such as a signal-to-noise ratio (S/N) threshold. The S/N of the transit signal depends on several factors in addition to the photometric precision of the data, such as the depth of the transit and the number of data points taken during the transit event. In addition, ground-based photometry typically exhibits substantial autocorrelation in the time series data points, on the timescales of the transits themselves. This additional red noise, which can come from a number of environmental and instrumental sources, substantially reduces the statistical power of the data (Pont et al. 2006).

In this paper we describe a method to statistically calculate the number of planets that a given transiting planet survey should be able to detect. We have done so in a way that maximizes the flexibility of our approach, so that it may be applied in as many circumstances as possible. Therefore, we have allowed for survey-specific quantities such as the observation bandpass to be defined arbitrarily, and have also allowed for our astrophysical assumptions to be easily altered. We account for factors such as the low frequency of Jovian planets, the increased difficulty of detecting transits at high impact parameters, as well as variations in stellar density due to Galactic structure and extinction due to interstellar dust. We also use the present-day mass function to realistically account for the number of main-sequence stars that will be in a given field of view. We do not treat, or simulate, the distribution and effects of giant stars. We approximate the detectability of a planetary transit using a S/N formulation.

This differs from the previously published work in the field by virtue of the generalized nature of our approach. Other authors, such as Pepper et al. (2003), Gillon et al. (2005), Gould et al. (2006a), and Fressin et al. (2007) have restricted themselves to modeling specific transit surveys or types of transit surveys. Our approach, which allows for variations such as arbitrary telescope parameters, observing bandpasses, cadences, and fields of view, is not restricted to any one type of transit survey. Indeed, we offer examples of how our method may be applied to the four most frequently used modes for transit searches. Furthermore, we have conducted an extensive literature review of basic astrophysical relations, such as the present-day mass function (PDMF), Galactic structure, and the main-sequence mass-luminosity relation, and included these in our predictions. We do not use more complicated (but presumably more accurate) Galactic models like the Besancon model (Robin & Creze 1986) because these models are less flexible and considerably more time consuming to utilize.

---

[3] As of 2008 June. See the Extrasolar Planets Encyclopedia at http://exoplanet.eu for an up-to-date list.

We first describe the mathematical formalism with which we have chosen to address this problem, and then move on to discuss our specific assumptions. As examples, we then offer our predictions for several different types of photometric surveys: point-and-stare, drift-scanning, space-based, and all-sky. For each, we choose a representative survey for that particular mode, and compare our predictions to actual survey results, or to predictions elsewhere in the literature. We also examine the possible effects of red noise on our predictions, and conclude by discussing what our predictions imply for the general state of photometric transiting planet surveys.

2. GENERAL FORMALISM

In the most general sense, we can describe the average number of planets that a transit survey should detect as the probability of detecting a transit multiplied by the local stellar mass function, integrated over mass, distance, and the size of the observed field (described by the Galactic coordinates (l, b):

$$\frac{d^6 N_{\text{det}}}{dR_p \, dp \, dM \, dr \, dl \, db} = \rho_*(r, l, b) r^2 \cos b \, \frac{dn}{dM} \, \frac{df(R_p, p)}{dR_p \, dp}$$
$$\times P_{\text{det}}(M, r, R_p, p), \quad (1)$$

where $P_{\text{det}}(M, r, R_p, p)$ is the probability that a given star of mass M and distance r orbited by a planet with radius R_p and period p will present a detectable transit to the observing setup. The quantity $df(R_p, p)/(dR_p \, dp)$ is the probability that a star will possess a planet of radius R_p and period p. The term dn/dM is the present day mass function in the local solar neighborhood, and ρ_* is the local stellar density for the three-dimensional position defined by (r, l, b). We use $r^2 \cos b$ instead of the usual volume element for spherical coordinates, $r^2 \sin \phi$, because b is defined opposite to ϕ: $b = 90°$ occurs at the pole.

First, we will integrate over both planetary radius and orbital period

$$\frac{d^4 N_{\text{det}}}{dM \, dr \, dl \, db} = \rho_*(r, l, b) r^2 \cos b \, \frac{dn}{dM}$$
$$\times \int_{R_{p,\min}}^{R_{p,\max}} \int_{p_{\min}}^{p_{\max}} \frac{df(R_p, p)}{dR_p \, dp} P_{\text{det}}(M, r, R_p, p) \, dR_p \, dp. \quad (2)$$

To do this, we must first specify the detection probability P_{det}, as well as the distribution of planet radius and periods $df(R_p, p)/dR_p \, d_p$, which we treat in § 2.2.

2.1. The Detection Probability P_{det}

The probability that a planet orbiting a given star shows a detectable transit may be broken up into two separate probabilities: the chance that a planet with radius R_p and period p around a star of mass M at a distance r will transit with a sufficient signal-to-noise ratio to be detected, and the window probability that a transit will be visible for the particular observing setup:

$$P_{\text{det}}(M, r, R_p, p) = P_{\text{S/N}}(M, r, R_p, p) P_{\text{win}}(p) \quad (3)$$

2.1.1. $P_{\text{S/N}}$

To calculate the probability that a planetary system will show detectable transits, we first select the appropriate statistical test to determine whether or not our notional data shows a discernible transit. In the world of transiting planet searches, many different methods are used either individually or together to test the statistical significance of a possible transit and alert the researcher

to its existence. For example, the Hungarian-made Automated Telescope Network (HATNet) uses a combination of the dip-significance parameter (DSP) and a box least-squares (BLS) power spectrum (among others) to test its photometric data. In general although, whatever test one uses, it is generally the S/N of a transit shape in the observed data is what determines whether or not that transit will be identified as an event.

As the square of the signal-to-noise can be approximated by the χ^2 test, we therefore chose to use the χ^2 value to determine if a given transit will be detectable for a set of given observation parameters. For a specific transit, the χ^2 statistic, assuming uncorrelated white noise,[4] is given by

$$\chi^2 = N_{\text{tr}} \left(\frac{\delta}{\sigma} \right)^2. \quad (4)$$

Where N_{tr} is the number of data points observed in a transit of fractional depth δ and with fractional standard deviation σ, assuming a boxcar transit shape with no ingress or egress. To be detected, the transit must have a χ^2 higher than a certain minimum $\chi^2_{\min}$

$$\chi^2 \geq \chi^2_{\min}. \quad (5)$$

The fractional depth of the transit is given by the square of the planet-to-star radius ratio

$$\delta = \left(\frac{R_p}{R_*} \right)^2. \quad (6)$$

While this formulation ignores the contributions of stellar limb-darkening to the depth of the transit, Gould et al. (2006a) demonstrate that, to a first order approximation, transit detection rates are not affected by limb-darkening. Briefly, they point out that the effects of limb-darkening on the transit light curve will be offset by the effects of the planet's ingress and egress from the stellar disk. Therefore, we have not included stellar limb-darkening terms in our calculations.

To determine σ, we assume Poisson statistics, such that for a given exposure with an observed number of N_S source photons from the target star and N_B background photons,

$$\sigma = \sqrt{\frac{N_S + N_B}{N_S^2} + \sigma_{\text{scint}}^2}, \quad (7)$$

where σ_{scint} accounts for scintillation as per Young (1967) and N_S is defined as

$$N_S = e_\lambda F_{S,\lambda} t_{\text{exp}} A, \quad (8)$$

where e_λ is the efficiency of the overall observing setup, and can take values from 0 to 1. It accounts for photon losses in places such as the filter, the mirrors of the telescope, the CCDs, and passage through the atmosphere. The value $F_{S,\lambda}$ is the source photon flux in the observed bandpass in units of $\gamma \, \text{m}^{-2} \, \text{s}^{-1}$. The term t_{exp} is the exposure time of each observation, and A is simply the light collecting area of the telescope.

The observed photon flux from the source in a certain bandpass, $F_{S,\lambda}$, can be defined as a function of the photon luminosity

---

[4] We consider red noise later, in § 4.5.

of the source, the distance to the source, and the effects of extinction due to interstellar dust:

$$F_{S,\lambda} = \frac{\Phi_\lambda(M)}{4\pi r^2} 10^{-A_\lambda/2.5}, \tag{9}$$

where $10^{-A_\lambda/2.5}$ accounts for extinction, and will be treated later. It has the effect of decreasing the number of observed photons from more distant stars, particularly in the bluer bands.

The photon luminosity of the source star, $\Phi_\lambda(M)$, we set as a function of the stellar mass M and the observational bandpass λ. We derive this quantity in Appendix A and only note here that we approximate the source stars as blackbodies. While this is a simplification, we show in Appendix A that it is sufficient for our purposes. Using $\Phi_\lambda(M)$ in equation (9) yields the observed photon flux at a distance r from that same star.

We thus now have an expression for N_S as a function of the stellar mass, distance, and bandpass choice, that we can use in our calculation of σ. We now only need N_B, the number of background photons observed. Similar to equation (8), N_B is defined as

$$N_B = S_{\text{sky},\lambda} \Omega t_{\exp} A, \tag{10}$$

where $S_{\text{sky},\lambda}$ is the photon surface brightness of the sky at the particular wavelength λ (or, in our case, for a specific band), and Ω is the effective area of the seeing disk. Assuming that the point-spread function (PSF) is a Gaussian with a full width half-maximum (FWHM) of Θ_{FWHM} arcseconds, Ω is defined by

$$\Omega = \frac{\pi}{\ln 4} \Theta_{\text{FWHM}}^2. \tag{11}$$

We can now place both of the expressions for N_S and N_B back into equation (7) to arrive at a determination of σ, the photometric uncertainty in the data points.

In certain cases, N_S and N_B will be so great that the CCD detectors of the telescope will be saturated by the number of photons received from a star during the exposure time. In this case, further measurements become useless. Mathematically, we may describe saturation as occurring when the number of photons collected by a single CCD pixel during an exposure exceeds the full-well depth (FWD) of that pixel:

$$N_{\text{pix}} \geq N_{\text{FWD}}. \tag{12}$$

N_{pix} is related to N_S and N_B as

$$N_{\text{pix}} = N_S \left(1 - \exp\left[-\ln 2 \left(\frac{\Theta_{\text{pix}}}{\Theta_{\text{FWHM}}} \right)^2 \right] \right)$$
$$+ \frac{N_B \Theta_{\text{pix}}^2}{\Omega}, \tag{13}$$

where Θ_{pix} is the angular size of an individual pixel.

The only part of equation (4) that remains to be defined is the number of data points observed while the system is in transit, N_{tr}. Assuming a circular orbit, we can say that the number of points seen in transit is directly proportional to the fraction of time that the system spends in eclipse, and given a total number of observed data points N_{tot}, N_{tr} is thus

$$N_{\text{tr}} = N_{\text{tot}} \left(\frac{R_*}{\pi a} \right) \sqrt{1 - b^2}, \tag{14}$$

where b is the impact parameter of the transit in units of the stellar radius, $b \equiv (a/R_*)\cos i$, i is the orbital inclination, and we

have assumed that $R_* \ll a$. We note that this does not hold for correlated noise, see Aigrain & Pont (2007).

Taking equation (14) and putting it back into our original formulation for χ^2 gives

$$N_{\text{tot}} \left(\frac{R_*}{\pi a} \right) \sqrt{1 - b^2} \left(\frac{\delta}{\sigma} \right)^2 \geq \chi_{\min}^2. \tag{15}$$

We define a new variable, χ_{eq}^2, as the value of equation (15) when $b = 0$, corresponding to the χ^2 of an equatorial transit about a given star. This gives us a new form that makes it easier to determine when a system will have a transit geometry sufficient to be detected by the given observing setup:

$$\chi_{\text{eq}}^2 \sqrt{1 - b^2} \geq \chi_{\min}^2. \tag{16}$$

Rearranging, we find that we are able to detect transits up to a $b_{\max}$ of

$$b_{\max}(M, r, p) = \sqrt{1 - \left(\frac{\chi_{\min}^2}{\chi_{\text{eq}}^2} \right)^2}. \tag{17}$$

Therefore, $b_{\max}$ is a function of M, r, and p as a result of the relations contained within the χ_{eq}^2 term.

Finally, we can now deduce an explicit expression for $P_{\text{S/N}}$. Given that detectable transits can only occur between $0 \leq b \leq b_{\max}$, we may simply integrate the normalized probability density function of allowable impact parameters from 0 to $b_{\max}$ to find a value for $P_{\text{S/N}}$.

Given that $b \equiv (a/R_*)\cos i$, the maximum value of the impact parameter for a circular orbit (different from the maximum at which we will detect transits) is $b = a/R_*$. The normalization for the impact parameter's probability density function is therefore

$$1 = \int_0^{a/R_*} P_b(b)\, db. \tag{18}$$

Since impact parameters are uniformly distributed,[5] this gives the solution $P_b = R_*/a$. It should be noted that this form of P_b is only valid in the case of circular orbits. For the period ranges probed so far by transit surveys, this is not a very restrictive assumption, since only $\sim 10\%$ of the known transiting planets have eccentricities inconsistent with zero. At longer periods, however, the effects of eccentricity must be considered (Barnes 2007; Burke 2008).

If we now integrate this over the range of impact parameters that yield detectable transits, we find

$$P_{\text{S/N}}(M, r, R_p, p) = \int_0^{b_{\max}} \frac{R_*}{a}\, db$$
$$= \frac{R_*}{a} \sqrt{1 - \left(\frac{\chi_{\min}^2}{\chi_{\text{eq}}^2} \right)^2}. \tag{19}$$

---

[5] The distribution of $\cos i$ is uniform, which follows from the assumption that planetary systems are randomly oriented. The distribution of $\cos i$ may be derived by considering the normalized angular momentum vector of the planetary orbit, which is free to point in any direction, and thus has a parameter space that forms an imaginary sphere around the parent star. Particular values of $\cos i$ are "hoops" along the surface of this sphere, and the ratio of their area to that of the total sphere is the probability that $\cos i$ will take that value.
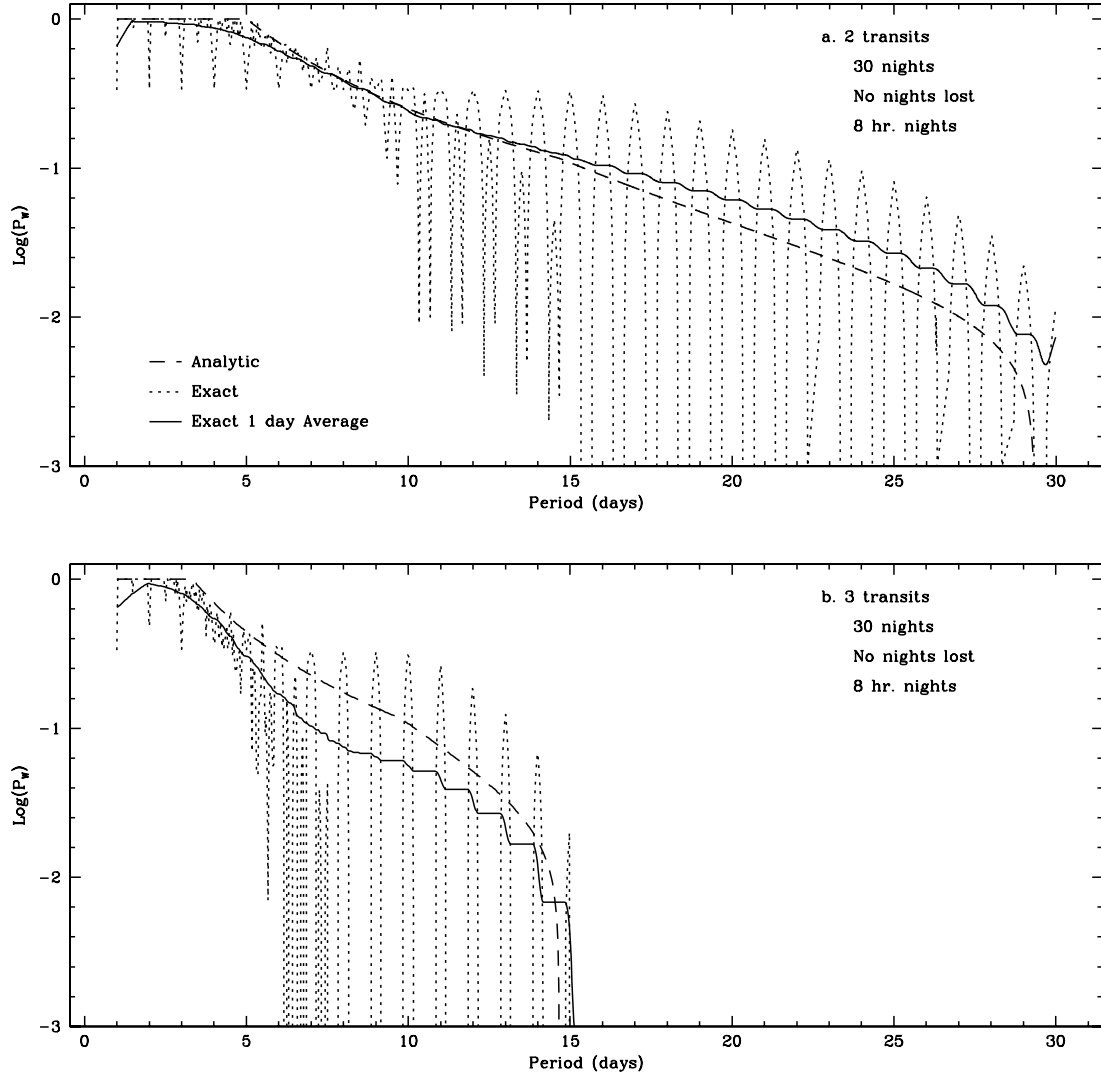
FIG. 1.—Comparing the analytic and exact formulations of the window probability. *Top*: Window probability for detecting two transits over a 30 day observing run with no nights lost and 8 hr of observing per night. *Bottom*: Same situation as top, but is instead for a minimum of three transits. [*See the electronic edition of the Journal for a color version of this figure.*]

2.1.2. P_{win}

The window probability (or window function) for a transit survey describes the probability that a transiting planet with a certain period will be visible to the survey, given the observational cadence of the survey (i.e., X hours of observing per night for Y nights). While many numerical routines exist that can exactly calculate the window probability for a given observing setup, for our purposes calculating an exact numerical answer is generally too time intensive to be practical.

We therefore developed an analytical framework that allows for the approximate calculation of a window probability. This has the virtue of requiring much less computing time. This is gained at the cost of a small loss of accuracy.

We give a full description of our analytic window probability in Appendix B. Briefly, we assume randomly sampled observations, which leads to simple expressions for the average window probability. This implicitly ignores the effects of aliasing at integer and half-integer day periods. Thus we are not able to address issues related to nonuniform or biased a posteriori distributions of planet parameters resulting from the effect of aliasing (e.g., Pont et al. 2006). Nevertheless, the analytic solution does follow

the "outline" of the exact calculations, and closely tracks the moving average value of the exact calculations sufficiently accurately for our purposes, as is shown in Figure 1.

2.2. The Frequency of Planets $df(R_p, p)/(dR_p\, dp)$

The term $df(R_p, p)/(dR_p\, dp)$ denotes the probability that a given star has a planet with radius R_p and orbital period p. In the example surveys that we consider later, we chose to focus on the detection of short period very hot and hot Jupiters (VHJs and HJs) with orbital periods of 1 to 5 days, as this is the regime being probed by these surveys. The exception of this focus is the *Kepler* survey, where we also examine the detectability of transiting planets over a range of radii and orbital periods. In both cases, we consider a fixed planet radius, so that

$$\frac{df(R_p, p)}{dR_p\, dp} = k(p)f(p)\delta(R_p - R_p'),\qquad (20)$$

where $k(p)$ is the correct normalization for a given period range, $f(p)$ is the distribution within this range, and R_p' is the planet

radius. For VHJs and HJs we use $R'_p = 1.1R_J$, the average radius of the transiting planets found to date.[6] In general, we have left the exact distribution of $f(p)$ to be assigned as is deemed appropriate. For the VHJ and HJ period distributions described above, we adopt a distribution that is uniform in period, $f(p) = 1$.

The frequency of hot Jupiters ($P \leq 5$ days, $M_p = 0.1-10\ M_J$) is commonly taken to be ~1% from the results of the RV planet surveys (e.g., Marcy et al. 2005a). However, adopting frequencies from RV surveys can lead to biased estimates of the expected yields from transit surveys, because RV surveys constrain the frequency of planets as a function of mass, rather than radius. Furthermore, as argued by Gould et al. (2006a) most RV surveys are more metallicity biased than transit surveys, and therefore the planet frequency found by RV surveys will be correspondingly higher, since planet occurrence is an increasing function of metallicity (Santos et al. 2004; Fischer & Valenti 2005). RV surveys are metallicity biased because the target samples for these surveys are magnitude limited. Since metal-rich stars are brighter at fixed color, they are overrepresented in magnitude-limited samples with respect to a volume-limited sample (Fischer & Valenti 2005). Indeed, Marcy et al. (2005b) note that because they use a magnitude-limited target list, their target stars have a definite metallicity bias. Transit surveys, on the other hand, are typically signal-to-noise ratio limited. At fixed color, metal-rich stars are both brighter and bigger. These two effects on the signal-to-noise ratio roughly cancel, and as a result transit surveys are very weakly biased with respect to metallicity (see § 8.1 of Gould et al. [2006a] for a quantitative discussion.)

Indeed, there is evidence for the effects of this metallicity bias in the comparison of various surveys. From the California and Carnegie Keck magnitude-limited and metallicity-biased (Fischer & Valenti 2005; Marcy et al. 2005b) survey sample, Cumming et al. (2008) find a frequency of ~1.3% for planets with $P \leq 5$ days and $M \geq 0.2\ M_J$. On the other hand, for the volume-limited (Udry et al. 2000b) CORALIE sample, Udry & Santos (2007) report a frequency of ~0.8% for $M \geq 0.2\ M_J$ and $P \lesssim 11.5$ days ($a < 0.1$ AU). Assuming that ~75% of these planets have $P \leq 5$ days, this corresponds to a frequency of 0.6% for $P \leq 5$ days and $M \geq 0.2\ M_J$, a factor of ~2 lower than that found for the Keck RV survey. Gould et al. (2006a) examined the OGLE-III transit survey, and found a frequency of ~0.5% for $P \leq 5$ days, which was subsequently confirmed and strengthened by (Fressin et al. 2007). Thus, as would be expected based on the arguments above, the frequencies inferred from transit surveys and volume-limited RV surveys are similar (and entirely consistent considering Poisson uncertainties and uncertainties in the mass-radius relation of giant planets), whereas they are substantially lower than the frequencies inferred from magnitude-limited (and metallicity-biased) RV surveys.

For the VHJ and HJ predictions that we later make using TrES, XO, and *Kepler*, we will adopt the normalizations from Gould et al. (2006a) of $k(p) = 1/690$ for VHJs with orbital periods uniformly distributed between 1–3 days, and $k(p) = 1/310$ for HJs uniformly distributed between 3–5 day periods. By using statistics from OGLE-III, instead of the RV statistics, we are avoiding the difficulties discussed above, and furthermore any unaccounted for biases inherent in photometric surveys will be preserved, and therefore any systematic effects unknown to us will still be accurately captured in the final predictions. In other words, it is an "apples-to-apples" comparison, instead of "apples-to-oranges" using RV data.

---

[6] See http://www.inscience.ch/transits/.

2.3. *The Present-Day Mass Function*

The next step is to integrate equation (1) over mass:

$$\frac{d^3 N_{\text{det}}}{dr\,dl\,db} = \rho_*(r, l, b) r^2 \cos b$$
$$\times \int_{M_{\text{min}}}^{M_{\text{max}}} P_{\text{det}}(M, r) \frac{dn}{dM}\,dM. \quad (21)$$

Having previously derived an expression for P_{det}, we need only determine the appropriate formulation for the PDMF, dn/dM. Using the results of Reid et al. (2002) who used the *Hipparcos* data sets along with their own data from the Palomar/Michigan State University survey to create a volume-limited survey out to 25 pc, we adopt

$$\frac{dn}{dM} =$$
$$\begin{cases} k_{\text{norm}} \left(\dfrac{M_*}{M_\odot} \right)^{-1.35} & \text{for } 0.1 \leq M_*/M_\odot \leq 1, \\[2mm] k_{\text{norm}} \left(\dfrac{M_*}{M_\odot} \right)^{-5.2} & \text{for } 1 < M_*/M_\odot, \end{cases} \quad (22)$$

where k_{norm} is the appropriate normalization constant. A fuller discussion for our decision to use this particular form for the PDMF is in Appendix C.

Adopting a stellar mass density in the solar neighborhood of $0.032\ M_\odot\ \text{pc}^{-3}$ (Reid et al. 2002), we find a normalization of $k_{\text{norm}} = 0.02124\ \text{pc}^{-3}$.

2.4. *Galactic Structure*

While our normalization of the mass function properly describes the density of stars in the solar neighborhood, beyond about 100 pc we must account for variations in stellar density due to Galactic structure. In addition, as mentioned previously, we must also take into consideration the effects of extinction due to interstellar dust. This is the final stage of integration that needs to be done on equation (1):

$$N_{\text{det}} = \int_0^{r_{\text{max}}} \int_{l_{\text{min}}}^{l_{\text{max}}} \int_{b_{\text{min}}}^{b_{\text{max}}} \Big[P_{\text{det}}(r) \rho_*(r, l, b)$$
$$\times r^2 \cos b\,dr\,dl\,db \Big]. \quad (23)$$

2.4.1. *Galactic Density Model*

We adopt the Bahcall & Soneira model for the Galactic thin disk (Bahcall & Soneira 1980). The model treats stellar density in the thin disk as a double exponential function involving both the distance from the Galactic center and height above the Galactic plane. Specifically, the mass density relative to the local density is given by

$$\rho_*(r, l, b) = exp\left[-\frac{d - d_{\text{GC}}}{h_{d,*}} - \frac{|z|}{h_{z,*}} \right], \quad (24)$$

where $d_{\text{GC}} = 8$ kpc is the distance from the Galactic center, $h_{d,*}$ is the scale length, and $h_{z,*}$ is the scale height.

We adopt a value of $h_{d,*} = 2.5$ kpc for the scale length of the disk. For the vertical scale height, we use the following relation,

which is a slightly modified form of that originally proposed by Bahcall & Soneira (1980):

$$
h_{z,*} = \begin{cases} 90 & \text{for } M_V \leq 2, \\ 90 + 200\left(\dfrac{M_V - 2}{3}\right) & \text{for } 2 < M_V < 5, \\ 290 & \text{for } 5 \leq M_V. \end{cases}
$$

We calculate the absolute V-band magnitudes of stars of a given mass using the blackbody assumptions, and the mass-radius and mass-luminosity relations described in Appendix A.

We defer the complete discussion of our reasoning behind our choices for $h_{d,*}$ and $h_{z,*}$ to Appendix D. The exact values for these parameters are not well constrained in the literature, and we show the effects of using different values in Appendix D. We note that within the range of values found in the literature, the final prediction results may differ by as much as 20%.

We only treat the thin disk in our modeling because we decided that an expanded analysis that also looked at stars in the thick disk, halo, and bulge would not add a significant number of detections to our final result. Both the thick disk and the halo have much lower volume densities than the thin disk at distances typically surveyed by transit searches. In the same vein, the bulge is too far away. These populations are also typically older than the thin disk, and so have a fainter turnoff point, meaning that these stars are more difficult to detect. That being said, for deeper transit surveys, such as some of the all-sky surveys, these populations may play an interesting role. In that case, it would be fairly simple to add them into our model.

We chose the Bahcall model instead of other more detailed models due to speed and portability considerations. While others, such as the Besancon models (Robin & Creze 1986), take into account features such as density variations from the spiral structure of the Galaxy and the Galactic bulge, they typically require over 30 minutes to simulate a given star field. In addition, many are only accessible through Web page interfaces, and therefore difficult to include in a self-contained prediction algorithm.

We do note that comparisons between our Galactic model and the Besancon model revealed star count differences of only 10% in the *Kepler* and TrES fields over a range of magnitude limits.

2.4.2. Interstellar Extinction Modeling

We also model the density of interstellar dust as a double exponential

$$
\rho_{\text{dust}}(r, l, b) = \exp\left[-\frac{d - d_{\text{GC}}}{h_{d,\text{dust}}} - \frac{|z|}{h_{z,\text{dust}}} \right]. \tag{25}
$$

Which gives the dust density at a distance r along a particular line of sight (l, b) relative to the density in the solar neighborhood (i.e., $z = 0$, $d = d_{\text{GC}}$). To calculate the extinction in a given bandpass for a star at a given distance, we then integrate equation (25) over r to arrive at the effective path length through an equivalent solar neighborhood dust density,

$$
\ell_{\text{dust}} = \int_0^r \rho_{\text{dust}}(r, l, b)\, dr. \tag{26}
$$

The V-band extinction in magnitudes, A_V, is given by this path length multiplied by the appropriate normalization constant: $A_V = k_{\text{dust},v}\, \ell_{\text{dust}}$. We adopt $k_{\text{dust},v} = 1$ mag kpc^{-1} (Struve 1847). We determine the extinction in other bands by using the extinction ratios (A_λ/A_V) from Cox (2000) assuming that $R_V = 3.1$.

3. ASSUMPTIONS

It is important to keep in mind that in the preceding derivations, we made several assumptions about the nature of extrasolar planetary systems, as well as some simplifications to keep the math and computing time more manageable.

1. *Detectability is best measured through χ^2.*—While S/N is a better metric than photometric precision to determine if a transit will be detectable, and the χ^2 parameter is a good proxy for S/N, actual transit surveys use a wide variety of tests to detect transits. Many tests, such as the visual evaluation of the data by a human being, are difficult to simulate numerically. Thus, using just a χ^2 test is a simplification, as human inspection may impose a much higher effective S/N threshold. As we show for typical surveys, the number of detections is a strong function of the S/N threshold. Thus, a proper determination of the effective S/N threshold imposed by all cuts is necessary before any robust conclusions can be drawn about the rate of detections relative to expectations. This is a fundamental—and unavoidable—limitation of our study.

2. *No stellar limb darkening.*—We have neglected to include terms for stellar limb darkening because, as stated above, to first order its effects are canceled by the effects of ingress and egress on the transit light curve (Gould et al. 2006a).

3. *Stars are blackbodies.*—We assume that the radiated energy of a star can be described as a blackbody. While in general this is a fair approximation (to within 0.5 mag, see Figs. 11 and 12, below), it does not hold true for the bluer bandpasses we consider (U band in particular) and for very low-mass stars.

4. *Simplified galactic structure.*—While we do treat variations in stellar density and interstellar dust densities in a general sense, our formalism is a simplified double exponential model that does not account for the spiral structure of the Galaxy or patchiness in the dust distribution. Therefore detailed, accurate predictions for an individual field may be better served by incorporating more sophisticated Galactic models (e.g., Fressin et al. 2007).

5. *No binary systems.*—While we only consider main-sequence stars, we do not model the statistics of binary systems; in our formulation all stars are treated as single. Although this is clearly not the case, it was outside the scope of this project to go deeply into binary frequencies, the ability of planets to form in such systems, and whether or not such planets would be detectable.

However, we can make a rough estimate of the expected magnitude of the effect of binaries. First, we assume that 1/3 of our stars have no binary companion (Duquennoy & Mayor 1991). We assume the remaining 2/3 of the stars are in binary systems. A typical wide-field survey has a PSF size of $\sim 20''$, which corresponds to a separation of ~ 4000 AU, and a period of $\log(P/\text{day}) \sim 7.8$ for an equal-mass solar-mass binary. Adopting the period distribution for binary stars from Duquennoy & Mayor (1991) we estimate that $\sim 90\%$ of the binary systems are unresolved. Of these unresolved binaries, we conservatively assume that those with separations $\lesssim 5$ AU, or $\log(P/\text{day}) \sim 3.5$, do not host hot Jupiters because the binary components are too close.

Thus, we estimate that $\sim 62\%$ of unresolved binaries can host hot Jupiters. To estimate the detectability of transiting planets in these, consider a binary where F_1, R_1, M_1 and F_2, R_2, M_2 are the flux, radius, and mass of the primary and secondary, respectively. In addition, define $\ell = F_2/F_1$, $\mathcal{R} = R_2/R_1$, and $q = M_2/M_1 \leq 1$. Then, for fixed planet radius and semimajor axis, assuming photon-limited uncorrelated uncertainties, the signal-to-noise ratios if the stars are resolved are $(\text{S/N})_1 \propto R_1^{-3/2} F_1^{1/2}$ and $(\text{S/N})_2 \propto R_2^{-3/2} F_2^{1/2}$, while if the stars are unresolved the signal-to-noise ratios for transits

across each component are $(S/N)_{B,1} \propto R_1^{-3/2} F_1^{1/2} (1+\ell)^{-1/2}$ and $(S/N)_{B,2} \propto R_2^{-3/2} F_2^{1/2} (1+\ell)^{-1/2}$. Assuming a constant volume density of stars and no interstellar extinction, the number of detected planets scales as $(S/N)_{\min}^{-3}$ (Gould et al. 2003). Thus, the number of planet detections for a population of binaries in the unblended case can be written as $N_0 \propto (S/N)_1^3 + f(q)(S/N)_2^3$, where $f(q)$ is the mass-ratio distribution. Similarly, for the blended case, this is $N_B \propto (S/N)_{B,1}^3 + f(q)(S/N)_{B,2}^3$. Using the scalings for the signal-to-noise ratios just derived, we evaluate the ratio of the number of detections in the blended case, relative to the number of detections one would expect if they were unblended,

$$ \frac{N_B}{N_0} = (1+\ell)^{-3/2} \frac{1 + f(q)\ell^3 \mathcal{R}^{-9/2}}{1 + f(q)\ell^{3/2} \mathcal{R}^{-9/2}}. \qquad (27) $$

We adopt $\ell = q^5$, which is roughly appropriate for the R band (see Fig. 11, below), and $\mathcal{R} = q^{0.8}$ (eq. [A4]). We consider two different distributions for $f(q)$. First, we adopt $f(q)$ from Duquennoy & Mayor (1991) which is weighted toward low mass-ratio systems. Second, we consider a distribution with a substantial population of "twins" with $q = 1$. Specifically, we choose $f(q) \propto 1 + 3.5 \exp[-0.5(q-1)^2/\sigma_q^2]$, with $\sigma_q = 0.2$, which approximates the distribution found by Halbwachs et al. (2003). When we then average equation (27) over q, weighting by $f(q)$, we find $\langle N_B/N_0 \rangle \simeq 0.89$ for the Duquennoy & Mayor (1991) mass-ratio distribution, and $\langle N_B/N_0 \rangle \simeq 0.68$ for the Halbwachs et al. (2003) mass-ratio distribution. Thus, we estimate that roughly 70%–90% of planets in unresolved binaries will still be detectable. The basic point is that, because the mass-luminosity relation is so steep, only those binaries that have $q \sim 1$ result in a substantial reduction in the signal-to-noise ratio. Low mass-ratio companions have little affect on the ability to detect transit signals. For the mass-ratio distribution of Duquennoy & Mayor (1991) these provide the bulk of the companion population, and so in this case the effects of binarity are suppressed even further.

Finally, we can summarize these various components to provide a rough estimate of the fraction of planets that will be detected when we consider binary systems, relative to our fiducial assumption that all stars are single. We find $1/3 + (0.1 + 0.62\langle N_B/N_0 \rangle)(2/3) \simeq 0.68$–$0.77$. Thus we expect that our estimated yields are systematically high by at most 30%. This is generally of order or smaller than other sources of uncertainty in our predicted yields.

We note that a corollary of our assertion that binaries have a small effect on the yields of transit surveys is that a substantial fraction of the planets that have been found in wide-field transit surveys should be in binary systems. Only two planets in confirmed common proper-motion binary systems have been reported (Burke et al. 2007; Bakos et al. 2007); however, systematic surveys for stellar companions to the host stars of the known transiting planets have not been performed, and therefore the true incidence of binary systems among the sample of transiting planets is not known.

4. RESULTS

Using the preceding formalism, we simulated four different surveys modes used by transit searches: point-and-stare, drift-scanning, space-based, and all-sky surveys. For the point-and-stare and drift-scanning cases, we use the TrES and XO transit surveys as our examples, respectively. Both surveys have been successful in their hunt for planets, and so we are able to directly compare the results of our calculations with the actual number of detections. For VHJs and HJs, we simulate the detection of

TABLE 1
TrES Survey Parameters

Parameter	Value
Band	550–700 nm
m_R limit	≈ 13
Observation time	Varies
Exposure time	90 s
CCD read time	25 s
Telescope diameter............................	0.1 m
Throughput...	0.6
PSF (FWHM)...................................	$25''$
$\theta_{\rm pix}$	$10''$
CCD full-well depth	$1.2 \times 10^5\ e^-$
Field of view	36 deg^2
Galactic longitude	Varies
Galactic latitude	Varies

planets with $R_p = 1.1 R_J$, and with the period distribution from Gould et al. (2006a) that we describe in § 2.2.

We demonstrate our method for space-based surveys by simulating the upcoming *Kepler* mission and comparing our predictions with other predictions for *Kepler* in the literature. Similarly for the all-sky mode, we examine the future planned observing programs of LSST, SDSS-II, and Pan-STARRS and make predictions for each.

For each mode, we have tried to approximate the effect of red noise on detectability by increasing our S/N cut (Pont et al. 2006). To show the effects of a more accurate red noise calculation, in § 4.5 we evaluate its effects on the TrES survey.

4.1. *Point-and-Stare TrES*

Point-and-stare observing is the traditional observing mode of tracking one field of sky continuously for all, or part, of the night. There are several operational photometric transit surveys that use this particular method of observing, including HATnet (Bakos et al. 2004), SuperWasp (Pollacco et al. 2006), and KELT (Pepper et al. 2007), but we chose to model is the Trans-Atlantic Exoplanet Survey (TrES). TrES is composed of a network of three telescopes: STARE on the Canary Islands, PSST at Lowell Observatory in Arizona, and Sleuth on Mt. Palomar, California. All are small, 10 cm aperture, wide-field ($6°$), CCD cameras that operate in unison to observe the same field of sky nearly continuously over one to two month periods. The survey specifics are described in more detail by Dunham et al. (2004) and by Brown & Charbonneau (1999). The relevant quantities for our modeling are listed in Table 1.

We fit to the RMS plots posted on the Sleuth observing Web site[7] to determine the background sky brightness in the TrES bandpass (550–700 nm), using the TrES team's stated photon throughput of 0.6, and we estimate that the TrES bandpass has a sky background of 19.6 mag arcsec^{-2}.

In our simulations, we treat the three TrES telescopes as one single instrument. This is clearly an oversimplification. While these telescopes are quite similar, there are differences between the data quality and quantity obtained from each site. For example, Sleuth must contend with a higher sky background due to its proximity to urban areas, whereas the fraction of clear nights for PSST is typically lower than for the other two sites. However, accounting for these differences is difficult, and furthermore their effect on our predicted detections is small compared to the other

---

[7] See http://www.astro.caltech.edu/~ftod/tres/sleuthObs.html.

TABLE 2
Properties of the Simulated TrES Survey Fields

Field	Observation Time (hr)	Galactic Coordinates
And0	765	[126.11, −015.52]
Cyg1	609	[084.49, +010.28]
Cas0	628	[120.88, −013.47]
Per1	643	[156.37, −014.04]
UMa0	497	[168.87, +047.70]
CrB0	537	[053.49, +048.92]
Lyr1	582	[077.15, +017.86]
And1	435	[109.03, −017.62]
And2	332	[115.52, −016.21]
Tau0	406	[169.83, −015.94]
UMa1	500	[156.32, +054.01]
Her1	592	[063.26, +026.42]
Lac0	686	[097.79, −008.86]

uncertainties in our model. We therefore ignore these effects. In dealing with the window probability of the TrES network, we were greatly aided by the plots posted on the Sleuth Web site which showed the window probability for each field, as well as the total network observing time each field had received.

We simulated thirteen different TrES fields using this methodology. The network observation times and the location of each field are shown in Table 2. The data for this table was collected off from the Sleuth observing Web site. The expected detections for each of the fields (and the total number of detections over all simulated fields) are shown in Table 3. The effects of our Galactic structure model can clearly be seen in distribution of expected detections. Both the UMa0 and CrB0 fields are substantially above the plane of the Galaxy ($b \approx 48°$), and thus show a much lower detection rate than those fields directed toward denser regions of the sky.

Figure 2 shows the distributions of detections over various parameters in the Lyr1 field. For Figure 2, we show two distributions, one assumes white noise in the photometry, the other assumes 3 mmag of red noise. For the white noise estimates, we used a signal-to-noise limit of S/N ≥ 30 and a magnitude limit of $m_R \leq 13$. Lyr1 is the field in which TrES-2 was discovered, and is adjacent to Lyr0, which contains TrES-1. We therefore have plotted the location of these two stars on the distributions shown in Figure 2. Both stars reside close to the peaks of our parent star mass and radius distributions. The third and fourth planets discovered by TrES, TrES-3 and TrES-4, are not plotted in Figure 2, as they reside in the Hercules fields, which are at a higher Galactic latitude than Lyra. Also of note is that with our assumptions half of the expected detections the F dwarf range. Among the false positives encountered by the TrES team, the most frequent are eclipsing F/M dwarf binaries. Figure 2 offers a partial explanation for this, showing that roughly half of TrES's planet-like signals come from F dwarfs.

Our predicted number of detections for the selected TrES fields of $N_{det} = 6.50$ covers only half of the fields observed by

TrES, so we may roughly extrapolate this result to 13.00 detections over all of the fields. This is roughly 3 times the four planets that TrES has actually detected, although we hesitate to draw strong conclusions from this comparison. While we have the luxury in our predictions of imposing draconian signal-to-noise ratio and magnitude cuts in the data, in real life the TrES detections are identified in data using more flexible selection effects. A better comparison between our predictions and the actual results of TrES (or any survey) would therefore require a better understanding of the selection effects in the observed data set.

To see how different selection criteria change our predictions for TrES, Tables 4, 5, and 6 show the effects of changing the S/N limit, the magnitude limit, and the radius of the targeted planets on the detection numbers.

A changing magnitude limit is particularly interesting, since all four of the TrES detections have been around stars with $m_R \leq 12$. This might indicate that the effective magnitude limit of the TrES survey is brighter than what is described in the published literature. If this is the case, then using a limit of $m_R \leq 12$ lowers the number of expected detections in our simulated fields to $N_{det} = 4.08$, and total detections to 8.16.

4.2. Drift-Scanning XO

The XO survey is a wide-angle ground-based photometric survey that uses drift scanning for its observations. The XO telescope is a pair of 11 cm cameras in place at the Haleakala Observatory in Hawaii. It monitors thousands of stars with $m_V \leq 12$ in six fields spaced evenly around the sky in right ascension.[8] The cameras scan within these fields, sweeping in declination from $0°$ to $+63°$ every 10 minutes. Typically, this provides over 3000 observations per star in each of the XO fields. More on the specifics of the XO system can be found in McCullough et al. (2005). The relevant parameters for our predictions are listed in Table 7.

To determine the throughput of XO and the background sky magnitude in the XO band (400–700 nm), we fit to the RMS diagram shown in Figure 8 of McCullough et al. (2005). With the addition of their stated calibration and scintillation error, we were able to exactly reproduce the relations shown in that figure using a background sky magnitude of 19.4 mag arcsec^{-1} and a system throughput of 0.16.

In addition, we checked to ensure that the values we used for the observation time and the number of nights observed generated an analytic window probability similar to that shown in Figure 7 of McCullough et al. (2005). Our ability to do this was greatly aided by the XO team's description of how many nights they lost to weather and mechanical trouble.

We simulated each of the six XO fields for our predictions. The results for each field using the base parameters of S/N ≥ 30, $m_V \leq 12$, and $R_p = 1.1R_J$ are shown in Table 8. The first planet discovered by the XO survey, XO-1b, (McCullough et al. 2006) is located in the field at 16 hr, and so we have also plotted the distribution of detections within this field over a variety of parameters

---

[8] These are 00, 04, 08, 12, 16, and 20 hr.

TABLE 3
Predicted TrES Detections for $m_R \leq 13$, S/N ≥ 30, and $R_p = 1.1R_J$

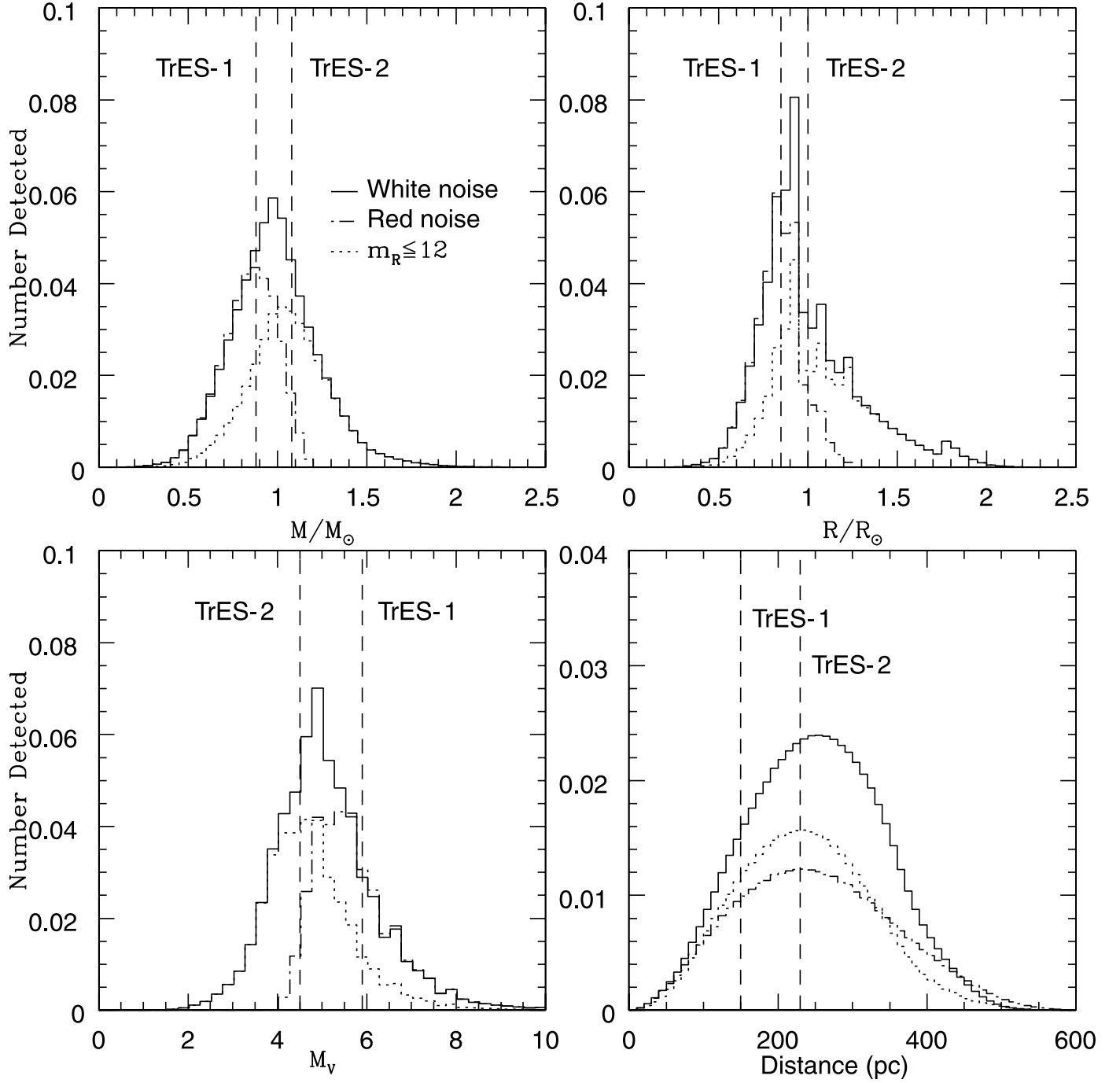| Detection Type	And0	Cyg1	Cas0	Per1	UMa0	CrB0	Lyr1	And1	And2	Tau0	UMa1	Her1	Lac0	Total
VHJs	0.38	0.37	0.34	0.34	0.18	0.21	0.32	0.24	0.19	0.22	0.17	0.21	0.41	3.59
HJs	0.34	0.30	0.29	0.28	0.16	0.18	0.27	0.18	0.09	0.15	0.16	0.19	0.34	2.91
Both	0.72	0.67	0.63	0.62	0.34	0.39	0.59	0.42	0.28	0.37	0.33	0.40	0.75	6.50

FIG. 2.—Distribution of predicted TrES detections in the Lyr1 field with and without red noise, and assuming a brighter magnitude limit than our fiducial value of $m_R \leq 13$. The location of the stars TrES-1 and -2 are also shown. TrES-2 lies within Lyr1, while TrES-1 is in the nearby Lyr0 field.

TABLE 4
PREDICTED TrES DETECTIONS FOR $m_R \leq 13$ AND $R_p = 1.1\,R_J$

S/N	VHJ	HJ	Both
≥20	6.06	5.72	11.78
≥25	4.68	4.05	8.73
≥30	3.59	2.91	6.50
≥35	2.77	2.13	4.90
≥40	2.16	1.59	3.75

TABLE 5
PREDICTED TrES DETECTIONS FOR S/N ≥ 30 AND $R_p = 1.1\,R_J$

m_R	VHJ	HJ	Both
≤12.0	2.13	1.95	4.08
≤12.5	2.95	2.58	5.53
≤13.0	3.59	2.91	6.50
≤13.5	3.86	3.01	6.87
≤14.0	3.94	3.03	6.97
≤14.5	3.96	3.05	7.01

TABLE 6
PREDICTED TrES DETECTIONS FOR S/N ≥ 30 AND $m_R \leq 13$

R_p (R_J)	VHJ	HJ	Both
0.9	1.66	1.17	2.83
1.0	2.52	1.90	4.42
1.1	3.59	2.91	6.50
1.2	4.53	3.88	8.41
1.3	5.52	5.04	10.56

TABLE 8
PREDICTED XO DETECTIONS FOR $m_V \leq 12$,
S/N ≥ 30, AND $R_p = 1.1$ R_J

Field (hr)	Detection
00	0.34
04	0.35
08	0.34
12	0.30
16	0.32
20	0.37
Total	2.02

in Figure 3. XO-2b (Burke et al. 2007) and XO-3b (Johns-Krull et al. 2008) are in noticeably different fields (08 and 04 hr, respectively), so were not included on the plot.

Tables 9, 10, and 11 show the effects of changing the S/N limit, the magnitude limit, and the radii of the targeted planets on XO's predicted detection numbers. For our base case of $m_V \leq 12$ and S/N ≥ 30, we expect that XO will detect 2.02 planets. If we consider the three actual detections from XO, it seems that, similar to TrES, the actual S/N and magnitude limit of the XO survey may differ from what is in the literature. For example, all three XO parent stars are at $m_V \leq 11.5$ (XO-1, the dimmest, is at $m_V = 11.3$). To see how this affected our calculated distributions, we also simulated the field at 16 hr assuming S/N ≥ 20, $m_V \leq 12$, and S/N ≥ 20, $m_V \leq 11.5$. The results are included on Figure 3. In these cases, the total number of detections increases to 5.86 and 4.21, respectively.

4.3. Space-based Kepler

To make predictions for a space-based mission, we chose NASA's Kepler mission, which is primarily designed to look for Earth-sized extrasolar planets. It consists of a single spacecraft that is currently scheduled to be launched into an Earth-trailing heliocentric orbit in early 2009. On board, Kepler will have 42 CCDs that will measure the light from 100,000 main-sequence stars within the telescope's field of view. Ideally, Kepler will identify planets orbiting within the "habitable zone" of their parent stars which would be capable of possessing liquid water on their surfaces.

Beyond looking for extrasolar Earths, the immense amount of photometry that Kepler is expected to produce over its four year operational lifetime will also allow for the detection of larger Jovian worlds transiting other stars. Previous estimates of the number of VHJs and HJs expected to be found by Kepler have predicted the discovery of about 180 HJs (no VHJs had been

TABLE 7
XO SURVEY PARAMETERS

Parameter	Value
Band	400–700 nm
m_V limit	12
Observation time	≈400 hr
Exposure time	54 s
CCD read time	n/a
Telescope diameter	0.11 m (two)
Throughput	0.16
PSF (FWHM)	45″
θ_{pix}	25.4″
CCD full-well depth	1.5×10^5 e^-
Field area	7.2×63 deg^2
Galactic longitude	Varies
Galactic latitude	Varies

detected at the time, and so they were not treated; Jenkins & Doyle 2003). Clearly, given that there are only 46 currently known transiting planets, this would be a significant increase in the number of known planets

To simulate Kepler with our methodology, we used the values in Table 12 to simulate Kepler observations out to a distance of 8 kpc, using a signal-to-noise ratio of S/N ≥ 7. We gathered the information in Table 12 from both the Kepler mission Web site[9] as well as published descriptions of the satellite (Borucki et al. 2004). Three items are of note. First, our value for the photon detection efficiency of Kepler is based on information from the Kepler Web site that a 12th magnitude G2 dwarf will show 7.8×10^8 γ hr^{-1}. Secondly, the value we use for the extinction ratio in the Kepler bandpass, $A_K/A_V = 0.861$, is an average that we calculated by integrating the extinction curve over the Kepler bandpass (400–850 nm; Fitzpatrick 1999). Finally, the Kepler Web site and much of the literature states that Kepler will use a dim magnitude cutoff of $m_V \leq 14$. In reality, there is no "hard" magnitude cutoff at which stars will be excluded from observations. Indeed, Kepler is expected to observe promising stars that are much dimmer than $m_V = 14$ (D. W. Latham 2007, private communication). Therefore, for our more explicit calculations of the expected number of HJ and VHJ detections for Kepler, we adjusted the magnitude limit such that the number of stars in the field agreed with the expected number of Kepler targets (100,000 in the case of the new field). For other, more general, simulations, we used several different magnitude limits to determine what effects these had on Kepler's expected detections.

While the location of the Kepler field is currently centered on Galactic coordinates ($l = 76.32°$, $b = 13.5°$), up until 2004 it was instead situated closer to the Galactic plane at ($l = 70°$, $b = 5°$). The switch in the position of the field to higher Galactic latitudes was done in order to avoid sorting through the larger number of giant stars that are found in fields directed along the plane of the Galaxy. Much of the previously published literature examining the Kepler mission have used the coordinates of the old field. For the purpose of comparison, we therefore simulated both the old and new Kepler fields.

4.3.1. General Characterization

Looking in the new Kepler field with a magnitude limit of $m_V \leq 14$, we simulated the number of detections as a function of orbital period for Jupiter-, Earth-, and Mars-sized worlds, assuming that every star possessed a planet with the given period and radius (Fig. 4), and that red noise was negligible. Interestingly, the difference between the number Jupiter- and Earth-sized planets detected is minimal for $p \leq 10$ days, which indicates that Kepler is able to detect all the planets within this radius range,
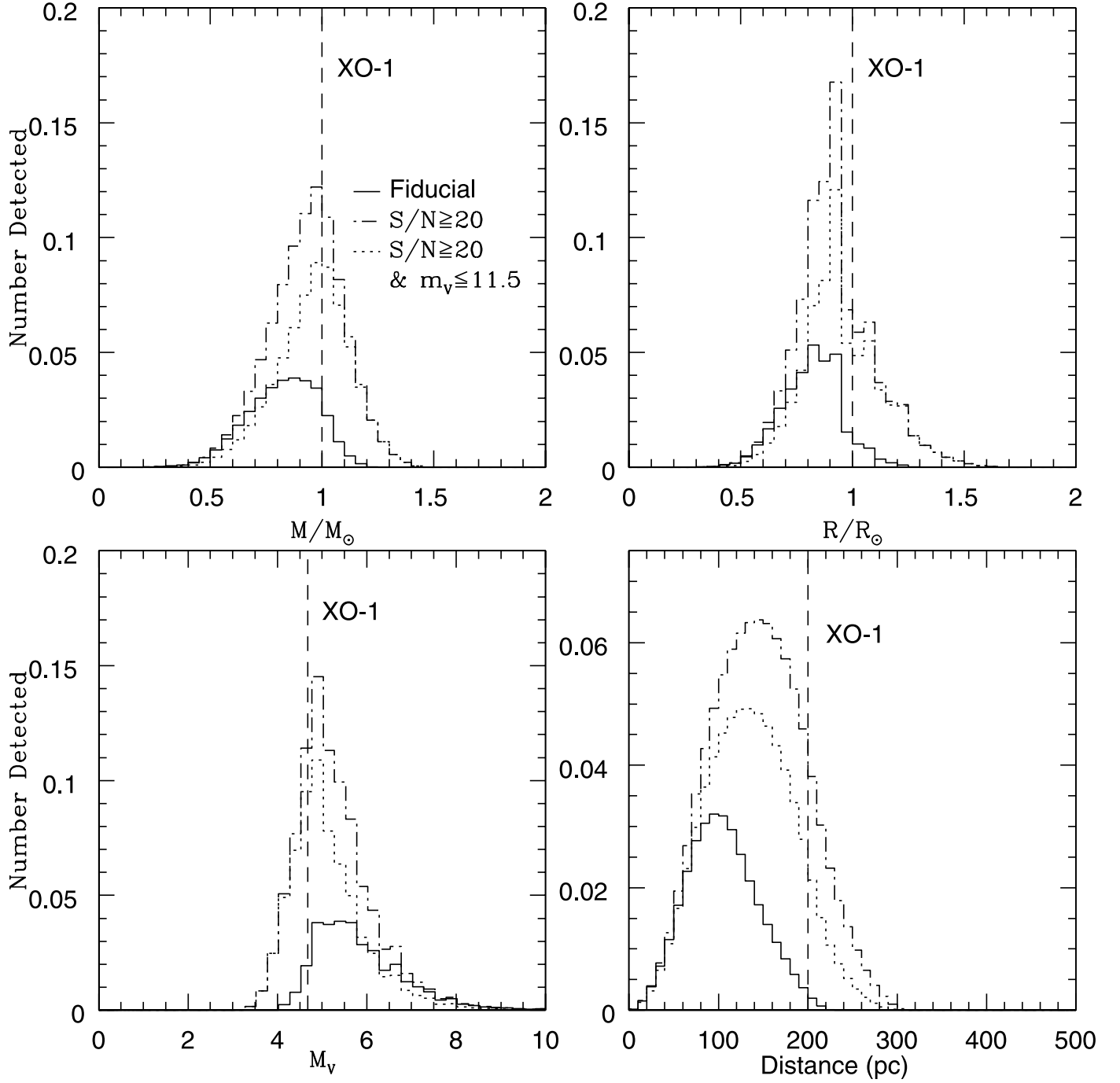
[9] See http://kepler.nasa.gov/.

Fig. 3.—Distribution of predicted XO detections in the field at 16 hr plotted against various parameters. Three cases are shown: our fiducial S/N ≥ 30, $m_V \leq 12$ case, S/N ≥ 20, $m_V \leq 12$, and S/N ≥ 20, $m_V \leq 11.5$. Also shown is the actual location of the star XO-1, which is located within this field.

TABLE 9
PREDICTED XO DETECTIONS FOR $m_V \leq 12$ AND $R_p = 1.1\ R_J$

S/N	VHJ	HJ	Both
≥20	3.45	2.41	5.86
≥25	2.20	1.27	3.47
≥30	1.37	0.65	2.02
≥35	0.83	0.33	1.16
≥40	0.50	0.17	0.67

TABLE 10
PREDICTED XO DETECTIONS FOR S/N ≥ 30 AND $R_p = 1.1\ R_J$

m_V	VHJ	HJ	Both
≤11.5	1.09	0.55	1.64
≤12.0	1.37	0.65	2.02
≤12.5	1.50	0.69	2.19
≤13.0	1.54	0.70	2.24
≤13.5	1.55	0.71	2.46

TABLE 11
PREDICTED XO DETECTIONS FOR S/N $\geq$ 30 AND $m_V \leq 12$

R_p (R_J)	VHJ	HJ	Both
0.9...	0.29	0.09	0.38
1.0...	0.69	0.25	0.94
1.1...	1.37	0.65	2.02
1.2...	2.07	1.18	3.25
1.3...	2.93	1.92	4.85

assuming that they transit. Therefore, for $m_V \leq 14$ and $p \leq 10$ days, the S/N of transiting Jupiter- and Earth-sized planets is large enough that *Kepler*'s S/N limit is irrelevant, and the $m_V \leq 14$ magnitude cutoff is the limiting factor in the number of detections. For smaller planets the size of Mars, the S/N of their transits are in the neighborhood of the *Kepler* S/N limit for the brightest stars and shorter periods, and so Mars-sized worlds are detected at a much lower rate.

To more accurately characterize the magnitude- and S/N-limited detection regimes, we plotted the number of detections as a function of planetary radius for orbital periods of 1 day and 365 days (Fig. 5). In addition in a magnitude limit of $m_V = 14$, we also plot the same relations for a limit of $m_V = 20$. One can clearly see the flattening of these curves as the radius of the target planets increase, indicating the increasing effect of the magnitude limit on the detection numbers. As is expected in the S/N limited regime, detections go as R_p^6 (Pepper et al. 2003; Gaudi 2005; Pepper & Gaudi 2005), although at the lower end saturation of the *Kepler* CCDs serves to cut out some of these detections.

4.3.2. *Habitable Planets*

One of the major goals of the *Kepler* mission is the detection of planets in a star's habitable zone that would be capable of supporting life. The habitable zone is defined as the distance from the parent star at which liquid water could exist on a planet's surface. We estimate the detection rate of habitable zone planets by assuming that all stars have a planet with a semimajor axis equal to

$$\frac{a}{\text{AU}} = \sqrt{\frac{L_{\text{bol},*}}{L_{\text{bol},\odot}}}. \qquad (28)$$

TABLE 12
Kepler MISSION SURVEY PARAMETERS

Parameter	Value
Band	400–850 nm
A_K	0.86
m_V limit..	$\approx$14
Observation time................................	4 yr
Exposure time	3 s
CCD read time.....................................	Negligible
Telescope diameter..............................	0.95 m
Throughput...	0.3
PSF (FWHM)....................................	10″
θ_{pix}	3.98″
CCD full-well depth	$1.2 \times 10^6 \ e^-$
Field of view	106 deg^2
Galactic longitude (new)	76.32°
Galactic longitude (old).....................	70°
Galactic latitude (new).......................	13.5°
Galactic latitude (old)........................	5°

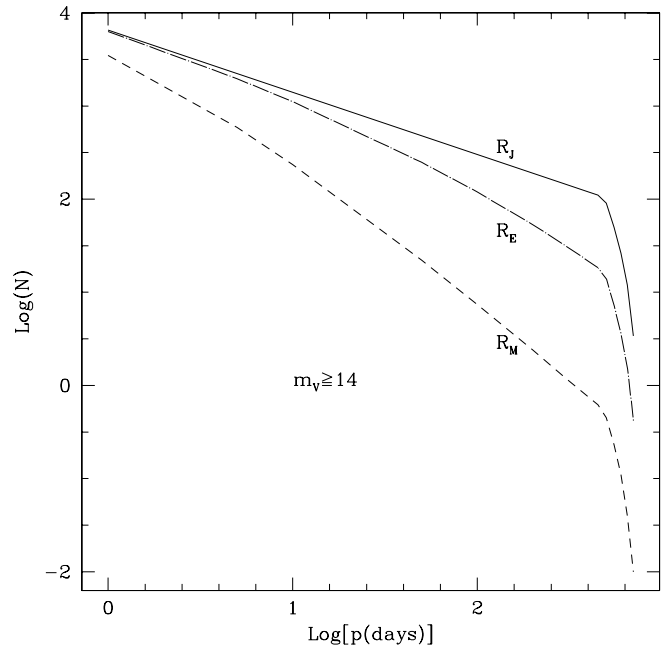FIG. 4.—$\log(N)$ vs. $\log(p)$ for *Kepler* detections down to $m_V = 14$ assuming that every star possess a planet, for Jupiter-, Earth-, and Mars-sized worlds. Note that both the Jupiter- and Earth-sized detection numbers are similar for $p \leq 10$ days; this implies that in this radius and period regime the number of *Kepler* detections is limited by the magnitude cutoff, and not the S/N (as is the case for the Mars-size planets). The strong downturn around $\log(p) = 2.7$ is a result of *Kepler*'s window probability (i.e., the probability of seeing two or more transits).

This assumption is simple and likely optimistic. In reality, the habitable zone has a finite width, and habitable planets will have a range of radii. Thus, a better estimate would be to adopt a distribution of planet masses and radii, and integrate over the range of radii and periods considered habitable. Furthermore, the assumption of one habitable planet per star is probably optimistic, and so the actual number of detected habitable Earths will be smaller than we estimate, depending on the real frequency of Earth-like planets. For the sake of simplicity, however, we will adopt the assumptions above.

With this assumption, we simulated the detection of habitable Earths in both the old and new *Kepler* fields (Fig. 6). Surprisingly, despite there being nearly twice as many main-sequence stars in the old field, the number of detected habitable Earths remained nearly the same in the two fields. This occurred because at the magnitude relevant for *Kepler*, the increased star count in the old field was mainly due to the larger number of A and F main-sequence stars found closer to the Galactic plane; habitable planets are doubly hard to detect around these stars due to the lessened depth of transit, and the longer periods required for a planet to be in the habitable zone. The exact distribution of stars by spectral type for the old and new fields is given in Table 13, and is discussed further below. For the 106,000 stars in the new *Kepler* field, we find that *Kepler* will be able to detect 79 habitable Earth radius planets, assuming that every star posses one.

In general, Figure 6 provides an excellent example of the benefit of repositioning the *Kepler* field away from the Galactic plane. Namely, the number of large dwarfs and giant stars (which we have not treated) found in the Galactic plane can be avoided, reducing the false-positive rate, while the number of habitable Earths detected remains nearly the same (Gould et al. 2003).

It is also possible that life may arise on non–Earth-sized worlds. A world the size of Mars might develop quite differently than it has in our own Solar System if it were closer to its parent star, and
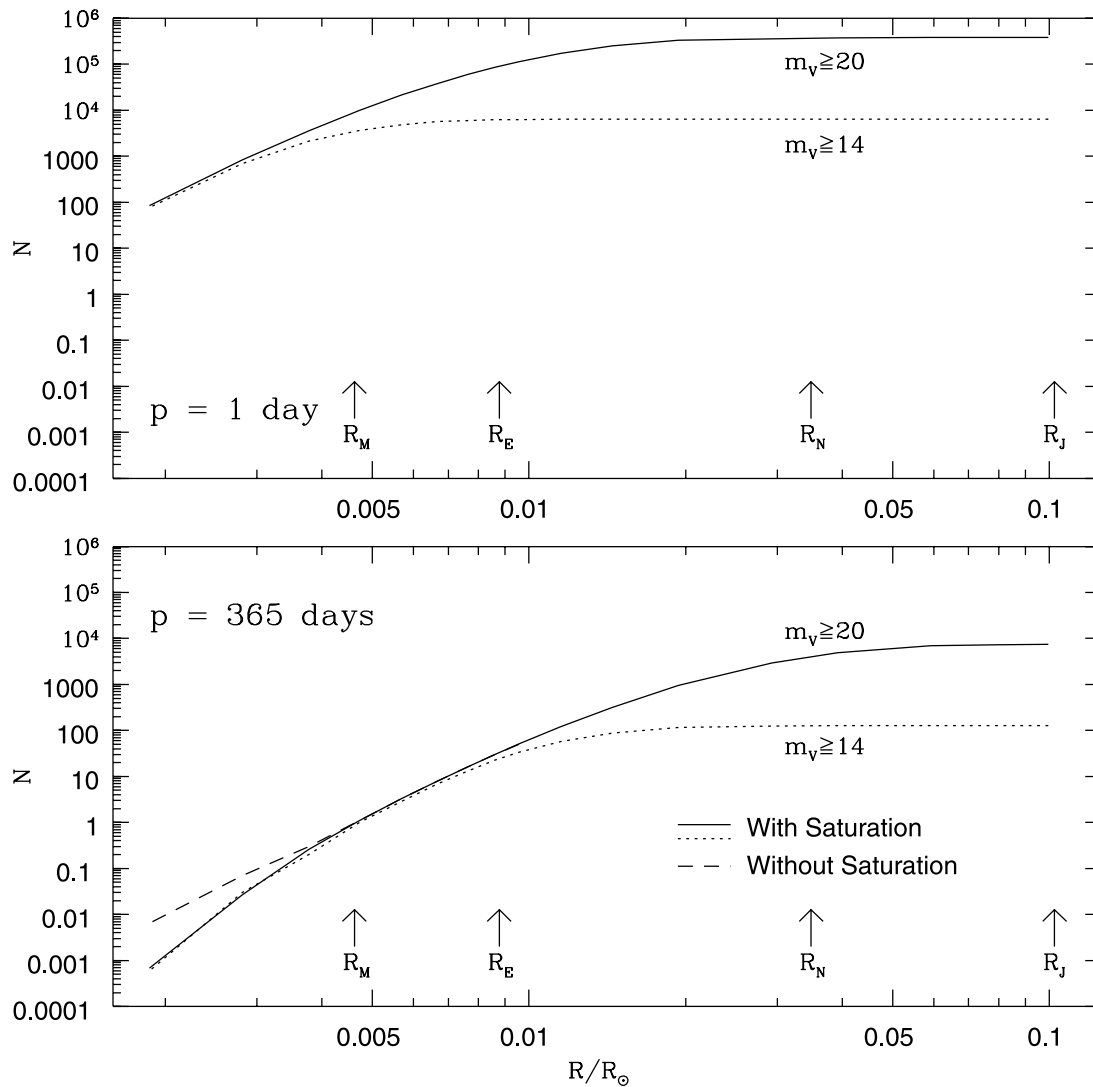
Fig. 5.—*Kepler* detections as a function of the planetary radius for magnitude limits of $m_V \geq 14$ and $m_V \geq 20$ assuming that all stars have planets. Also shown are markers for the radii of Jupiter, Neptune, Earth, and Mars. One can see that as the period increases, the planetary radius at which *Kepler* becomes magnitude-limited (*flat portions of the graphs*) increases, while the overall number of detections falls.

Jovian worlds in the habitable zone could possess large moons capable of supporting life. Figure 7 shows the number of detections of Jupiter-, Earth-, and Mars-sized habitable worlds, assuming that every star has a planet, against a variety of parameters in the new *Kepler* field. For completeness, the magnitude limit was set to $m_V \geq 20$. One can see in the mass and radius plots an explicit demonstration of why changing the number of bright, massive main-sequence stars in the field does not effect the number of habitable planet detections. As was mentioned earlier all detections drop to zero beyond 1.4 $M_\odot$. In the plot showing the number of detections binned against the absolute V magnitude of the parent star we have additionally shown the effect of changing the apparent V magnitude limit on the distribution of Earth-sized planets.

4.3.3. *Hot and Very Hot Jupiters*

Using the HJ and VHJ frequencies provided by Gould et al. (2006a) we were able to make explicit predictions about the number of these worlds that *Kepler* should find. As mentioned previously, Jenkins & Doyle (2003, hereafter JD03) have made their own estimates of *Kepler*'s HJ detection ability (as noted

previously, no VHJs were known at the time, and so were not considered), thus providing a useful point of comparison. As before, we used a limit of S/N ≥ 7, although for Jupiter-sized planets *Kepler* is magnitude, and not signal-to-noise, limited.

At the time of the JD03 estimate the *Kepler* field was centered on its old coordinates. For the purposes of comparison, we therefore examined the old *Kepler* field in addition to the new field.

Tables 13 and 14 show the distribution of our predictions by spectral type. To facilitate comparison with JD03, we have included not only the final prediction numbers for both old and new *Kepler* fields, but also the distribution of the calculated total number of stars in each field. As stated above, in both fields we set the magnitude limit of the simulation such that the total number of stars in the fields agreed with the number of expected *Kepler* targets. In the case of the new field, this is 100,000 stars, while for the old field we matched the number of stars predicted by JD03. This corresponded to a magnitude limit of $m_V = 15.7$ in the new field and $m_V = 15.9$ in the old field.

Looking at the detections, one can see in Table 14 that our final prediction of $N_{\text{det}} = 127$ for the old *Kepler* field is lower
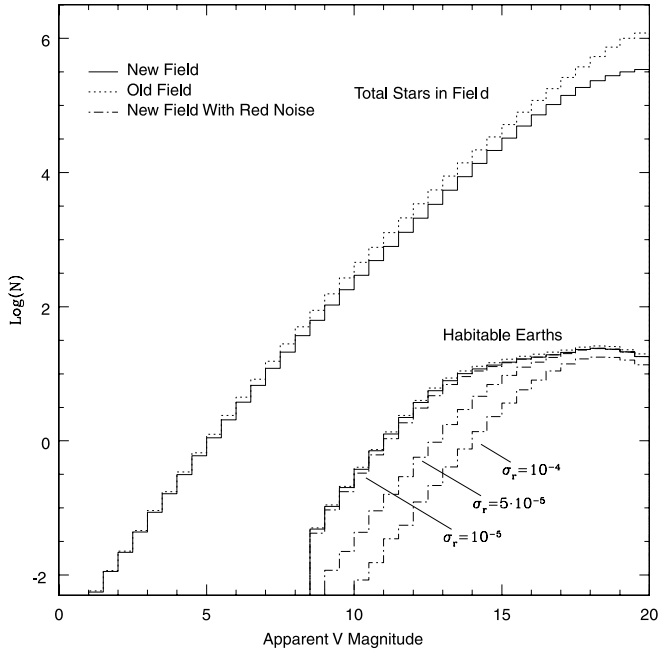
Fig. 6.—Number of habitable Earths detected by *Kepler* in the old and new fields binned by the apparent *V* magnitude of the parent star, as well as the results assuming varying amounts of red noise on the new field. Also shown are the star counts for both fields binned in a similar manner.

than the JD03 prediction. This is primarily a result of the fact that the Gould et al. (2006a) results that we use for $k(p)$ (see § 2.2) posit a HJ and VHJ frequency approximately 1/3 of the frequency used in JD03. An important difference, however, between our work and JD03 is our consideration of VHJs, which were not treated by JD03. This has the effect of "doubling" the number of planets we expect *Kepler* to detect (as demonstrated by the final values in parentheses in Table 14). We therefore arrive at detection rates for VHJs and HJs that are 2/3 that of JD03 for the old *Kepler* field, and 1/3 that of JD03 in the new *Kepler* field. For the old field, we predict 54.2 VHJs and 72.8 HJs. For the new field, we predict 26.9 VHJs, and 36.0 HJs.

In the new *Kepler* field, we see that looking away from the plane of the Galaxy reduces the number of HJs and VHJs detected by half to $N_{det} = 62.9$, which matches the overall drop in the star count in the new field. The distribution of detections changes substantially, shifting more toward F, G, and K dwarfs. This is the caused by our model of the Galaxy's vertical structure, which uses reduced scale heights for earlier stars. Thus, most of the difference in star counts between the old and new fields is caused by fewer B, A, and F dwarfs, leading to proportionally smaller detection numbers around these types of stars.

While the prediction number of detections in the new *Kepler* field is lower than the corresponding JD03 prediction would be for the same star count, this is accounted for in the same way as in the old field. Indeed, we find that our predictions are consistent with those of JD03, except for the difference in our assumption

about the frequency of HJs and VHJs. The eventual results from *Kepler* therefore offer an excellent opportunity to more robustly determine the actual prevalence of short-period Jovian worlds.

4.4. *All-Sky Surveys*

In addition to wide-field photometric surveys, several current and planned projects aim to repeatedly gather photometry over gigantic swathes of the sky. While not all of them have extrasolar planet detection as one of their goals, the large areas covered by these surveys, together with their long durations (from 3 to 10 yr), offer the possibility of detecting hundreds of transiting VHJs and HJs.

We examined three all-sky surveys, LSST, SDSS-II, and Pan-STARRS. The Large Synoptic Survey Telescope (LSST) will be a large 8.4 m telescope, with an effective aperture of 6.4 m, situated in northern Chile. It is scheduled to finish construction sometime in 2014. Using a 9.6 deg^2 field of view, the LSST team intends to image 20,000 deg^2 of the sky repeatedly over the course of 10 yr.[10]

SDSS-II is the second phase of the Sloan Digital Sky Survey (SDSS) which began in 2005 June and is scheduled to last for three years until 2008 June. It is composed of three distinct surveys, the Sloan Legacy Survey, SEGUE, and the Sloan Supernova Survey (Sako et al. 2005). The Supernova Survey is the only one of these three that will collect photometry capable of finding transiting planets; the other two survey modes are focused on non-variable objects within and outside the Galaxy. To execute the Supernova Survey, SDSS uses a 2.5 m telescope on Apache Point in New Mexico with a 6.25 deg^2 field of view. Over the three year run of SDSS-II, the team expects the Supernova Survey to image about 300 deg^2 of the sky near the Galactic poles every two nights for the three months a year the poles are visible. Given the observing cadence the survey expects to use, this translates into a total observing time of 37.5 days.

Finally, Pan-STARRS (the Panoramic Survey Telescope and Rapid Response System) will be a group of four 1.8 m telescopes atop the summit of Haleakala, Hawaii (Kaiser 2004). The first telescope, PS-1, is nearing completion. When fully operational, Pan-STARRS will embark on a 3 yr long "medium deep" survey that will observe 1200 deg^2 of the sky near the Galactic poles for a total of five months.[11] We also looked at a notional "wide" Pan-STARRS survey that looked at 10 times the amount of sky as the medium deep survey, but used the same amount of observing time.

To characterize each of these surveys and find how many transiting VHJs and HJs they should discover, we first calculated the *V* magnitude limits at which each of the surveys will be able to detect transiting Jovian worlds. The complete derivation of this limit is given in Appendix E. Table 15 shows the results of these

---

[10] See http://www.lsst.org.
[11] The remaining balance of the three years will be taken up by the other Pan-STARRS science modes; see http://pan-starrs.ifa.hawaii.edu/project/reviews/PreCoDR/documents/PSCoDD_1_4_design_reference_mission.pdf.

TABLE 13
NUMBER OF MAIN-SEQUENCE STARS IN *Kepler*'s FIELD OF VIEW

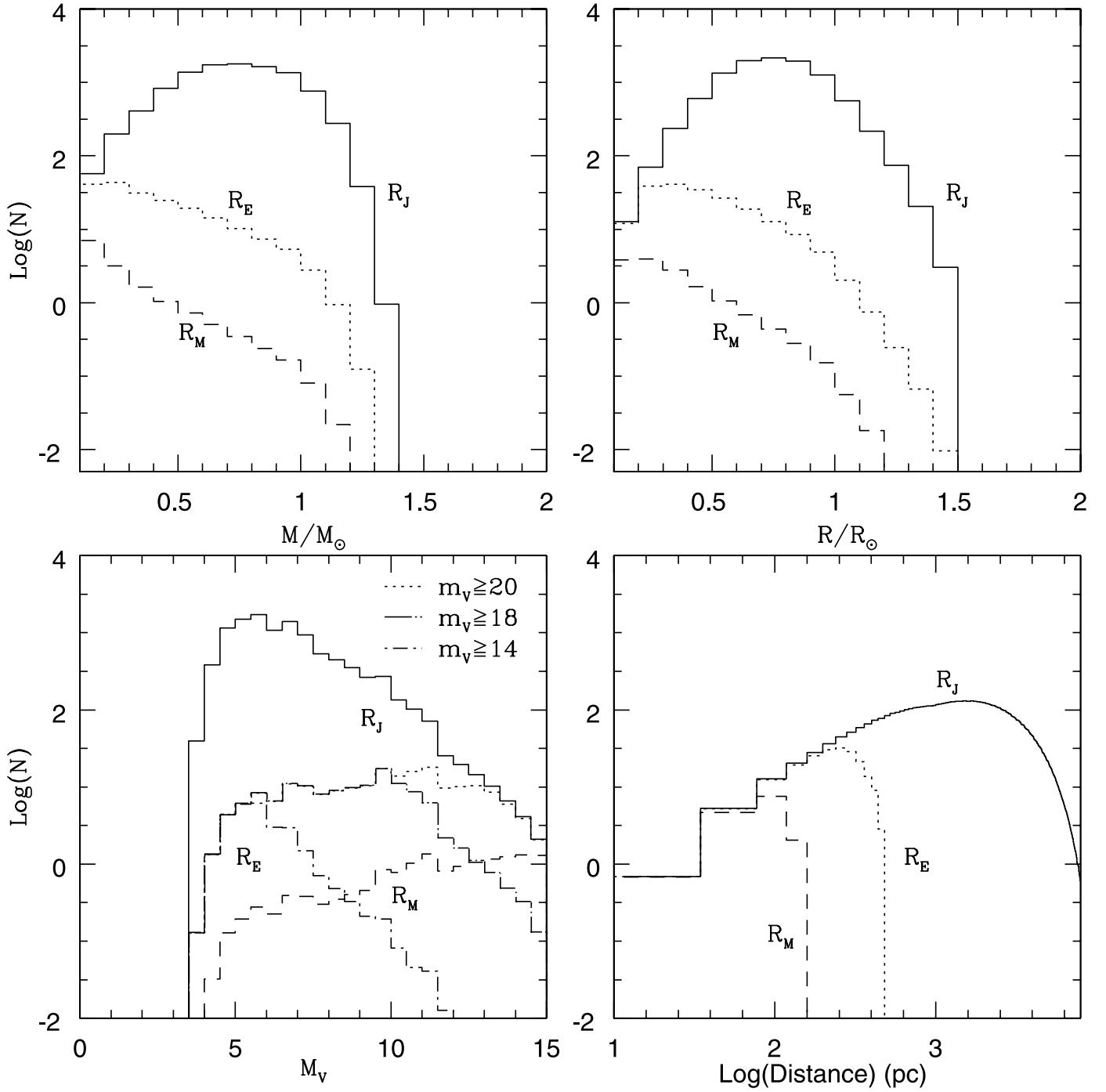| Source	B	A	F	G	K	M	All
Jenkins & Doyle 2003	11014	36708	81085	61347	15573	831	206558
Calculated (old field)	4278	31370	76630	62891	18243	1264	194678
Calculated (new field)	338	5066	41094	45568	13351	942	106342

FIG. 7.—Number of *Kepler* detections of Jupiter-, Earth-, and Mars-sized worlds binned against a variety of parameters. These plots assume that every star possesses a planet, and use an apparent *V* magnitude limit of $m_V \geq 20$. In the plot showing the distribution by absolute *V* magnitude of the parent star, we have also plotted distributions showing the effect of changing the apparent magnitude limit on the distribution of detected Earth-sized worlds. [*See the electronic edition of the Journal for a color version of this figure.*]

TABLE 14
EXPECTED TRANSITING HJs AND VHJs IN *Kepler*'s FIELD OF VIEW

Source	B	A	F	G	K	M	All (HJs)
Jenkins & Doyle 2003.................	10	32	71	54	14	1	181 (181)
Calculated (old field)..................	4.2	27.3	55.5	33	8	0.4	127 (72.8)
Calculated (new field).................	0.3	4.4	28.6	23.5	5.8	0.33	62.9 (36)

TABLE 15
ALL-SKY SURVEY PARAMETERS

Survey	LSST	SDSS-II	Pan-STARRS	Pan-STARRS Wide
m_V limit (Sun-like)	18.51	15.56	14.99	12.49
m_V limit (M dwarfs).........................	23.13	20.18	19.61	17.11
Observation time..............................	10 yr	37.5 days	5 months	5 months
Telescope diameter............................	6.5 m	2.5 m	1.8 m	1.8 m
Throughput...	0.5	0.5	0.5	0.5
Field of view	9.6 deg^2	6.25 deg^2	7 deg^2	7 deg^2
Area surveyed	20,000 deg^2	300 deg^2	1200 deg^2	12,000 deg^2

calculations, as well as the other relevant parameters for each of the surveys. We calculated magnitude limits for detecting giant planets around both Sun-like ($0.8 \leq M_*/M_\oplus \leq 1.2$) stars and M dwarfs ($0.1 \leq M_*/M_\oplus \leq 0.5$) for S/N ≥ 20.

To translate these magnitude limits into the number of transiting planets each of the surveys should detect, we used our methodology to calculate the density of VHJs and HJs for various limiting magnitudes, in numbers per square degree, along the Galactic plane, at the Galactic poles, and averaged around the sky. Table 16 shows the resultant values for a series of fiducial magnitude limits, as well as the planet densities at the magnitude limits of our modeled surveys for Sun-like stars; Table 17 shows transiting planet densities around M dwarfs. In addition, in Figure 8 we show the transiting planet density around Sun-like stars and M dwarfs as a function of Galactic latitude. It is interesting to note the dip near the Galactic plane, which occurs as a result of extinction from interstellar dust.

Knowing the angular size of each survey, we use these densities to predict the number of planets that will be seen. For SDSS-II and Pan-STARRS, which are both pointed toward the Galactic poles, we used the polar densities, while for LSST and the wide Pan-STARRS, with their much larger and wide-ranging survey area, we used the all-sky average density. The final predictions for both Sun-like stars and M dwarfs are shown in Table 18.

We note that these predictions are based on a simpler model than what we have used in other portions of this paper. For instance, we do not include the effects of CCD saturation or sky background. Nevertheless, our calculations should provide a good expectation for the range of the detection numbers that these all-sky surveys can expect.

The large difference between the number of planets expected to be detected by LSST and the other surveys is caused by two effects. First, SDSS-II and Pan-STARRS cover much smaller areas of the sky to shallower depths than LSST, and second, LSST is not primarily directed toward the Galactic poles, greatly increasing the density of stars for which planet detection is possible.

Pan-STARRS wide, meanwhile, has a substantially brighter magnitude limit than any of the other surveys, which prevents it from detecting as many planets as LSST, despite the similarities in the size of their survey areas. Pan-STARRS wide does detect more planets than the planned narrow survey, primarily because the latter "runs up" against the scale height of the Galaxy.

The real difficulty, however, in managing the 7740 or 15,530 detections that we expect LSST to make will be properly identifying and separating the false positives from the actual planetary transit events. Indeed, to date the wide-angle surveys have generated approximately 15 false positive candidates for every actual transiting planet they have seen (D. W. Latham 2007, private communication). If we extend this ratio to LSST, SDSS-II, Pan-STARRS, and Pan-STARRS wide, we would expect them to create 116,000, 90, 290, and 720 false positives, respectively around Sun-like stars, and 233,000, 180, 550, and 1200 around M dwarfs, respectively. Working through this many transit candidates (especially in the case of LSST) is an intensive, time-consuming task. We would therefore expect the all-sky surveys, despite the high absolute number of predicted detections, to discover transiting planets at about the same rate as the operational wide-angle surveys, as the identification rate is predominantly limited by the ability to make follow-up observations of candidates to determine whether or not it is a false positive.

4.5. Effects of Red Noise

As pointed out by Pont et al. (2006) the general mathematical description we have used for the S/N of a transit, equation (4), assumes uncorrelated errors on the collected photometric data points (white noise). In practice this may not be the case, as phenomenon such as changes in the seeing, airmass, telescope tracking, etc. introduce systematic trends into the observed data that are correlated on the timescales of transits (red noise). This has the effect of substantially decreasing the effective S/N displayed by a transit, thus making detections more difficult. Specifically, Pont et al. (2006) show that to account for red noise, equation (4)

TABLE 16
TRANSITING VHJ AND HJ DENSITIES AROUND SUN-LIKE (G2) STARS (deg^{-2})

Magnitude Limit	Galactic Plane	Galactic Poles	All-Sky Average
$m_V \leq 20$	4.052	0.027	0.800
LSST (18.51)	1.609	0.027	0.387
$m_V \leq 18$	1.125	0.026	0.293
$m_V \leq 16$	0.219	0.025	0.087
SDSS-II (15.56)	0.146	0.020	0.063
Pan-STARRS (14.99)........................	0.083	0.016	0.041
$m_V \leq 14$	0.029	0.009	0.017
Pan-STARRS Wide (12.49)...............	0.005	0.003	0.004
$m_V \leq 12$	0.003	0.002	0.002

TABLE 17
Transiting VHJ and HJ Densities Around M dwarfs (M5) (deg^{-2})

Magnitude Limit	Galactic Plane	Galactic Poles	All-Sky Average
LSST (23.13)	2.7558	0.0888	0.7764
SDSS-II (20.18)	0.2473	0.0398	0.1136
$m_V \leq 20$	0.2081	0.0368	0.0989
Pan-STARRS (19.61)	0.1422	0.0304	0.0725
$m_V \leq 18$	0.0257	0.0105	0.0169
Pan-STARRS Wide (17.11)	0.0092	0.0048	0.0068
$m_V \leq 16$	0.0047	0.0015	0.0028
$m_V \leq 14$	0.00017	0.00015	0.00016
$m_V \leq 12$	0.00001	0.00001	0.00001

must be modified by adding a term to include a covariance coefficient C_{ij} between the ith and jth measurements,

$$\chi^2 = \delta^2 \left(\frac{\sigma^2}{N_{tr}} + \frac{1}{N_{tr}^2} \sum_{i \neq j} C_{ij} \right)^{-1}. \qquad (29)$$

This is equation (4) in Pont et al. (2006) using the variables we previously defined for our equation (4). The authors give a more

manageable version on page 8 of their work, which relates the S/N with red noise to the S/N without red noise:

$$\chi_r^2 = \chi_w^2 \left[1 + N_k \left(\frac{\sigma_r}{\sigma} \right)^2 \right], \qquad (30)$$

where χ_w^2 is the white noise χ^2 statistic we calculate in equation (4), N_k is the number of data points in the single kth transit
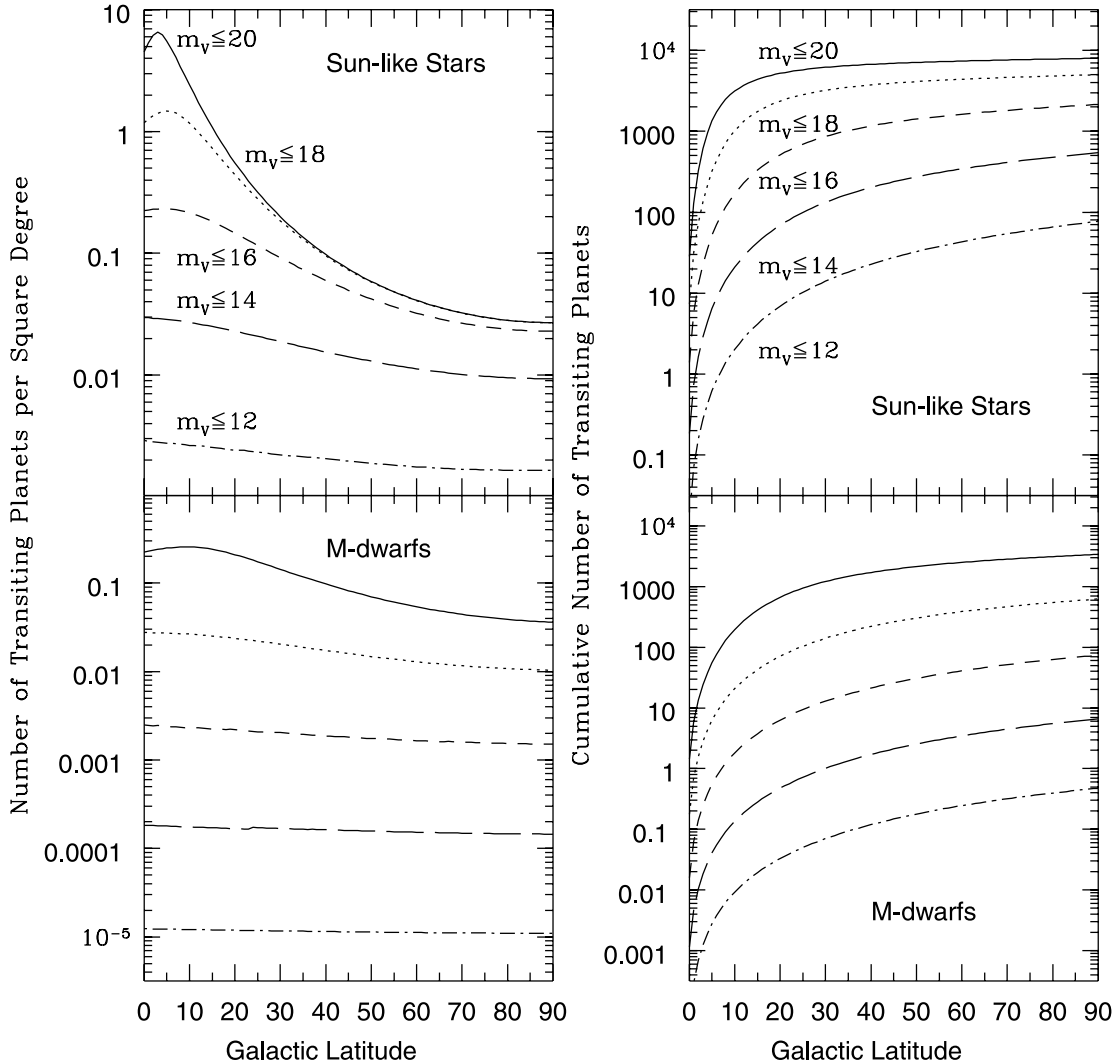
Fig. 8.— Number of transiting planets per square degree around Sun-like stars and M dwarfs for several V magnitude limits. *Left*: Differential plots showing the value at a particular Galactic latitude. *Right*: Cumulative plots showing the total planets from zero degrees up to a particular Galactic latitude.

TABLE 18
ALL-SKY NUMBER OF VHJs AND HJs DETECTED

Survey	Sun-Like (G2)	M Dwarfs (M5)
LSST	7740	15530
SDSS-II	6.0	11.9
Pan-STARRS.......................	19.2	36.5
Pan-STARRS wide	48.0	81.6

TABLE 20
Kepler AND RED NOISE

σ_r	Habitable Earths
10^{-4}	14.37
5×10^{-5}	32.53
10^{-5}	72.88
0	78.89

(which on average is just $N_{\rm tr}$/number of transits), and σ_r is the red component of the noise in the data.

To properly use this formulation, we must therefore have some idea of what value σ_r takes. Pont et al. (2006, 2007) infer values of 3 mmag of red noise for typical ground-based surveys, which can be reduced to approximately 2 mmag after applying de-trending algorithms to the data. We therefore took this tentative value of $\sigma_r = 0.002$ and used it to resimulate the TrES fields to see what effect red noise would have on detection numbers. As noted in Pont et al. (2007) we set a signal-to-noise ratio limit of $S/N_r \geq 12$, which is the S/N_r value of the detection of TrES-1b; TrES-2b has a higher value of $S/N_r = 14$. The results are shown in Table 19. Figure 2 shows (in addition to TrES with only white noise) the distribution of detections using 2 mmag of red noise and $S/N_r \geq 12$.

Compared to a white S/N limit of $S/N_w \geq 30$, using $\sigma_r = 2$ mmag and $S/N_r \geq 12$ drops the total number of expected detections by about 20%. Looking at the split between HJs and VHJs, the HJ population is disproportionately effected, which is a result of the relatively fewer transits HJs show, lowering the overall S/N of their transit signal. Similarly, Figure 2 shows that red noise also cuts out the detections around higher mass stars, as these have a lower S/N overall. Without de-trending the data, a red noise of 3 mmag would cause an even further reduction in detections, down 60% as compared to white noise.

We also examined the possible effect that red noise would have on the yield of habitable Earth-radius planets from *Kepler* mission. Given that *Kepler* is a space-based mission designed to detect transit signals with fractional depths at the level of $\sim 10^{-4}$, we would expect that the amount of red noise due to instrumental and environmental sources would be small, if it exists at all. However, this is generally difficult to predict before launch. Furthermore, there may exist significant astrophysical red noise due to intrinsic variability.

We therefore estimated the expected *Kepler* yield for a set of three values for σ_r, assuming that every star has an Earth-radius planet with $a = {\rm AU}(L/L_\odot)^{1/2}$ (Table 20). For a relatively large amount of red noise at $\sigma_r = 10^{-4}$, we find the number of habitable Earths radius is drastically reduced from our baseline (uncorrelated noise) prediction, from 78.89 down to 14.37. This is not surprising, since Earth passing in front of a Sun-like star has a transit depth of only 8.5×10^{-5}. For this amount of red noise, *Kepler* is restricted to detecting Earth-like planets only around K and M dwarfs. This shifts the expected distribution of host

stars toward dimmer magnitudes (Fig. 6). For smaller values of $\sigma_r = 5 \times 10^{-5}$ and $\sigma_r = 10^{-5}$, the number of habitable Earths detected is reduced to 32.53 and 72.88, respectively. Both of these estimates also assume that every target star has an Earth-sized planet in the habitable zone, and therefore must be viewed as upper limits.

Thus, we conclude that *Kepler*'s yield of habitable zones of habitable planets will be not be strongly affected by red noise, provided that the red noise is $\lesssim 10^{-5}$ on the timescales of the transits, which is ~ 13 hr $(M/M_\odot)^{1.75}$ for solar-type stars. If the red noise is found to be larger than this, it may compromise the primary *Kepler* science by significantly reducing the number of expected habitable planet detections. For example, if the red noise is $\sigma_r = 10^{-4}$, and the frequency of habitable planets with $r = R_\oplus$ is ~ 20%, there is a ~ 5% chance that *Kepler* will not detect any. The effects of red noise on the habitable planet yield can be mitigated by including a larger number of K and M dwarfs in the target sample. Of course, it will also be possible to extrapolate the frequency of shorter period or more massive planets, where the effects of red noise are mitigated, into the habitable zone.

5. CONCLUSION

In this paper we have developed a method to more accurately predict the number of transiting planets detectable by wide-field ground- and space-based photometric surveys. We have taken into account such factors as the frequency of gas giants around main-sequence stars, the probability of transits, stellar density changes from Galactic structure, and the effects of interstellar extinction. We adopt a S/N detection threshold criterion to approximate the detectability of planetary transits.

To test our model, we compared our predictions to two operating photometric surveys, TrES and XO. For both, we calculated the number of statistically expected detections, as well as the population characteristics of the host stars. For TrES, we overestimate the number of detections (13 vs. the four actual detections), while for XO our simulations yield approximately the actual number of detections (two predicted vs. three actual). The exact reason for the discrepancy between our predictions and the actual TrES results could be one of many. As we discuss in the TrES section, from the properties of the known TrES planets, it may be that the survey has a brighter magnitude cutoff than the one we use in our fiducial case. If this were the case, and we changed the magnitude limit from $m_R \leq 13$ to $m_R \leq 12$, the number of expected detections would drop to 8.16 from 13. Red noise also has the potential

TABLE 19
PREDICTED TrES DETECTIONS WITH 2 mmag RED NOISE, $S/N \geq 12$

Object	And0	Cyg1	Cas0	Per1	UMa0	CrB0	Lyr1	And1	And2	Tau0	UMa1	Her1	Lac0	Total
VHJs........................	0.34	0.33	0.31	0.30	0.17	0.20	0.30	0.22	0.17	0.20	0.16	0.20	0.36	3.26
HJs...........................	0.25	0.22	0.21	0.21	0.11	0.13	0.19	0.12	0.07	0.11	0.11	0.14	0.25	2.15
Both........................	0.59	0.55	0.52	0.51	0.28	0.33	0.49	0.34	0.24	0.31	0.27	0.34	0.61	5.41

to substantially affect TrES; a specific estimate of the magnitude of red-noise in the TrES survey would be worthwhile.

Indeed, it is generally hard to draw specific statistical conclusions from the comparison of our results and those of the transit surveys because they typically do not adopt the strict detection criteria we have used. Promising planet candidates are often followed up even if they are beyond the stated S/N or magnitude limits of the survey. Understanding and quantifying how the survey teams select candidates is vital to appropriately deriving the statistical properties of extrasolar planets. Indeed, Tables 4–6 and 9–11 demonstrate that the actual predictions depend crucially on the specific magnitude and S/N threshold used.

As more transiting planets are discovered, these statistical properties are increasingly becoming the frontier of research, shifting the focus of the field away from individual detections. Given that transiting planets are so far one of the only ways that we may infer the radius of extrasolar planets and their exact orbital properties, the statistical characteristics of this group of objects is one of the only foreseeable ways that the areas of planetary interiors, system dynamics, migration, and formation will acquire more data.

In the future, besides using the preceding formalism to design more efficient transit surveys, we would hope that a more systematic approach toward transit surveys will allow this model to be used to make more specific statistical comparisons. As the number of known transiting planets grows, we would also expect that this formalism will be used to test different distributions of planet frequencies, periods, and radii against those observational results. This will allow us to better understand the statistics of extrasolar planetary systems, improve our ability to find new planets, and help to understand the implications of the ones that have already been detected.

APPENDIX A

PHOTON LUMINOSITY OF A SOURCE STAR

The photon luminosity in a bandpass of a given star of a given mass is calculated exactly by dividing the Planck radiation law for wavelength by the energy of a photon at that wavelength, multiplying the result by the surface area of the star, and then integrating over the wavelength range of the relevant bandpass:

$$\Phi_\lambda(M) = \int_{\lambda_{\min}}^{\lambda_{\max}} \frac{B_\lambda(T_*)}{E_\lambda} 4\pi^2 R_*^2 \, d\lambda, \tag{A1}$$

where $B_\lambda(T_*)$ is the Planck Law, and $E_\lambda = hc/\lambda$ is the energy of a photon. Note that this assumes that stellar radiation can be approximated as a blackbody, a crude assumption that still suffices for our purposes, as we show below.

Because the widths of the bandpasses in the visible and near-infrared that we are considering are generally narrow with respect to $dB_\lambda/d\lambda$, we can safely approximate this integral by the product of the central wavelength of each bandpass λ_c and a bandpass width $\Delta\lambda$:

$$\Phi_\lambda(M) = \frac{B_\lambda(T_*, \lambda_c)\Delta\lambda}{E_{\lambda_c}} 4\pi^2 R_*^2 = \frac{8\pi^2 c R_*^2 \lambda_c^{-4} \Delta\lambda}{\exp\left[hc/\lambda_c k T_*(M)\right] - 1}, \tag{A2}$$

which, once we select a bandpass, only leaves the temperature of the source star $T_*(M)$ to be determined.

Using the Stefan-Boltzmann Law, we may relate the temperature of the source star to its bolometric luminosity and radius:

$$T_* = \left(\frac{L_{\text{bol},*}}{4\pi\sigma R_*^2}\right)^{1/4}, \tag{A3}$$

where R_* and $L_{\text{bol},*}$ are the radius and the bolometric luminosity of the source star, respectively. We chose to treat both as functions of the mass of the source star, as this is an independent variable that allows the use of the present day mass function to describe the general stellar population.

Using stellar data from Harmanec (1988) and Popper (1980) we fit broken power laws for both the $R_*(M)$ and $L_{\text{bol},*}(M)$ relations. For the mass-radius relation, we fit for two breaks in the power law, while the fitting for the mass-luminosity relation was accomplished using only one break. The relations that we thus found are

$$R_*(M) = R_\odot \left(\frac{M}{M_\odot}\right)^\alpha \text{ with } \alpha = \begin{cases} 0.8 & \text{for } M_* \leq 1, \\ 1.5 & \text{for } 1 < M_* \leq 1.5, \\ 0.54 & \text{for } 1.5 < M_*, \end{cases}$$

$$L_{\text{bol},*}(M) = L_{\text{bol},\odot} \left(\frac{M}{M_\odot}\right)^\beta \text{ with } \beta = \begin{cases} 2.5 & \text{for } M_* \leq 0.4, \\ 3.8 & \text{for } 0.4 < M_*. \end{cases} \tag{A4}$$

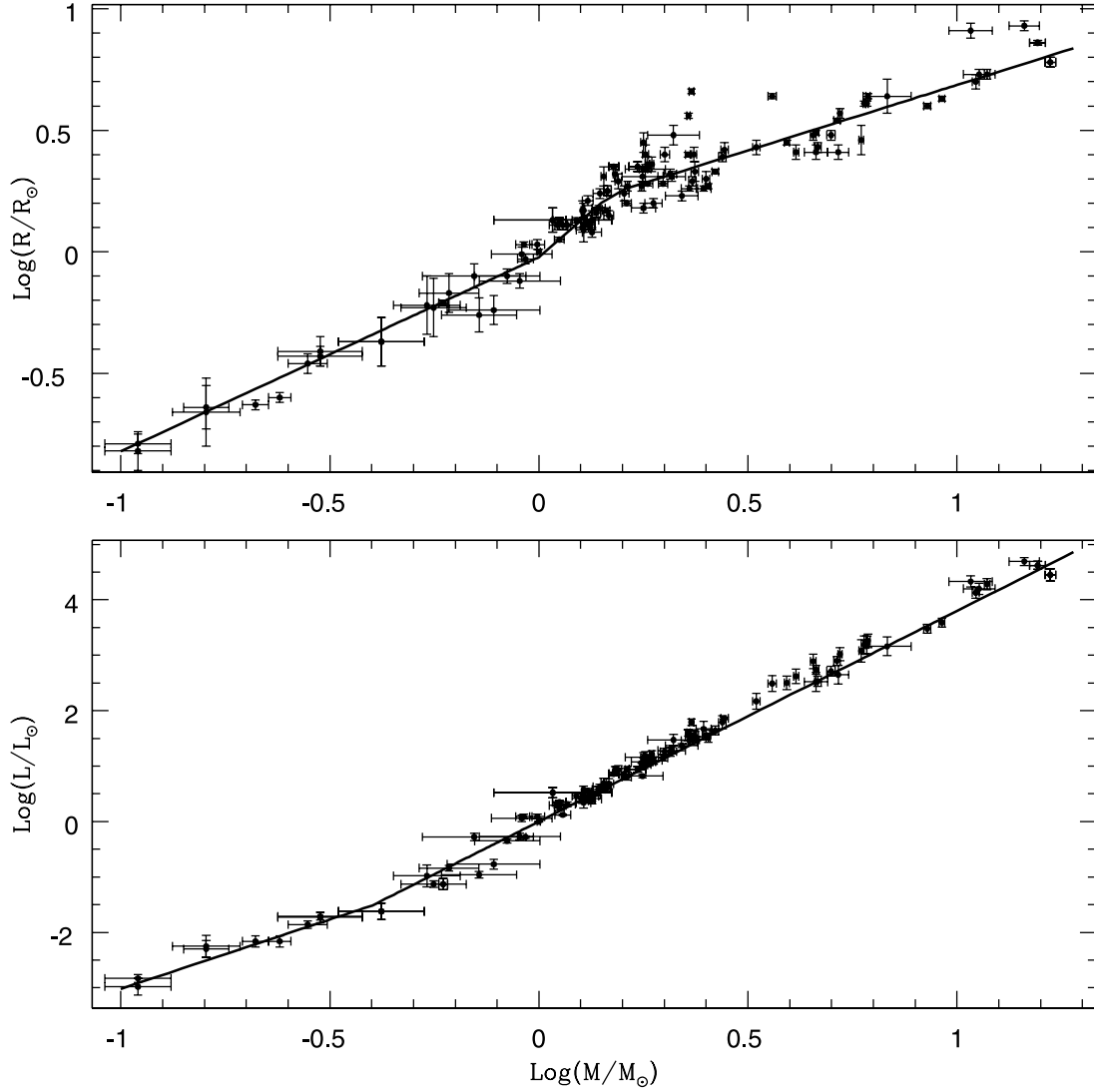Equations (A4) and (A5) are plotted against the Harmanec (1988) and Popper (1980) data in Figure 9.

Fig. 9.—Mass-radius and mass-luminosity relations. The data points from Harmanec (1988) and Popper (1980) are overlaid with our best-fit broken power laws for each. Note the mass-radius fit has two breaks as opposed to only one in the mass-luminosity fit. [*See the electronic edition of the Journal for a color version of this figure.*]

The real test, however, is how well these relations allow for the calculation of the effective temperature of a star at a given mass using equation (A3), as this feeds directly into the determination of the photon luminosity in a given band. Figure 10 shows a plot of our mass versus temperature relation along with data points from Harmanec (1988) and Popper (1980). In addition, we also show the analytical relation provided by Harmanec (1988). One can see that our formulation of equations (A4) and (A5) affords a close approximation to both the data and the relation provided by Harmanec (1988).

Substituting back into equation (A2) allows us to calculate the photon luminosity of the source star, as well as the absolute magnitude of the source star in various bandpasses. If we compare these calculated absolute magnitudes in different bands for various stellar masses against empirical magnitudes for those same masses (Figs. 11 and 12) we can see that the two relations are similar, to within half a magnitude, especially in longer wavelengths. Shorter wavelength bands such as *U*, *B*, and to a certain extent *V*, have calculated magnitudes that diverge from the empirical relation below about a solar mass, which is an unfortunate effect of using blackbody relations to determine the photon luminosity, as in these higher energy bands stellar spectra increasingly do not look like blackbodies. Nevertheless, the blackbody assumption is sufficiently accurate for our purposes.

APPENDIX B

ANALYTIC WINDOW PROBABILITY

Roughly, if we ignore the effects of aliasing, we may say that the window probability of seeing a minimum of $N_{\mathrm{tr,min}}$ transits of a planet with orbital period p over the course of N_n nights may be broken up into four different regimes.

First, if the average number of transits during the total time during which the target star is observed exceeds $N_{\mathrm{tr,min}}$, then it is statistically "certain" that the transit will occur during the times of observation.
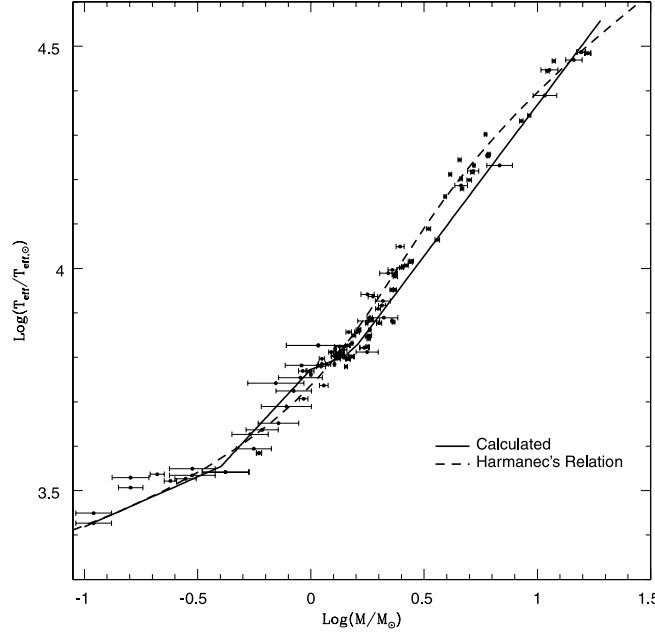
Fɪɢ. 10.—Mass-temperature relation, again using data from Harmanec (1988) and Popper (1980). Overlaid onto the plot are our calculated relationship, as well as the analytical relation described by Harmanec (1988). [*See the electronic edition of the Journal for a color version of this figure.*]

Second, it may be that the entire length of the survey is long enough that there are at least $N_{\rm tr,min}$ transits, but that during the total time spent observing the average number of transits is less than that required for detection. Then one is not certain to detect $N_{\rm tr,min}$ transits, and there is a decreasing window probability (for increasing planet period).

Third, one may have $N_{\rm tr,min}$ transits during an observing run, but only if the first transit occurs sufficiently early during that particular run. The probability of detecting the system then decreases at an even greater rate as the orbital period increases.

Finally, the orbital period of system may be so great that there is no possible way in which one could see the minimum number of transits required for a detection, and $P_{\rm win} = 0$. Mathematically then,

$$
P_{\rm win} = \begin{cases}
1 & \text{for } 1 \leq \dfrac{\langle N_{\rm tr} \rangle}{N_{\rm tr,min}}, \\[2ex]
\left(\dfrac{\langle N_{\rm tr} \rangle}{N_{\rm tr,min}} \right)^2 & \text{for } \dfrac{\langle N_{\rm tr} \rangle}{N_{\rm tr,min}} < 1 \leq \dfrac{(N_n - 1)\Lambda + T_n}{N_{\rm tr,min}P}, \\[2ex]
\left(\dfrac{\langle N_{\rm tr} \rangle}{N_{\rm tr,min}} \right)^2 \left\{ \left[(N_n - 1)\Lambda + T_n - (N_{\rm tr,min} - 1)P \right] (P)^{-1} \right\}^{1/2} & \text{for } \dfrac{(N_n - 1)\Lambda + T_n}{N_{\rm tr,min}P} < 1 \leq \dfrac{(N_n - 1)\Lambda + T_n}{(N_{\rm tr,min} - 1)P}, \\[2ex]
0 & \text{for } \dfrac{(N_n - 1)\Lambda + T_n}{(N_{\rm tr,min} - 1)P} < 1,
\end{cases}
\tag{B1}
$$

where $\Lambda = 1$ day. The term $\langle N_{\rm tr} \rangle$ is the average number of transits that will be seen during the time spent observing, and is given simply by:

$$
\langle N_{\rm tr} \rangle = \frac{N_g T_n}{P}.
\tag{B2}
$$

The value N_g is defined as the total number of "good" observing nights in a particular run, and may mathematically be described as $N_g = N_n - N_{\rm lost}$, where $N_{\rm lost}$ is the number of nights lost due to weather, technical problems, or other issues, and T_n is the length of each night's observing time. Alternatively, one may simply view the quantity $N_g T_n$ as the total amount of good observing time in a given run.

To test the accuracy of our formulation, we integrated and found the average window probability for both the analytic and exact formulations linearly averaged over several period ranges ($1-3$, $3-5$, and $5-10$ days). These are shown as a function of total observing nights under different viewing scenarios in Figure 13. One can see that as a result of the analytic solution not taking into account aliasing, there is divergence between the two solutions as the analytic form approaches unity, as well as in the cases requiring a minimum of three transits.

APPENDIX C

THE PRESENT DAY MASS FUNCTION

A review of the literature shows that while there is agreement that the PDMF beyond 1 $M_\odot$ can be treated as a power law with a slope of $\alpha \approx 5.3$, below a solar mass this agreement breaks down. This is because the low-mass end of the mass function is unfortunately not as well
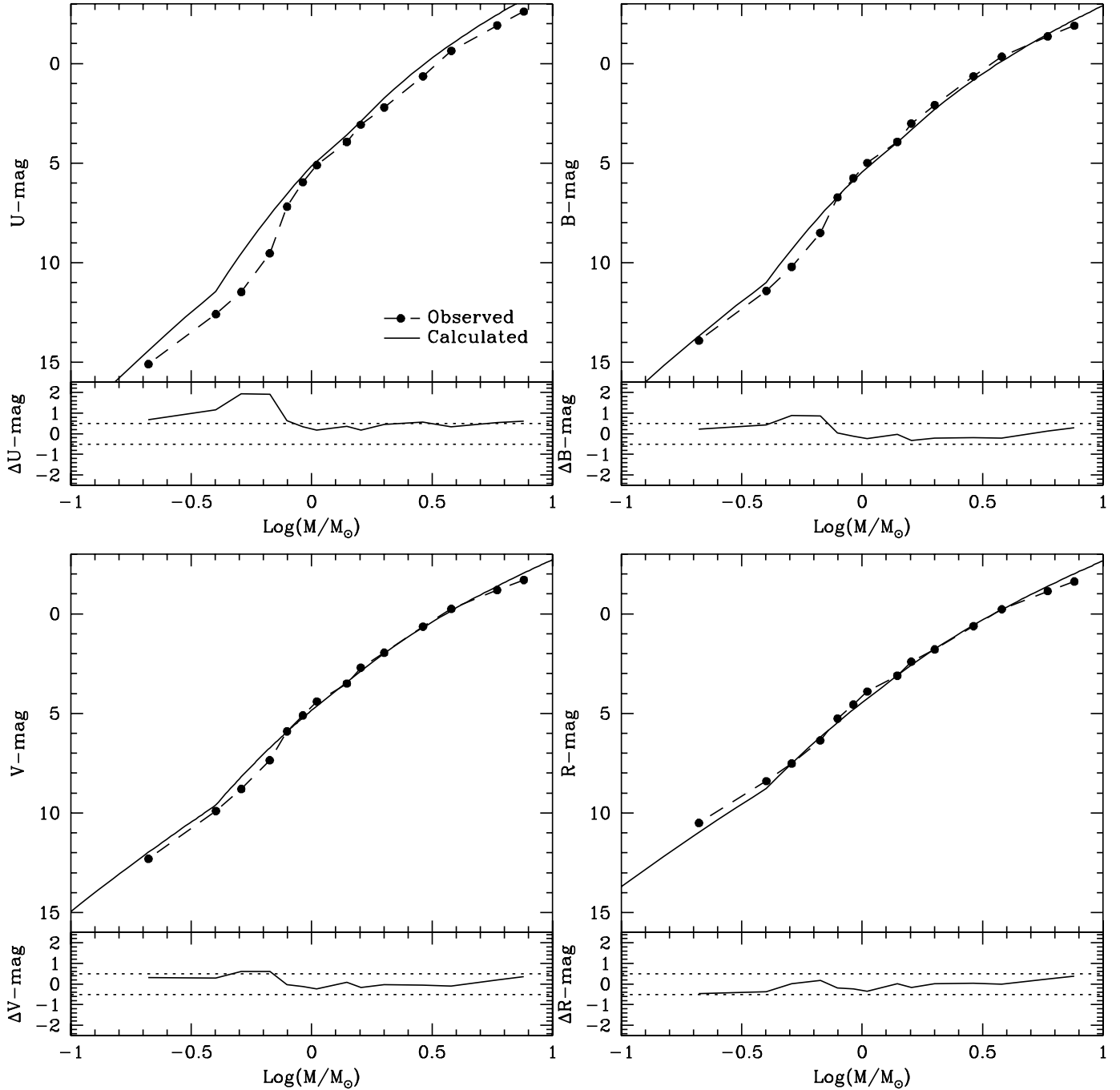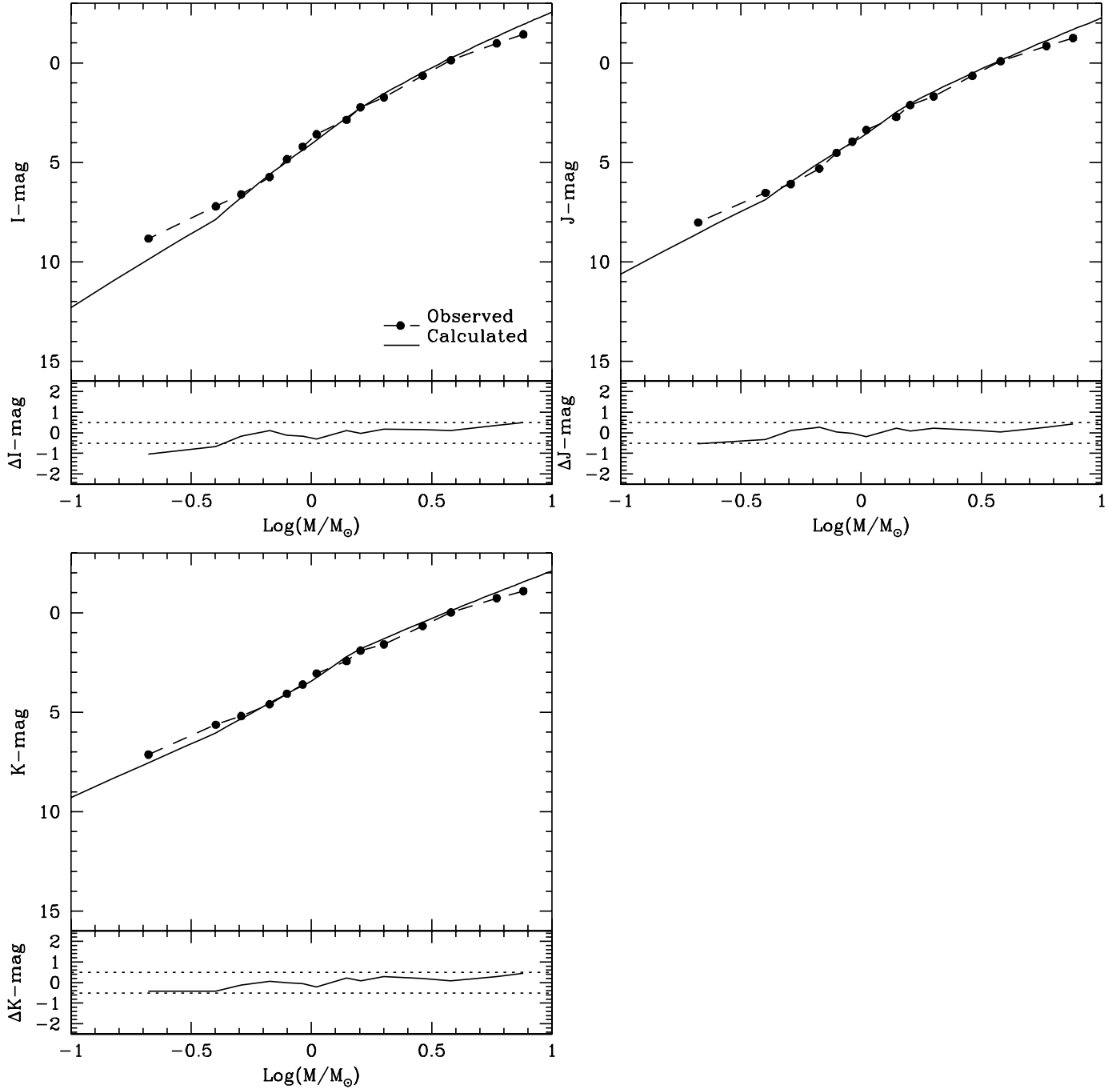
Fig. 11.—Our calculated *U*-, *B*-, *V*-, and *R*-band magnitudes, each as a function of mass, compared to observed stellar magnitudes in those bands (Cox 2000). Below each is a plot of the residuals in that band.

constrained as it is for higher mass stars, reflecting the difficulty of observing dim M dwarfs, as well as uncertainties in the mass-luminosity function (Henry et al. 1999).

For instance, Reid et al. (2002) who used the *Hipparcos* data sets along with their own data from the Palomar/Michigan State University (PMSU) survey to create a volume-limited sample out to 25 pc, propose a traditional power law formulation of the PDMF:

$$\Psi(M) = \frac{dn}{dM} = k_{\mathrm{PDMF}} M^{-\alpha}, \tag{C1}$$

with $\alpha = 5.2 \pm 0.4$ for $M > M_\odot$, and $\alpha = 1.35 \pm 0.2$ for $M \le M_\odot$, and k_{PDMF} as the correct normalization.

In a recent review, meanwhile, Chabrier (2003) argues for a lognormal function below a solar mass:

$$\Psi(M) = \frac{dn}{dM} = \frac{1}{M(\ln 10)} A \exp\left[\frac{-(\log M - \log M_c)^2}{2\sigma^2}\right], \tag{C2}$$

with $A = 0.158^{+0.051}_{-0.046}$, $M_c = 0.079^{+0.021}_{-0.016}$, and $\sigma = 0.69^{+0.05}_{-0.01}$.

Fig. 12.—Our calculated I-, J-, and K-band magnitudes, each as of function of mass, compared to observed stellar magnitudes in those bands (Cox 2000). Below each is a plot of the residuals in that band.

For comparison, if we normalize both to the Scalo (1986) normalization for 5 Gyr of $(dn/dM)_{1\,M_\odot} = k_{\mathrm{PDMF}} = 0.019\ M_\odot^{-1}\ \mathrm{pc}^{-3}$ and plot the two functions against each other (Fig. 14), we can clearly see the divergence both numerically and functionally of these two forms below a solar mass. Interestingly however, for most surveys this divergence makes little difference in the final number of predicted detections: typically Reid's relation gives only 3% fewer detections than Chabrier's does, if both are normalized at a solar mass. That this difference is so small comes from the fact that most wide-angle surveys are not sensitive to detections around stars with $\log(M) < -0.15$. Hence, the region of greatest uncertainty in the PDMF is left unseen by these surveys. Since the difference between the two relations is small, and because the data set used by Reid et al. (2002) is much more recent than that used by Chabrier (2003) we therefore chose to use the Reid et al. (2002) form of the PDMF.

For the purpose of comparing the Reid and Chabrier PDMFs in Figure 14, we normalized both to the value given by Scalo (1986). Going forward, however, we change the normalization to the integrated mass density found observationally by Reid et al. (2002) of $0.032\ M_\odot\ \mathrm{pc}^{-3}$.
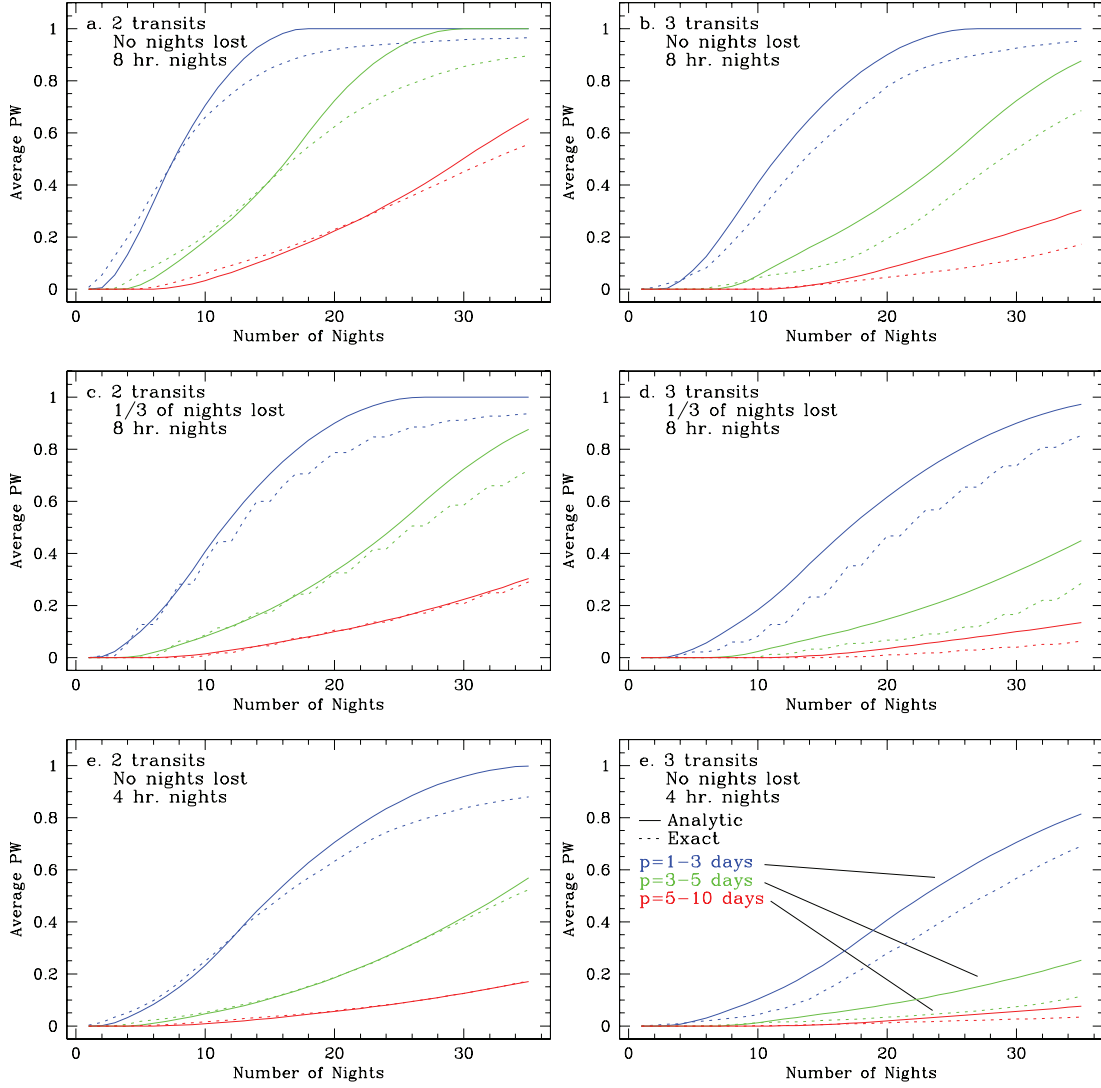
Fig. 13.—Comparison of our analytic formulation of the window probability vs. using an exact formulation to calculate the average window probability as a function of total observing nights averaged over three different orbital period ranges assuming a linear distribution of periods. The left side uses a two transit minimum, while the right requires three transits. (*a, b*) Idealized conditions of 8 hr nights and no observing losses due to weather. (*c, d*) 8 hr nights also, but now 1/3 of the total number of nights are lost due to bad weather. (*e, f*) Effects of shorter 4 hr nights but no losses due to weather.

APPENDIX D

GALACTIC STRUCTURE

The values for the scale length and scale height of the Galactic disk that we use in our modeling are $h_{d,*} = 2.5$ kpc and $h_{z,*}$ as a function of absolute magnitude. Specifically, we calculate the vertical scale height of the disk as:

$$h_{z,*} = \begin{cases} 90 & \text{for } M_V \leq 2, \\ 90 + 200\left(\dfrac{M_V - 2}{3}\right) & \text{for } 2 < M_V < 5, \\ 290 & \text{for } 5 \leq M_V, \end{cases} \tag{D1}$$

which is a modified form of the relation found in Bahcall & Soneira (1980) who instead used a value of 320 pc for the scale height of stars with $5 \leq M_V$.

The large uncertainties in the value of the scale length of the disk reflects the uncertainty in the relevant literature. In their initial paper, Bahcall & Soneira (1980) use $h_{d,*} = 2.5$–3.0 kpc. Since then various sources have proposed an even $h_{d,*} = 3.0$ kpc (Kent et al. 1991), have covered all the bases by using several values between $h_{d,*} = 2.0 - 3.0$ kpc (Olling & Merrifield 2001), and have given the very exact result of $h_{d,*} = 2.264 \pm .09$ kpc (Drimmel & Spergel 2001).

To see what effect different scale lengths have on our final detection numbers, we simulated observations of a $6° \times 6°$ field along the Galactic plane at intervals of $10°$ in Galactic longitude, as well as fixing the Galactic longitude at certain points ($0°$, $90°$, and $180°$) and varying the latitude of the fields. Figure 15 shows these results for scale lengths of 2, 2.5, and 3 kpc. At the extremes (directly toward
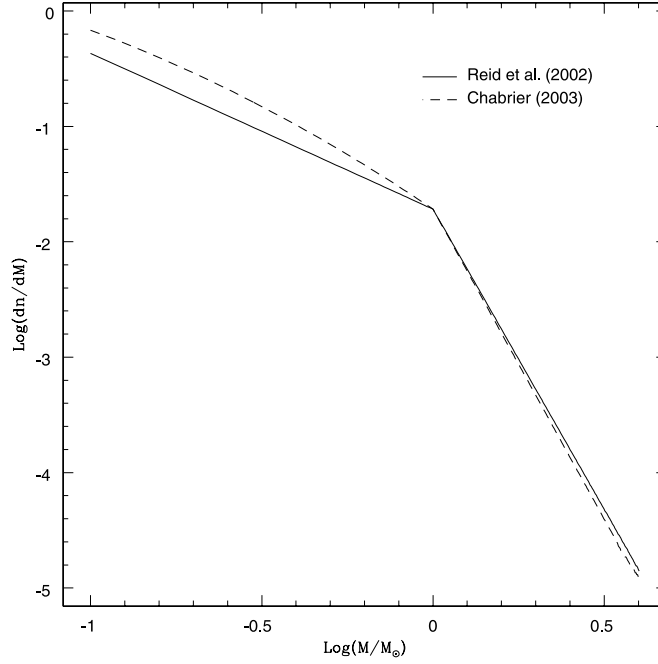
Fɪɢ. 14.—Comparison of the PDMFs proposed by Reid et al. (2002) and Chabrier (2003). Reid et al. (2002) uses a power law (eq. [C1]) below a solar mass, while Chabrier (2003) uses a lognormal function (eq. [C2]).

or directly away from the Galactic core) the 2 and 3 kpc scale lengths differ by only 0.15 detections per field. We chose 2.5 kpc not only to split this difference, but also because more recent results place the scale length much lower than earlier 3 kpc estimates (e.g., Drimmel & Spergel [2001]).

In the case of the scale height of the Galactic disk, the literature divides between those who derive the scale height from optical star counts and those who rely on near-infrared observations. In the optical case, the scale height is presented as a function of absolute visual magnitude that varies from 90 pc to values between 250–320 pc. The near-infrared observations, on the other hand, lead to one single value of the scale height, generally around 290 pc. This happens because measurements in the infrared are dominated by older M dwarfs that have dispersed in an even fashion away from the Galactic plane.

To compare these different estimates, we have plotted the number of detections in a $6° \times 6°$ field along $l = 90°$ as a function of Galactic latitude in Figure 16. We used two of the optically derived scale height relations, as well as a constant scale height of 290 pc. One of the optical relations, the Bahcall model, is given by

$$h_{z,*} = \begin{cases} 90 & \text{for } M_V \leq 2, \\ 90 + 230\left(\dfrac{M_V - 2}{3}\right) & \text{for } 2 < M_V < 5, \\ 320 & \text{for } 5 \leq M_V. \end{cases} \tag{D2}$$

The other optical model is taken from Allen's Astrophysical Quantities (AAQ; Cox 2000) and is

$$h_{z,*} = \begin{cases} 90 & \text{for } M_V \leq 2, \\ 90 + 160\left(\dfrac{M_V - 2}{3}\right) & \text{for } 2 < M_V < 5, \\ 250 & \text{for } 5 \leq M_V. \end{cases} \tag{D3}$$

Also plotted in Figure 16 are the number of detections above and below $M_V = 5$ for the three different scale height relations, as well as the residuals of the constant 290 pc and the AAQ relation against the Bahcall formulation. The exact choice of how the Galactic scale height is treated leads to rather different final predictions. Specifically, while the optical relations are similar, the constant 290 pc scale height from near-infrared observations leads to substantially more detections around stars with $M_V < 5$.

For our predictions, we have adopted the form of the Bahcall & Soneira (1980) and Allen's Astrophysical Quantities scale height relations. We find the gradated scale height in these models appealing, because it accounts for the fact that more luminous stars with shorter lifetimes should not be dispersed far away from star-forming regions in the plane of the Galaxy. The constant 290 pc relation, because it is based on near-infrared data, reflects the sensitivity of those observations to less massive stars that are brighter in longer wavelengths. Nevertheless, we have used the 290 pc result as our value for the scale height of late-type stars with $5 < M_V$, in contrast to the values used in the AAQ and Bahcall models.
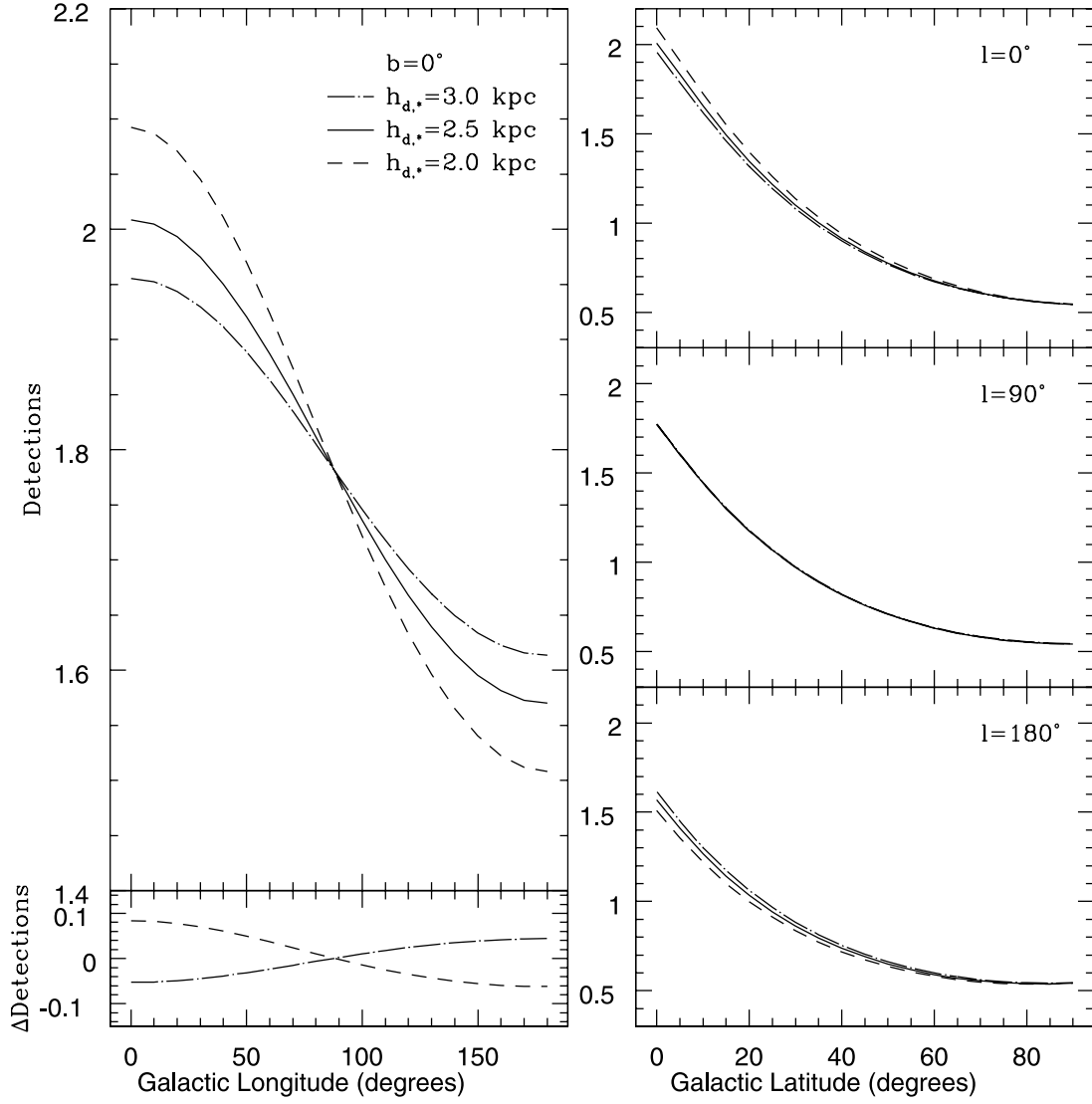
FIG. 15.—Effects of various scale lengths on the predicted number of detections in a $6° \times 6°$ field. *Left*: Detections as a function of Galactic longitude along the Galactic plane, as well as the difference in the number of detections from the assumption of a scale height of 2.5 kpc. *Right*: Detections as a function of Galactic latitude for three different Galactic longitudes. Note the lack of effect of the scale length at $l = 90°$, as well as the convergence of all three at $b = 90°$, as is to be expected.

APPENDIX E

ALL-SKY SURVEYS' LIMITING MAGNITUDES

As per equation (4), we have

$$\sqrt{\chi^2} = \mathrm{S/N} = (N_{\mathrm{tr}})^{1/2} \frac{\delta}{\sigma}. \tag{E1}$$

For the all-sky surveys that we are considering, the number of in-transit data points observed is simply the number of data points acquired in each field multiplied by the fraction of time a planet will spend in transit during its orbit:

$$\mathrm{S/N} = \left(\frac{t_{\mathrm{obs}} \Theta_{\mathrm{FOV}}}{t_{\mathrm{exp}} \Omega_{\mathrm{sur}}} \frac{R_*}{\pi a} \right)^{1/2} \frac{\delta}{\sigma}, \tag{E2}$$

where t_{obs} is the total observation time of the survey, Θ_{FOV} is solid angle of the survey telescope's field of view, and Ω_{sur} is the total solid angle of the sky covered by the survey. Rearranging, we get

$$\sigma = \left(\frac{t_{\mathrm{obs}} \Theta_{\mathrm{FOV}}}{t_{\mathrm{exp}} \Omega_{\mathrm{sur}}} \frac{R_*}{\pi a} \right)^{1/2} \frac{\delta}{\mathrm{S/N}} \tag{E3}$$

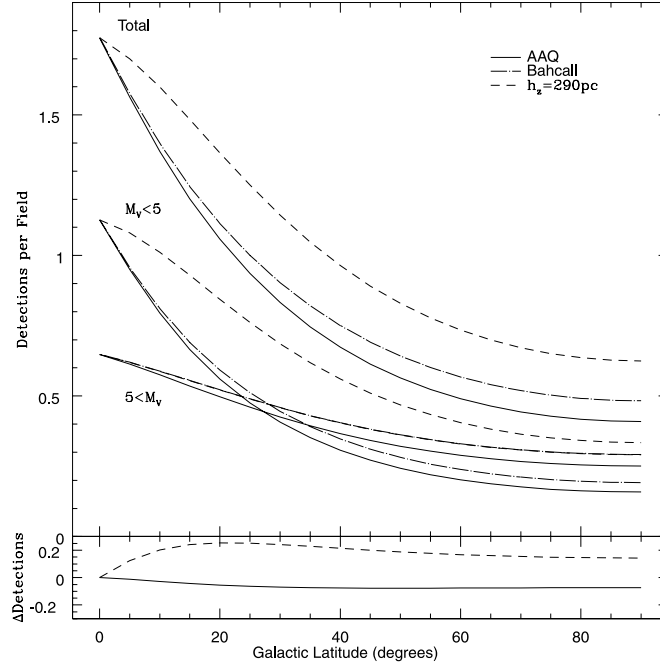Fig. 16.—Total number of detections in a $6° \times 6°$ field as a function of Galactic latitude for a variety of scale height relations. Also shown are the number of detections around stars with $M_V < 5$ and $5 < M_V$, as well as the difference in the predicted number of detections using the constant 290 pc and Allen Astrophysical Quantities (Cox 2000) scale height relations relative to the predictions using Bahcall & Soneira (1980).

For Poisson, source-noise dominated errors, we have

$$\sigma = N_{\rm src}^{-1/2} \tag{E4}$$

where $N_{\rm src}$ is the total number of photons from the source collected during $t_{\rm exp}$. We can then write,

$$\sigma = \left(\frac{t_{\rm exp} e_\lambda}{t_{\rm exp,0}}\right)^{-1/2} \left(\frac{D_{\rm scope}}{D_{\rm scope,0}}\right)^{-1} \sigma_0 10^{0.2(m_V - m_{V,0})} \tag{E5}$$

where σ_0 is the fiducial fractional photometric uncertainty achieved for a star with apparent magnitude $m_{V,0}$ using a telescope with diameter $D_{\rm scope,0}$ and an exposure time of $t_{\rm exp,0}$. We have defined e_λ as the efficiency of the observing setup relative to the fiducial case. If we rearrange equation (E5) to

$$m_V = 5 \log\left[\left(\frac{t_{\rm exp} e_\lambda}{t_{\rm exp,0}}\right)^{1/2} \left(\frac{D_{\rm scope}}{D_{\rm scope,0}}\right) \frac{\sigma}{\sigma_0}\right] + m_{V,0} \tag{E6}$$

we may now substitute in equation (E4) for σ. Doing so, and simplifying, we find that

$$m_V = 5 \log\left[\left(\frac{t_{\rm obs} e_\lambda}{t_{\rm exp,0}} \frac{\Theta_{\rm FOV}}{\Omega_{\rm sur}} \frac{R_*}{\pi a}\right)^{1/2} \frac{D_{\rm scope}\delta}{D_{\rm scope,0}\sigma_0 {\rm S/N}}\right] + m_{V,0} \tag{E7}$$

Which is the limiting magnitude around which an all-sky survey will detect planets.

REFERENCES

Aigrain, S., & Pont, F. 2007, MNRAS, 378, 741
Backer, D. C., Foster, R. S., & Sallmen, S. 1993, Nature, 365, 817
Bahcall, J. N., & Soneira, R. M. 1980, ApJS, 44, 73
Bakos, G., Noyes, R. W., Kovács, G., Stanek, K. Z., Sasselov, D. D., & Domsa, I. 2004, PASP, 116, 266
Bakos, G. Á., et al. 2007, ApJ, 656, 552
Barnes, J. W. 2007, PASP, 119, 986
Beaulieu, J.-P., et al. 2006, Nature, 439, 437
Bond, I. A., et al. 2004, ApJ, 606, L155
Borucki, W., et al. 2004, in Stellar Structure and Habitable Planet Finding, ed. F. Favata, S. Aigrain, & A. Wilson (ESA SP-538; Noordwijk: ESA), 177
Brown, T. M. 2003, ApJ, 593, L125
Brown, T. M., & Charbonneau, D. 1999, BAAS, 31, 1534

Burke, C. J. 2008, ApJ, 679, 1566
Burke, C. J., et al. 2007, ApJ, 671, 2115
Chabrier, G. 2003, PASP, 115, 763
Charbonneau, D., Brown, T. M., Burrows, A., & Laughlin, G. 2007, Protostars and Planets V, ed. B. Reipurth, D. Jewitt, & K. Keil (Tucson: Univ. Arizona Press), 701
Cox, A. N. 2000, Allen's Astrophysical Quantities (4th ed.; New York: AIP)
Cumming, A., Butler, R. P., Marcy, G. W., Vogt, S. S., Wright, J. T., & Fischer, D. A. 2008, PASP, 120, 531
Drimmel, R., & Spergel, D. N. 2001, ApJ, 556, 181
Dunham, E. W., Mandushev, G. I., Taylor, B. W., & Oetiker, B. 2004, PASP, 116, 1072
Duquennoy, A., & Mayor, M. 1991, A&A, 248, 485

Fischer, D. A., & Valenti, J. 2005, ApJ, 622, 1102
Fitzpatrick, E. L. 1999, PASP, 111, 63
Fressin, F., Guillot, T., Morello, V., & Pont, F. 2007, A&A, 475, 729
Gaudi, B. S. 2005, ApJ, 628, L73
Gaudi, B. S., et al. 2008, Science, 319, 927
Gillon, M., Courbin, F., Magain, P., & Borguet, B. 2005, A&A, 442, 731
Gould, A., Dorsher, S., Gaudi, B. S., & Udalski, A. 2006a, Acta Astron., 56, 1
Gould, A., & Morgan, C. W. 2003, ApJ, 585, 1056
Gould, A., Pepper, J., & DePoy, D. L. 2003, ApJ, 594, 533
Gould, A., et al. 2006b, ApJ, 644, L37
Halbwachs, J. L., Mayor, M., Udry, S., & Arenou, F. 2003, A&A, 397, 159
Harmanec, P. 1988, Bull. Astron. Inst. Czechoslovakia, 39, 329
Henry, T. J., Franz, O. G., Wasserman, L. H., Benedict, G. F., Shelus, P. J.,
    Ianna, P. A., Kirkpatrick, J. D., & McCarthy, D. W., Jr. 1999, ApJ, 512, 864
Horne, K. 2003, ASP Conf. Ser. 294, Scientific Frontiers in Research on Extrasolar
    Planets, ed. D. Deming & S. Seager (San Francisco: ASP), 361
Jenkins, J. M., & Doyle, L. R. 2003, ApJ, 595, 429 (JD03)
Johns-Krull, C. M., et al. 2008, ApJ, 677, 657
Kaiser, N. 2004, Proc. SPIE, 5489, 11
Kent, S. M., Dame, T. M., & Fazio, G. 1991, ApJ, 378, 131
Lovis, C., et al. 2006, Nature, 441, 305
Marcy, G., Butler, R. P., Fischer, D., Vogt, S., Wright, J. T., Tinney, C. G., &
    Jones, H. R. A. 2005a, Prog. Theor. Phys. Suppl., 158, 24
Marcy, G. W., Butler, R. P., Vogt, S. S., Fischer, D. A., Henry, G. W., Laughlin,
    G., Wright, J. T., & Johnson, J. A. 2005b, ApJ, 619, 570
McCullough, P. R., Stys, J. E., Valenti, J. A., Fleming, S. W., Janes, K. A., &
    Heasley, J. N. 2005, PASP, 117, 783

McCullough, P. R., et al. 2006, ApJ, 648, 1228
Olling, R. P., & Merrifield, M. R. 2001, MNRAS, 326, 164
Pepper, J., & Gaudi, B. S. 2005, ApJ, 631, 581
Pepper, J., Gould, A., & Depoy, D. L. 2003, Acta Astron., 53, 213
Pepper, J., et al. 2007, PASP, 119, 923
Pollacco, D. L., et al. 2006, PASP, 118, 1407
Pont, F., Zucker, S., & Queloz, D. 2006, MNRAS, 373, 231
Pont, F., et al. 2007, in ASP Conf. Ser. 366, Transiting Extrapolar Planets Workshop,
    ed. C. Afonso, D. Weldrake, & T. Henning (San Francisco: ASP), 3
Popper, D. M. 1980, ARA&A, 18, 115
Reid, I. N., Gizis, J. E., & Hawley, S. L. 2002, AJ, 124, 2721
Robin, A., & Creze, M. 1986, A&A, 157, 71
Sako, M., et al. 2005, preprint (astro-ph/0504455)
Santos, N. C., Israelian, G., & Mayor, M. 2004, A&A, 415, 1153
Scalo, J. M. 1986, Fundam. Cosmic Phys., 11, 1
Struve, F. G. W. 1847, Études d'Astronomie Stellaire: Sur la Voie Lactée et Sur la
    Distance des Étoiles Fixes (St. Petersburg: Impr. l'Académie Impériale Sci.)
Udalski, A., et al. 2005, ApJ, 628, L109
Udry, S., & Santos, N. C. 2007, ARA&A, 45, 397
Udry, S., et al. 2000b, A&A, 356, 590
von Braun, K., & Ciardi, D. R. 2007, in IAU Symp. 249, Exoplanets: Detection,
    Formation and Dynamics, ed. Y.-S. Sun, S. Ferraz-Mello, J.-L. Zhou (Cambridge:
    Cambridge Univ. Press), 93
Wolszczan, A., & Frail, D. A. 1992, Nature, 355, 145
Young, A. T. 1967, AJ, 72, 747