Earth Detection by Astrometry and RV in Multi-Planet Systems

Wesley A. Traub[1], James Marr[1], Alan Boss[2], Andrew Gould[3], Chas Beichman[4], Eric Ford[5], Greg Laughlin[6], Hal Levison[7], Doug Lin[6], Sean Raymond[8], Andrew Boden[4], Valeri Makarov[4], Stefano Casertano[9], Debra Fischer[10], Jeremy Kasdin[11], Matt Mutterspaugh[12], and Michael Shao[1]

[1] Jet Propulsion Laboratory, California Institute of Technology
[2] Carnegie Institution
[3] Ohio State Univ.
[4] NASA Exoplanet Science Institute
[5] Univ. of Florida
[6] Univ. of California, Santa Cruz
[7] Southwest Research Institute
[8] Univ. of Colorado
[9] Space Telescope Science Institute
[10] San Francisco State Univ.
[11] Princeton Univ.
[12] UC Berkeley, & Tennessee State Univ.

Extrasolar Planets in Multi-body Systems: Theory and Observations
Torun, Poland 25-29 August 2008

Overview

- Q: Can SIM-Lite, with RV, detect Earths in multi-planet systems?
- A: Yes, as shown by the following double-blind study.

- Four teams:
    - Team-A: planetary models (5 groups, ~500 systems total)
    - Team-B: data simulation (1 group, 48 systems)
    - Team-C: data analysis (5 groups)
    - Team-D: synthesis (1 group)

- Phase-1: January–August 2008, results presented here.
- Phase-2: Through January 2009.

Participants

Team-A groups

- A-1: Eric Ford, Univ. of Florida
- A-2: Greg Laughlin, UC Santa Cruz
- A-3: Hal Levison, Southwest Research Institute
- A-4: Doug Lin, UC Santa Cruz
- A-5: Sean Raymond, Univ. of Colorado

Team-B

- Andy Boden, Michelson Science Center
- Valeri Makarov, Michelson Science Center

Team-C groups

- C-1: Stefano Casertano, STScI
- C-2: Debra Fischer, San Francisco State Univ.
- C-3: Jeremy Kasdin, Princeton Univ.
- C-4: Matt Muterspaugh, UC Berkeley
- C-5: Mike Shao, JPL

Team-D

- Chair: Wes Traub, JPL
- Vice-Chair: Alan Boss, Carnegie Institution
- Chas Beichman, MSC
- Andy Gould, Ohio State Univ.
- Each PI from Team-C groups

External Independent Readiness Board

- Chair: Vern Weyers, GSFC retired
- Alan Boss, Carnegie Institution
- Ed Groth, Princeton Univ.
- Joseph Wampler, consultant

Constraints

- Fully double-blind exercise.
- Time was very short.
- All teams worked to a common schedule.
  - Team-A groups all delivered on time, but for A-4.
  - Team-C groups all delivered on time, but for C-3.
  - Most experienced team (C5), with a head start, did the best.
  - Expect that with more experience, all teams will do very well.
- Detection by one team is a success for this test.
  - This is not a test of the teams, it is a test of the technique.
- Addressed more than just Solar System analogs so that teams would not know what to expect.
- Definition: expected SNR = RMS_signal / mission_noise

Summary of Blind Test Results

- Inputs: 48 planetary systems (all 1 Sun @ 10 pc).
  - 32 random
  - 8 Solar-system-analogs
  - 4 single terrestrial in HZ
  - 4 no-planets.
  - Noise added to all signals(4 levels for astro, 1 level for RV).
  - Two timelines: (5 yr astro, 15 yr RV) and (10 yr astro, 20 yr RV).
- Outputs: reliability of detections was 40% to 100% (3 teams > 80%)
  - 48 of 95 planets were reasonably detectable, i.e. above threshhold.
  - All were found by at least one team (most by 3 or 4 teams).
  - 16 HZ planets: all found by at least 2 teams.
  - 12 HZ terrestrials: all found by at least 2 teams.

Major Conclusions:

- Single-planet detection is not degraded by presence of other planets.
- Astrometry plus RV can find HZ Earths in multi-planet systems.
- Statistical testing methods need improvement.

Team-A Models

- Team-A groups each generated ~100 model planetary systems using their own planet formation theory.
- These 529 models formed the Random-System data pool.
- We randomly selected 32 systems for this study.
- Models were requested to be consistent with Cumming et al. 2008, i.e., 10.5% of FGK stars have a Jupiter (2<P<200 days, 0.3<M<10 Jup.)

Team-A Inputs: Planet-System Statistics

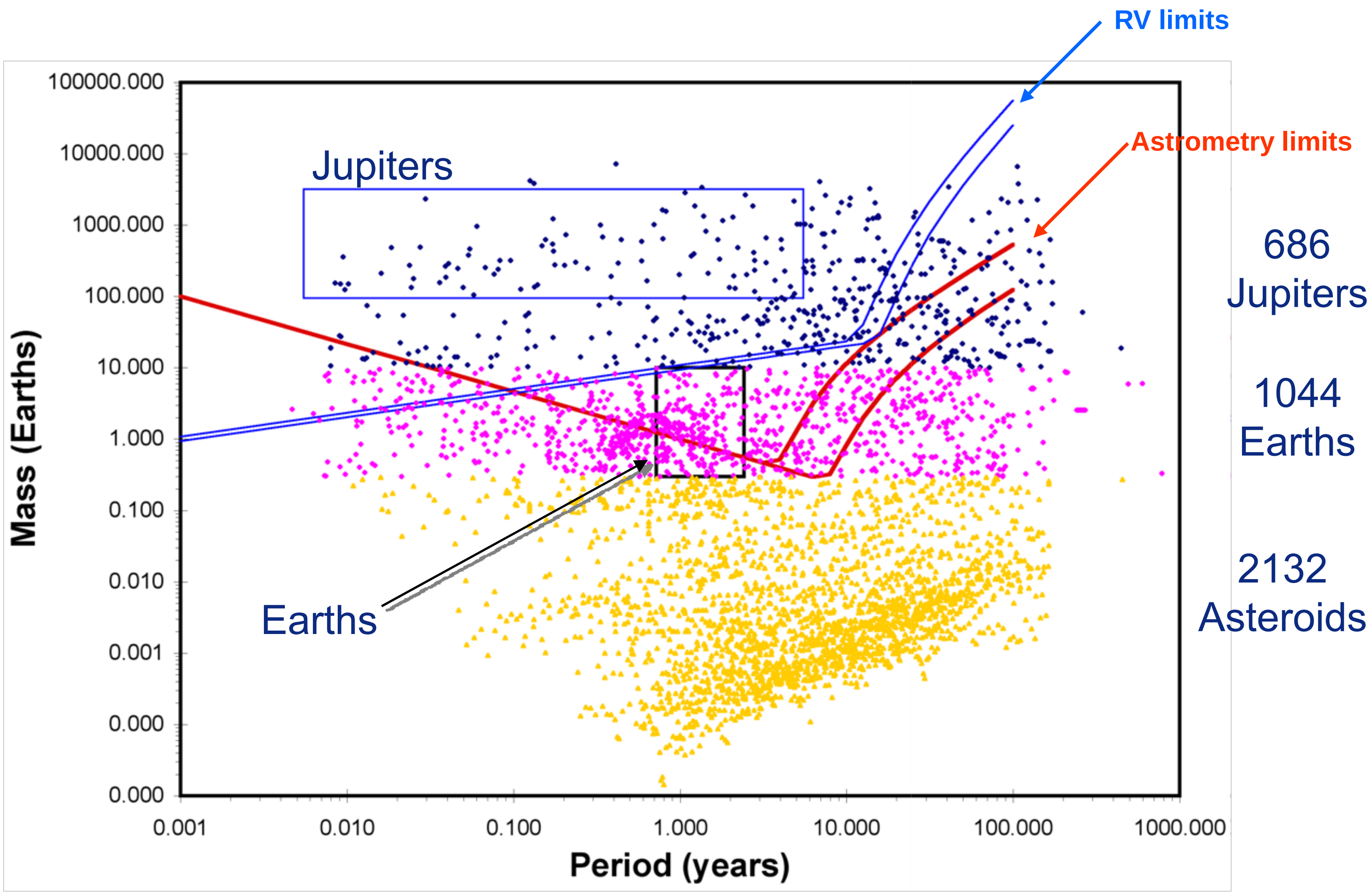
Group	PI	# Stars	Planets /Star (ave.)	Period (median) years	Mass (median) Earths	e (median)	i (median) degrees
A-1	E. Ford	156	5	2	5	0.11	5
A-2	G. Laughlin	159	2	1	1	0.09	4
A-3	H. Levison	74	5	12	17	0.06	2
A-4	D. Lin	190	20	0.6	0.05	0.005	-
A-5	S. Raymond	140	17	6	0.005	0.00	0.06

Medians

Group	Period (min, max) years	Mass (min, max) Earths	e (min, max)	i (min, max) degrees
A-1	0.007 -- 784.	0.05 – 7250.	0.001 – 0.99	0.03 – 175.
A-2	0.008 – 39.	0.001 – 1340.	0 -- 0.49	0.02 – 19.
A-3	0.2 – 270.	0.02 – 1270.	0.001 – 0.93	0.0003 – 58.
A-4	0.003 – 44.	0.01 – 51.	0.0001 – 0.57	---------
A-5	0.005 – 164.	0.00001 – 4060.	0 – 0.71	0 – 42.

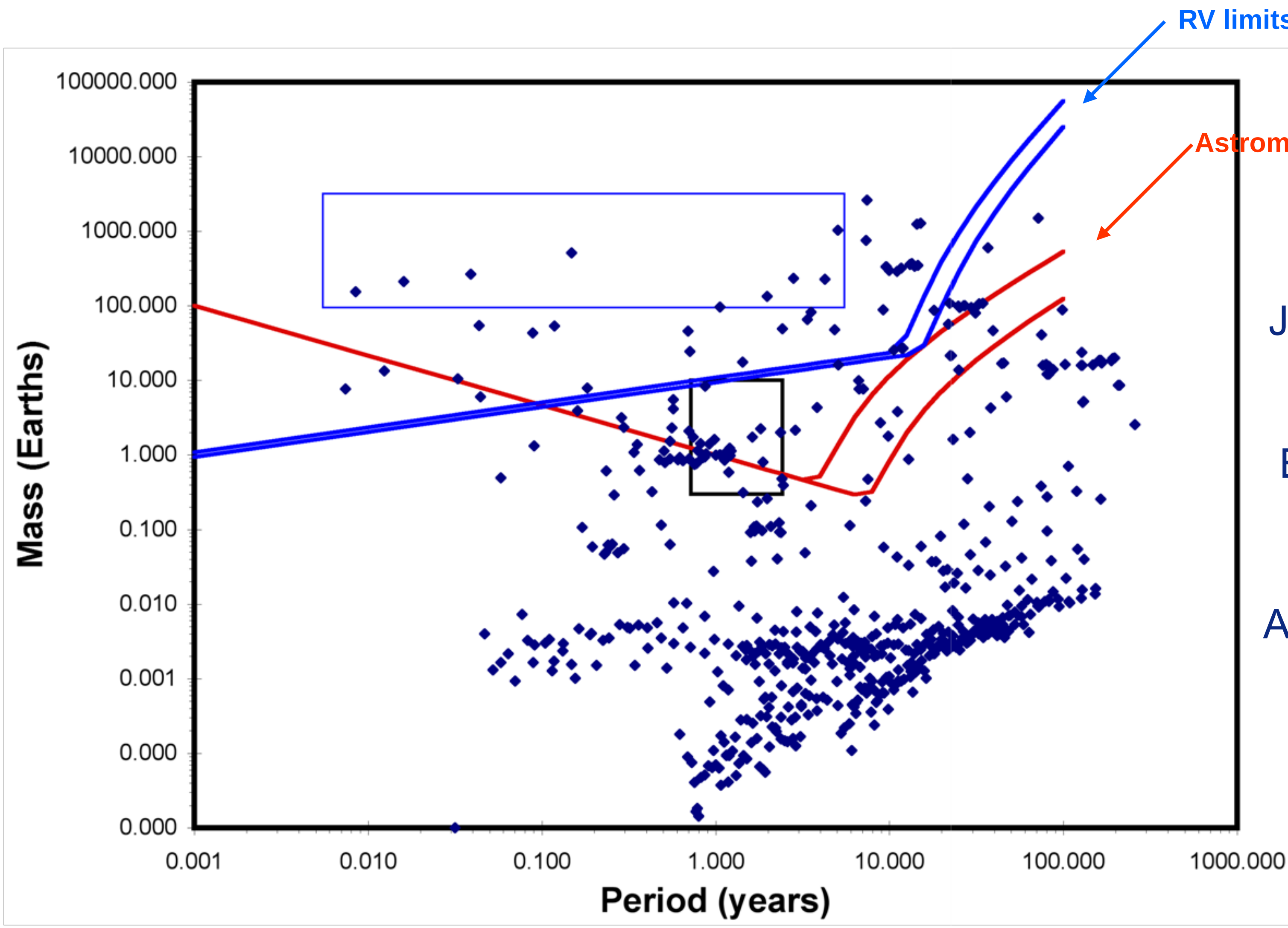
Extremes

Data Pool: 527 stars, 3862 Objects



National Aeronautics and Space Administration

Blind Test Data: 48 stars and 581 objects



82 Jupiters

77 Earths

422 Asteroids

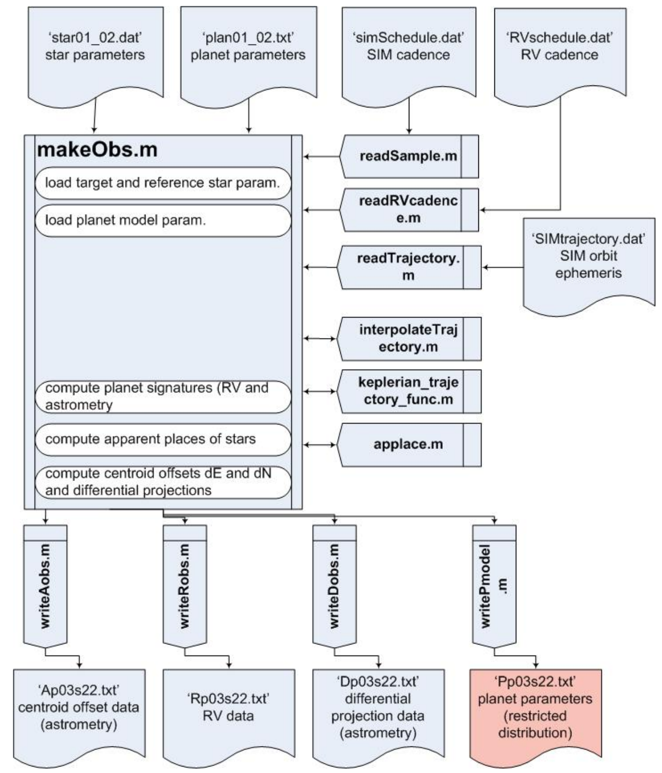
Team-B Synthetic Data

- Planetary Systems:
  - Random, Solar-system analogs, One-Earth, No-Planets.
  - Randomized orientations and orbital phase.
  - Generated synthetic SIM-Lite & RV data.
- Target stars:
  - One solar mass, 10 pc, 30° latitude.
- All significant effects are included in synthetic data:
  - Motion of observer (parallax effect).
  - Space motion of target star (3D space motion).
  - Realistic sampling cadence.
  - Astrophysical noise.
  - Instrument systematic noise.

Remi Soummer, AMNH and GSFC.

Team-B Data Generation

- MATLAB code, many modules.
- Planetary Systems:
  - Random, SS analogs, 1 & 0 Planets.
  - Random orientations & phases.
  - Generated synthetic SIM-Lite & RV data.
- Target stars:
  - One solar mass, 10 pc, 30° latitude.
- All significant effects are included:
  - Motion of observer (parallax effect).
  - Space motion of target star in 3D.
  - Realistic sampling cadence.
  - Astrophysical noise.
  - Instrument systematic noise.



Systems *vs* Type & Noise

- Four types: 32 random; 8 Solar-system-analog, 4 single Earths, 4 no planets
- Four levels of astrometric noise and RV @ 1 m/s:

Astro Noise	Planet Mass Detectable*	Random Data Pool	Solar-system clones	Single Terrestrial/HZ	No planet
2.46 μas	3 Earth	16	2	2	2
0.82 μas	1 Earth	16	2	2	2
0.57 μas	0.7 Earth		2		
0.41 μas	0.5 Earth		2		

* 50% probability of detection at SNR=5.8 with 1% FAP at mid-HZ for 1 Sun at 10pc with N=250 samples (visits).

- Two mission durations: 5 yrs, 10 yrs.
  - Note, the 250 samples distributed over the two periods (NOT that there were twice as many samples on the long duration mission as the short).
- Two minor errors in data generation were made, with no impact on the results.

SNR Primer (1 of 2)

The signal to noise ratio (SNR), for astrometry or RV, is defined as

$$SNR = \frac{signal}{mission\_noise} = \frac{\alpha}{\sigma_1 / \sqrt{N}}$$

$$so.........SNR(astro, Earth) = \frac{0.30\mu as}{0.82\mu as / \sqrt{250}} = \frac{0.30}{0.052} = 5.8$$

$$and.......SNR(RV, Earth) = \frac{0.090 m/s}{1.0 m/s/\sqrt{150}} = \frac{0.090}{0.082} = 1.1 << 5.8$$

where α = RMS motion, σ_1 = single-measurement noise, and N = # measurements.

SNR. From statistics, if we have **SNR = 5.8** or more,
then we get **good completeness** (over 50%) and **few false alarms** (under 1%).

SNR Primer (2 of 2)

For a given SNR, we find minimum mass:

$$m(\text{astro}) = (\sigma_1 \times SNR / \alpha_0 \times N^{1/2}) P^{-2/3}$$

$$m(RV) = (\sigma_1 \times SNR / \beta_0 \times N^{1/2}) P^{+1/3}$$

A minimum-variance bound analysis gives the expected uncertainties:

mass: $\sigma_m/m = \text{sqrt}(2) / SNR \sim 1.4 / SNR$

period: $\sigma_P/P = (\text{sqrt}(6)/\pi) \times (P/T) / SNR \sim 0.8 (P/T) / SNR$

For long-period planets, an approximate correction factor is

mass: $1 + ((P/T - 0.70)/0.18)^2$ for $P/T > 0.70$

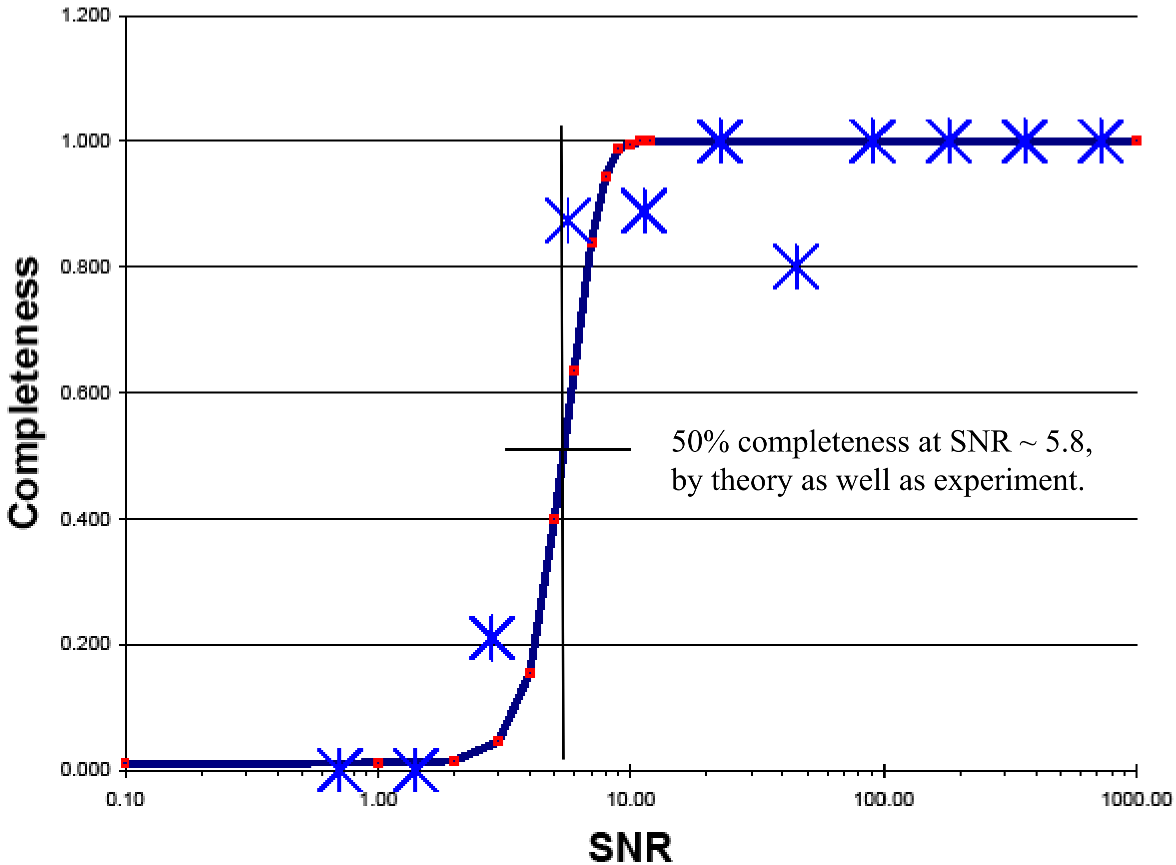
period: $1 + ((P/T - 0.52)/0.27)^2$ for $P/T > 0.52$

So the noise is roughly a factor of 4× worse at $P/T = 1$.

What are the Interesting Questions?

1. Is the expected threshold of SNR~5.8 valid?
2. Do other planets interfere with the detection of HZ terrestrials?
3. What is the reliability of detection (probability that a detection is true)?
4. What is the completeness (probability that a planet will be detected)?
5. We can ask the above questions for
   - All planets,
   - Terrestrial planets,
   - Habitable zone planets and
   - Habitable Terrestrial planets.

1. Completeness of detection (*vs* SNR)



- Completeness is the detected fraction of planets.
- Curve is theoretical for 1% FAP (Catanzarite et al 2006).
- Points are # correct planets / # total planets, for any team.
- Shows that at SNR > 5.8, measured completeness is excellent, as predicted.
- Here SNR is the RSS of the combined RV and Astro SNRs.

2. Errors are ~Gaussian plus a few outliers

For each Team-C group, when a good, marginal, or "clean-up" planet was reported, we calculated the period and mass offsets in units of the expected astro + RV error. A few cases with $\sigma_x/x > 1$ were rejected, and $\sigma_x/x<0.01$ were set to 0.01; little change.

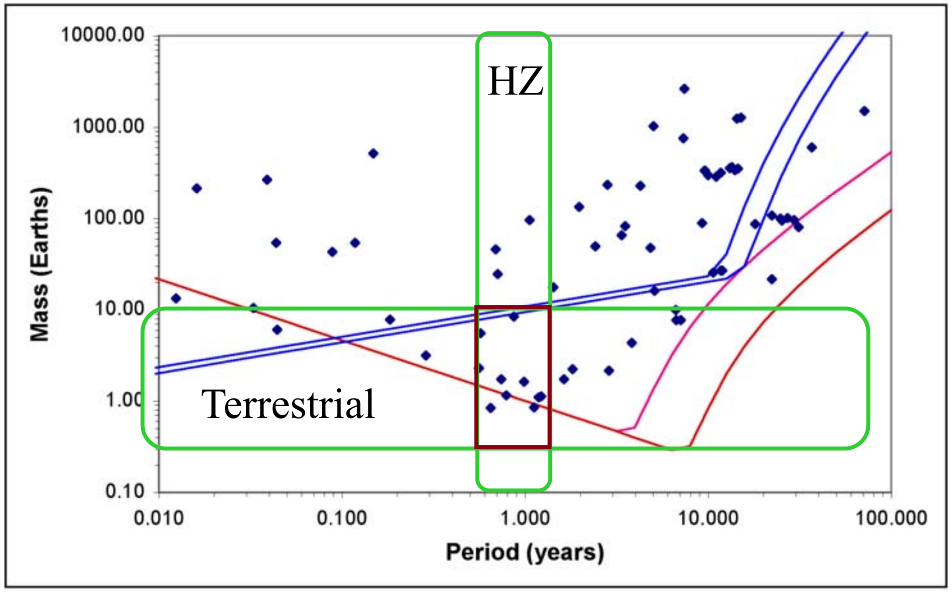
Table entries are number of detections per group

x(obs)-x(model) / sig(model) =	<-3	-3 to -2	-2 to -1	-1 to 0	0 to 1	1 to 2	2 to 3	>3
Period C1	3	2	9	11	15	4	1	6
Period C2	1	2	4	31	31	3	1	0
Period C4	0	1	0	29	34	2	0	2
Period C5	1	1	1	25	33	3	1	1
Mass C1	7	3	2	7	7	10	0	13
Mass C2	5	0	5	18	18	12	7	6
Mass C4	2	2	4	16	14	8	2	17
Mass C5	3	1	7	13	18	12	4	6
~Gaussian	0.1	1	9	22	22	9	1	0.1

Periods are slightly better than expected, masses slightly worse. May need a better theory. But this data suggests that planets are as detectable in multi-planet systems as in single-planet ones.

3. Reliability (*vs* planet type)

- Astrometric & RV detection uses a periodogram in the presence of noise.
  - A low threshold increases detections, but also increases false alarms.
- Reliability: if we claim to see a planet, what is the probability that it is true?
- Define: reliability = #detected **/** (#detected + false alarms)

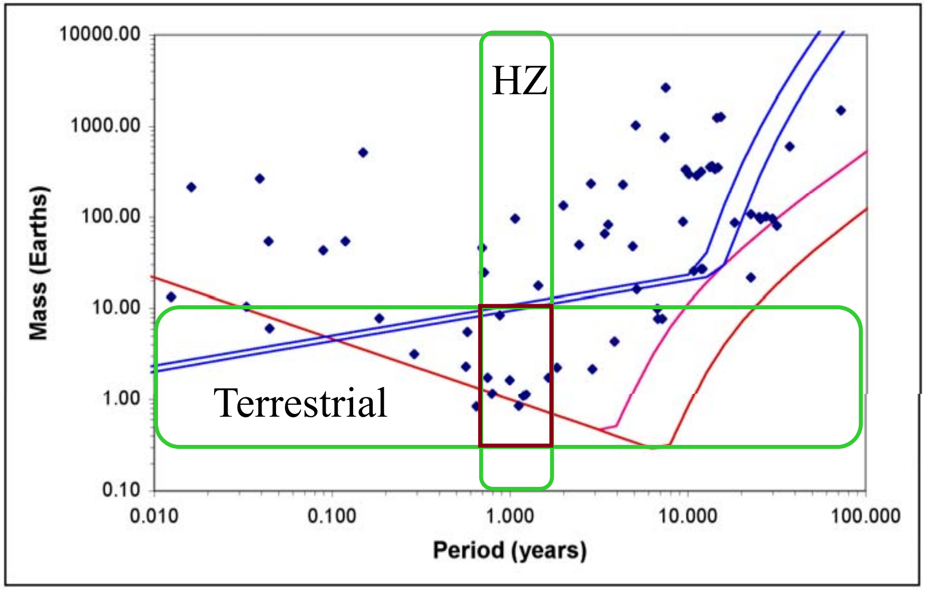


Reliability	Team C1	Team C2	Team C4	Team C5
All	70%	87%	89%	98%
Terrestrial	41%	86%	80%	96%
HZ	44%	76%	79%	100%
Terr & HZ	40%	80%	71%	100%

Figure shows SNR-based detection limits for
RV (blue, upper) 5 & 10 yrs and
SIM-Lite (red, lower) 5 & 10 yrs.

4. Completeness (*vs* planet type)

- There are 70 high SNR (>5.8) planets (plotted).
- 48 of these have a period shorter than 10 years.
- We should have detected all of these, and we did.
- Define: completeness = #detected / #detectable



Completeness	Team C1	Team C2	Team C4	Team C5
All	60%	91%	89%	95%
Terrestrial	28%	81%	81%	90%
HZ	53%	84%	84%	100%
Terr & HZ	42%	71%	71%	100%

Chart shows SNR-based detection limits for
RV (blue, upper), 5 & 10 yrs and
SIM-Lite (red, lower) 5 & 10 yrs.

Blind Test Summary/Conclusion

- Study set out to determine:
  - Can Earths be detected in Solar Systems at 10 pc?

    Answer: Yes
  - What is the sensitivity needed to detect Earths?

    Answer: 40% of 5-year, 0.82 μas, 6-m astrometry mission plus 15 years of RV data on ~60 stars.
- Study was constrained by time.
  - Given more time, team performance expected to improve and converge to the best team's results.
  - Phase 2 will address additional important questions.

Tentative Plans for Phase 2 Study

- Double-blind style continued.
- Extend study to real target stars.
- Complete by Winter AAS.
- Improve detection criteria (e.g., F-Test, stability, others TBD).
- Build theoretical model of astro plus RV.

- Future topics:
  - non-gaussian, non-stationary instrumental noise
  - explicit astrophysical noise
  - prediction of planet position for imaging instruments