A Curios Case Of Solar System Surface Gravities Clustering

Eugene V. Bobukh

Abstract

The values of **Log**(**g**) for 14 largest Solar System planetary bodies appear to form four notable clusters, clearly separated by three large gaps.

Under the assumption of **Log(g)** inherent continuity, two statistical methods were applied to assess the likelihood of producing the same or stronger clustering with 14 randomly chosen gravity values. Both arrived at the answer of **p** = $8*10^{-4}$. With giant planets excluded from the analysis, that probability remained small at **p** ≈ $(5\text{-}6)*10^{-3}$.

Several other distributions – such as known exoplanetary gravities and masses/radii/densities of Solar System bodies – were tested with the same methods. None of them conclusively demonstrated "unlikelihood" of clustering stronger than **p** = 0.05.

While confirming that the perceived Solar System surface gravities clustering is likely real, these results could only be partially explained.

Introduction

Planetary surface gravity is a complex parameter depending on planetary mass and radius, which in turn depend on body's composition and formation history. As a variable of a complex origin, it might be intuitively expected to follow a continuous distribution, with no prominent gaps or spikes.

Yet the **Log**(**g**) distribution of large Solar System planetary bodies appears to form four prominent clusters, well separated by large gaps:

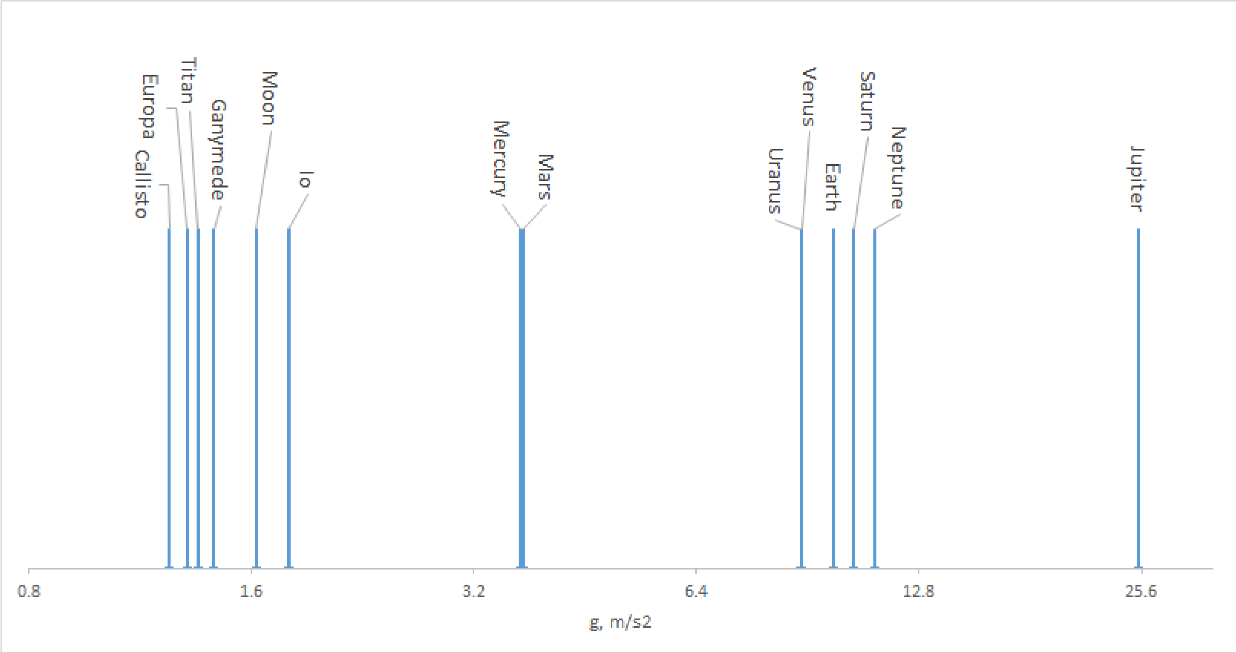


Figure 1 Gravities distribution for Solar System bodies with g > 0.8 m/s², log scale

Those groups consist of:

- Group 1: Moon, Io, Europa, Ganymede, Callisto, and Titan – loosely centered around 1.6 m/s²
- Group 2: Mercury and Mars, at 3.7 m/s²
- Group 3: Venus, Earth, Saturn[1], Uranus, Neptune – near 10 m/s²
- Group 4: Jupiter with gravity of 25 m/s²

It is understood that these are objects of very different nature and formation mode. But that makes it even more surprising to find gravity values so similar within (for example) Group 3 across bodies of rather diverse masses and compositions. And while surface gravity is not a primary parameter for describing a planetary body, it is still a physical variable, and as such, it is legitimate to study its distribution, at least formally.

For instance, one can ask a question: is that clustering real, or is it merely a visual appearance? And how likely is such grouping to arise by a pure chance, under the assumption of **Log(g)** inherent continuity?

Two statistical methods were implemented to address those questions, described in the subsequent sections.

Experiment Procedure

1. Introduce quantitative measure of clustering quality
2. Apply that measure to real objects being studied (e.g., Solar System gravities)
3. Generate large number of random "test" values on similar range

[1] For bodies with large difference between polar and equatorial gravities, average is used. This simplification, while not physically sound, does not seem to affect the results significantly.

4. Measure their clustering quality
5. Count occurrences where “test” clustering quality is better than that of the real objects
6. The rate of those occurrences quantifies the probability of the real clustering being a result of a random fluctuation

1. Introduce quantitative measure of clustering quality

Multiple and often incompatible methods for assessing clustering quality are known (e.g., [140, 160]) across data analysis industry. That is reflective of the fact that the very definition of “quality” is often subjective and domain-dependent, leaving no universal method to apply well across all contexts.

In the spirit of that understanding, it often makes sense to develop measures tailored for specific problems, keeping in mind industry practices as a reference.

In this case, two measures were introduced, further referred as Gaps Area (GA) and Blur Tolerance (BT). The expectation was that the use of two independent approaches would help validate consistency and correctness of the result.

Gaps Area GA

This measure, which could be viewed as a variation of Davies–Bouldin index [140], is defined as the average size of inter-cluster areas relative to the total span of values considered:

$$GA = \frac{\sum_{c=1}^{K} \left(\min\limits_{g \in Cluster_c} g - \max\limits_{g \in Cluster_{c-1}} g \right) / (K-1)}{\max g - \min g}$$

Here **c** enumerates clusters in the order of increasing gravity, and **K** is the total number of them. The definition is scale-invariant so there is no dependency on the choice of the modeling range.

To be applied, it requires that objects are clustered first using one of the standard methods. In this case, hierarchical agglomerative clustering with minimal (or “single-linkage”) merging [40, 50] criteria was chosen for its simplicity and ease of results interpretation, especially in one-dimensional case. The algorithm performs as following:

1. Initially, each element (i.e., each gravity value) is a cluster of its own.
2. For each pair of clusters **A** and **B**, distance **D**(**A**, **B**) is calculated. **D**(**A**, **B**) is defined here as the smallest distance between any element in **A** and any element in **B**. In one-dimensional case, this is simply **Min**(**B**) – **Max**(**A**) if **B** is to the right of **A**.
3. The pair of clusters with the least mutual distance is merged into a new cluster.
4. Steps 2-3 are repeated until the desired number of clusters **K** is reached.

Defined that way, clustering is “good” when objects sit in a few closely packed groups with large inter-cluster gaps:

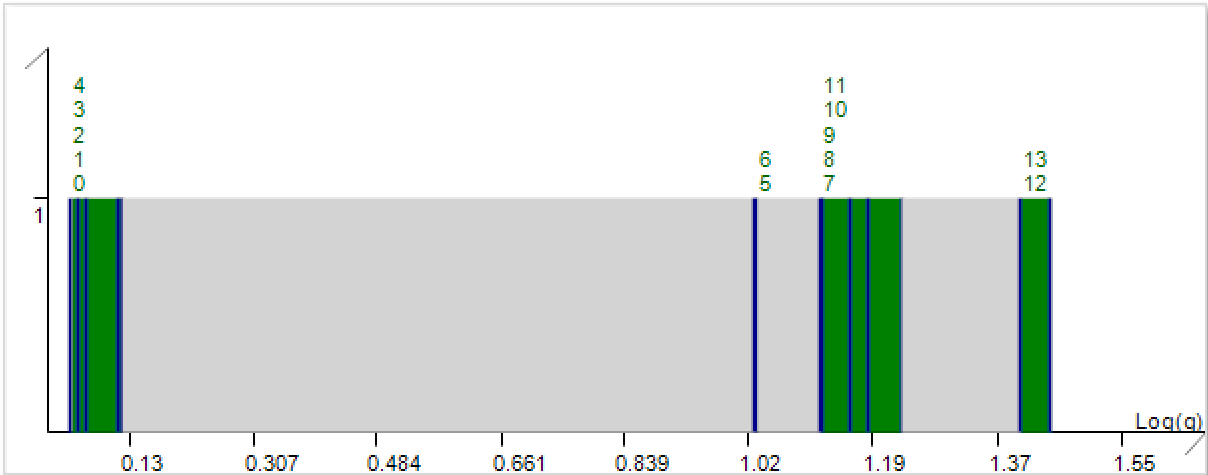


Figure 2. "Good" objects clustering (large inter-cluster gray gaps, tight green clustered groups)

The opposite of that are nearly-uniformly spaced values that produce relatively little gray area and thus lower clustering scores:

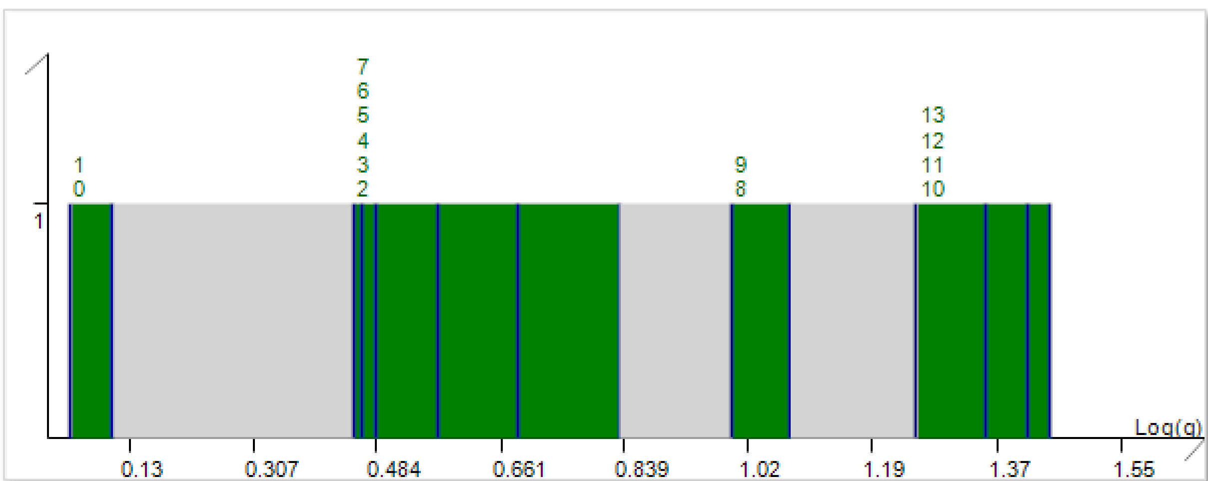


Figure 3. "Poor" clustering quality: small gray area, large and loose green groups.

While simple, this definition becomes degenerative in the presence of multiple single-element clusters, so an alternative definition is needed to augment it.

Blur Tolerance measure BT

This approach builds on the ideas behind Kernel density estimation [50, 150] and assesses how discernable clusters are under poor observations/measurement conditions.

Starting with a set of gravity values $\mathbf{g}_i$, it converts them to a continuous function $\mathbf{F}(\mathbf{g})$ via the blur transformation:

$$F(g;r) = \sum_i exp(-(Log(g) - Log(g_i))^2/r^2)$$

Intuitively, that mimics image quality loss caused by imperfect observation. The tighter and more pronounced the clusters are, the more blurring/smoothing their "image" can sustain while preserving enough contrast to distinguish the clusters:

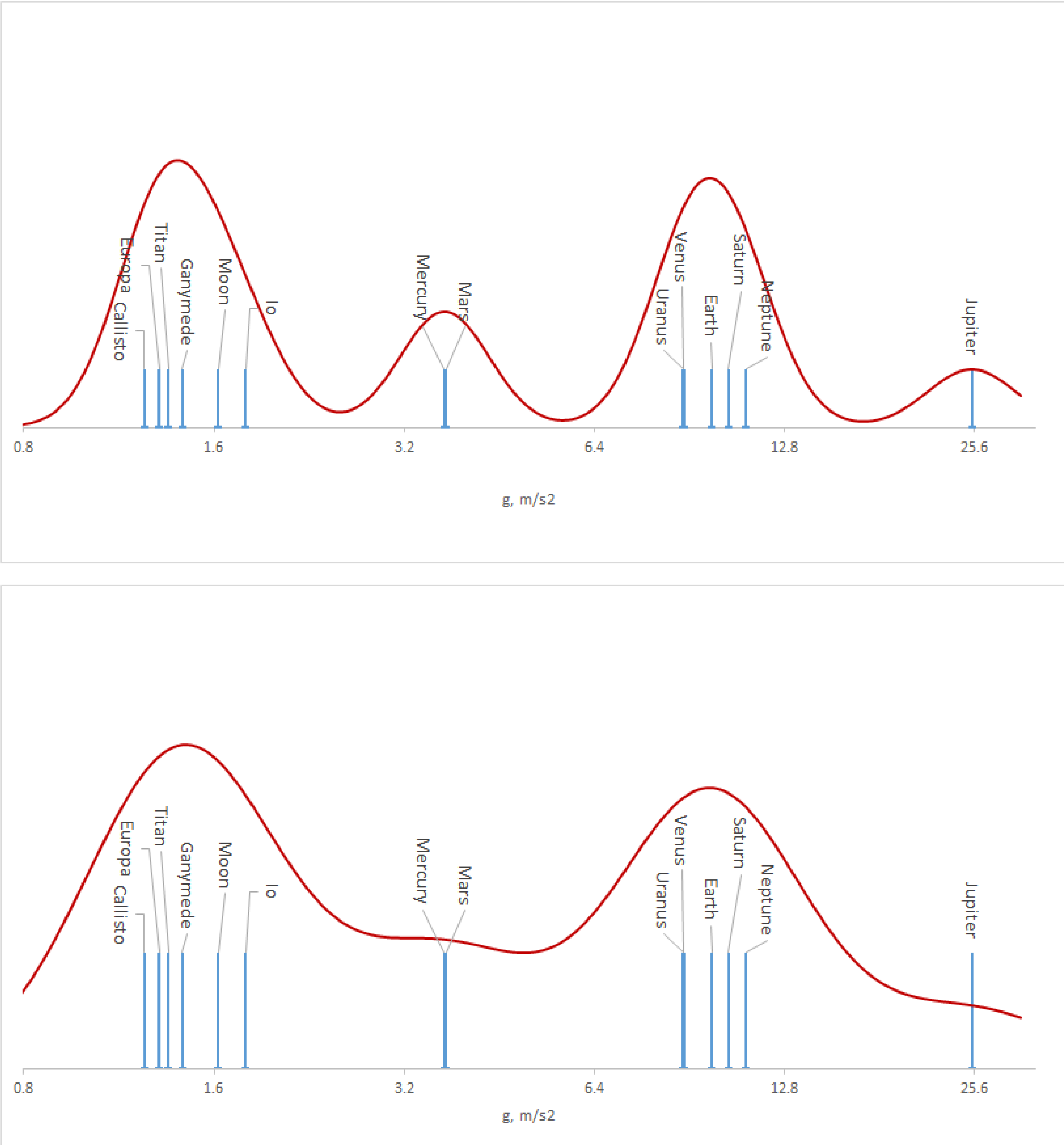


*Figure 4. **F(g)** (red curve) for blur radius **r** = 0.1 (above) and **r** = 0.2 (below) for Solar System gravities, arbitrary units. For **r** = 0.1 (above), four clusters are distinguishable. At increased blur **r** = 0.2 (below), only two groups remain detectable over the imposed minimum contrast ratio of 2:1.*

The blur radius **r** (relative to the total gravities span) that preserves at least 2:1 contrast between all **K** clusters in the group is defined as Blur Tolerance clustering quality measure (**BT**):

$$BT = \frac{r}{\max g - \min g}$$

This definition is also scale-invariant.

This metric produces only clustering quality, but not the actual clusters; however, those could be easily recognized visually by looking at the **F(g)** function.

GA vs BT?

While these two measures often reasonably agree, sometimes their assessments of clustering quality could conflict, resulting in the estimates of unlikelihood differing by orders of magnitude. A good example of that is Kepler-102 system [110].

Per Gap Area metric, Kepler-102 gravities make three strong (**p** = 0.006) clusters:

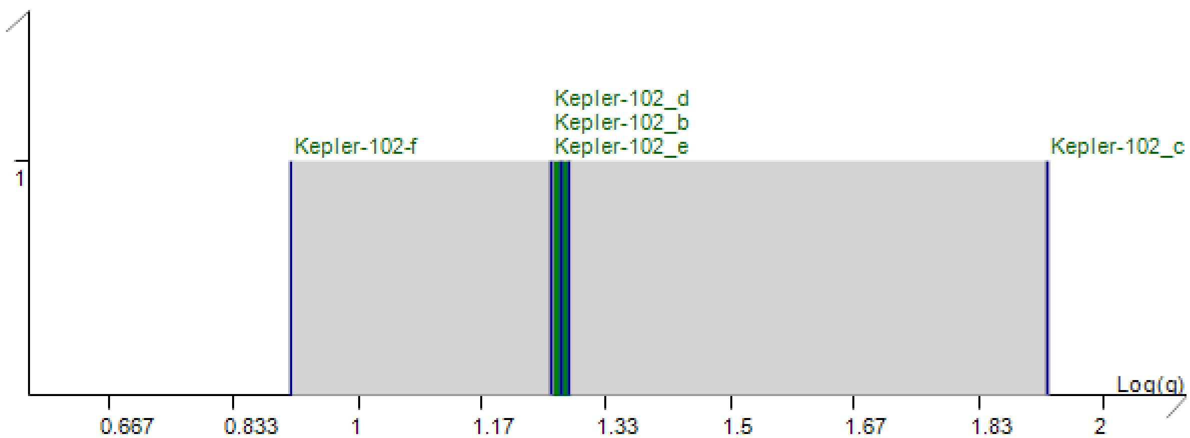


Figure 2. Gaps Area clustering of Kepler-102

Yet Blur Tolerance sees nothing extraordinary here (**p** = 0.34, the “blur” function is shown in green):

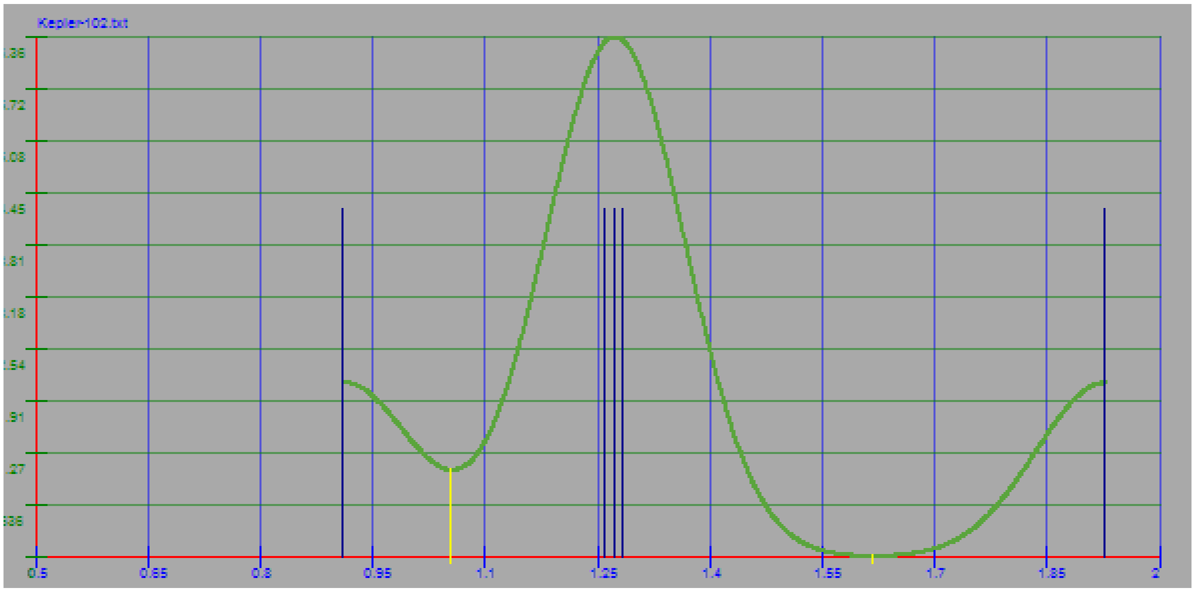


Figure 5.Blur Tolerance clustering of Kepler-102 gravities

How such disagreements should be reconciled?

Gaps Area measure is concerned only with the amount of space left between the clusters. If one keeps making the central cluster on Figure 5 narrower, they would get progressively better clustering scores, even though very little physically changes through such a progression. This mental experiment suggests that Gap Area metric becomes degenerative when cluster size is smaller than the uncertainties of the measurement. It would produce elevated scores in such a scenario, as well as when there are multiple single-element “clusters” in the result.

Blur Tolerance approach is more robust against those issues and its results should be favored in most conflicts, especially when clusters are tight or singular. At the same time, it also has a bias of its own, whereas it slightly favors periodically spaced clusters over non-periodic ones.

Overall, Blur Tolerance was found to be more robust than Gap Area and is treated as the primary scoring method throughout this work. Results obtained with Gap Area are typically discarded when there are two or more singular clusters in the output.

2. Apply clustering quality measures to real objects being studied

Target cluster counts **K** in the range of 2-8 were used to avoid accidentally picking subjectively appealing but unjustified count.

The data sets used were:

Solar System gravities [10, 60]

Primary Set: surface gravities of 14 largest Solar System bodies with **g** > 1 m/s^2. This included Jupiter, Neptune, Saturn, Earth, Uranus, Venus, Mars, Mercury, Io, Moon, Ganymede, Titan, Europa, and Callisto.

No Giants Set. Same as the previous, with giant planets removed (i.e., with no Jupiter, Saturn, Uranus, and Neptune). The set was considered to test whether the transition from rocky to gaseous planets is the reason sufficient to explain the apparent clustering.

Small Bodies. This set consists of 15 small Solar System bodies with 0.1 m/s^2 < **g** < 1 m/s^2 that are **not** expected to demonstrate significant grouping. It contains Triton, Eris, Pluto, Haumea, Titania, Oberon, Ceres, Charon, Ariel, Rhea, Umbriel, Dione, Iapetus, Tethys, and Quaoar in it. This set was used to test if the approach detects any clusters where (at least from the standpoint of visual inspection) we do not expect them.

Extrasolar Systems gravities

Four extrasolar systems with five or more planets were run through the pipeline to see whether similar kind of gravities clustering is observed in them. Gravity values were obtained from mass and radii values in [110]. Only transit detections (to have direct radii data) were included. Measurement uncertainties were ignored (which arguably might have led to significant biases). While clustering up to **K** = 4 was produced, most likely only **K** = 2 results are reliable.

The data set included four systems: **Kepler-11**, **Kepler-20, Kepler-62, and Kepler-102**.

Extrasolar High Precision Gravities

This consisted of twenty extrasolar planets across different systems with surface gravities known to within 25% precision (per [110]).

Solar System non-gravity parameters

These additional data sets were used primarily for testing and/or evaluation purposes. All parameters were clustered in the logarithmic space (e.g., Log(**M**), Log(**R**), etc.).

Solar System radii for the same 14 bodies as for gravities analysis. In theory, no significant clustering is expected in this case.

Solar System masses, which similarly should've exposed no clustering either.

Solar System bodies' distances to the Sun for the same 14 bodies as for gravity set. This set might show some weak clustering at higher values of **K** since giant's satellites have the same solar distances as their parent bodies.

Solar System densities for the same 14 bodies as for the gravities set. A weak form of clustering separating gas giants from rocky planets and (possibly) icy satellites may be expected.

3. Generate large number T of random values sets

Each set would contain **N** values randomly drawn from a selected distribution, where **N** is the number of real objects studied (e.g., 14 for Solar System gravities).

To be clear, no physical processes (like planets formation) are modeled here. All "gravities" were just random numbers from a certain fixed distribution.

As a null hypothesis, assume that groups and gaps are random fluctuations

In other words, that no physical process in the Solar System is expected to produce any clustering in gravities distribution.

Of course, even a quick glance at the Solar System suggests this is not the case. Formation processes of rocky planets, giant planets, and their satellites are likely different [30] and completely smooth transition between the properties of those classes is not necessarily expected. By explicitly ignoring these – known and unknown – effects, we seek to quantify their impact on clustering of gravities, via the comparison of grouping produced by the null hypothesis vs. the observations.

Draw test "gravities" from a log-uniform[2] distribution

Obviously, the reality is more complex. Even if all planetary bodies in the Solar Systems had formed via the same process, the resulting gravities distribution probably would not have been that simple.

Just as a consideration, if masses distribution follows the power law of

$$\frac{dN}{dm} \propto m^{-a}$$

Then, at least for smaller objects with negligible pressure-induced compression, the distribution of objects count over **L** = **Log**(**g**) would be exponential:

$$\frac{dN}{dL} \propto e^{3(1-a)L}$$

A log-uniform distribution is a special case of that for **a** = 1.

However, I feared that accounting for such fine details would complicate the model and introduce additional poorly known free parameters, so decided to stay with less precise but more robust log-uniform distribution. After all, most realistic choices of continuous distributions should not result in prominent and multiple clusters. A log-uniform distribution seems like a reasonable first approximation to assert that statement.

4. Measure clustering quality of each test set

That was done using both BT and GA measures.

5. Count cases where test clustering quality is better than that of the real objects

And derive the "likelihood" of the observed clustering, simply as:

$$p = \frac{Count\ of\ clustering\ outcomes\ with\ quality\ no\ worse\ than\ the\ Solar\ System's}{Trials}$$

The statistical error is crudely estimated as

[2] This is the distribution of variable **X** such that Log(**X**) has a uniform distribution on a specified range, with zero probability density outside of that range

$$Err \sim \sqrt{p * (1 - p) * Trials}$$

(Under the expectation that at least for small values, **p** should have roughly binomial distribution)

That probability is calculated for both Gap Area and Blur Tolerance approaches. Whenever a strong disagreement is seen, it is reconciled per the guidelines in the metrics definition section (typically, Blur Tolerance wins).

Results

For each of the set grouped into **K** clusters, the probability to generate the same or better clustering in **T** = 18,000 attempts was calculated for each value of **K**.

Solar System gravities

Table 1. Probabilities of generating random gravities with clustering better than observed in the Solar System

Clusters count K	Primary Set, Blur Tolerance	Primary Set, Gaps Area	No Giants Set, Blur Tolerance	No Giants Set, Gaps Area	Small Bodies, Blur Tolerance	Small Bodies Gaps Area
2	0.16711 ± 0.00278	0.22428 ± 0.00311	0.25006 ± 0.00323	0.12206 ± 0.00244	0.74456 ± 0.00325	0.40622 ± 0.00366
3	0.11606 ± 0.00239	0.04011 ± 0.00146	**0.00511 ± 0.00053**	**0.00589 ± 0.00057**	**0.14150 ± 0.00260**	**0.35383 ± 0.00356**
4	**0.00083 ± 0.00022**	**0.00078 ± 0.00021**	0.99661 ± 0.00043	0.02500 ± 0.00116	0.86261 ± 0.00257	0.51267 ± 0.00373
5	0.99911 ± 0.00022	0.00283 ± 0.00040	0.98756 ± 0.00083	0.06161 ± 0.00179	0.75494 ± 0.00321	0.56861 ± 0.00369
6	0.99906 ± 0.00023	0.01000 ± 0.00074			0.57244 ± 0.00369	0.59261 ± 0.00366
7	0.99044 ± 0.00073	0.01644 ± 0.00095			0.23589 ± 0.00316	0.53389 ± 0.00372

The minimal p-value in each column is marked with bold font. Gray indicates discarded Gaps Area measures (typically due to clusters being too tight or containing single elements).

The immediate interpretation of these results is the following:

1. **K** = 4 indeed appears to be the most natural cluster count for 14 heaviest Solar System bodies gravities, while the case of **K** = 3 cannot be seriously discussed due to metrics disagreement.
2. The probability to generate such clustering for **K** = 4 out of log-uniform gravities distribution by random chance is low (**p** ≈ $8*10^{-4}$)
3. Without giant planets, there are still three significant clusters observed, with the probability of them being a statistical noise of **p** = $(5\text{-}6)*10^{-3}$
4. Gravities clustering in Solar System bodies with 0.1 m/s^2 < **g** < 1 m/s^2 is nearly nonexistent (the best **p** = 0.14 is for **K** = 3)

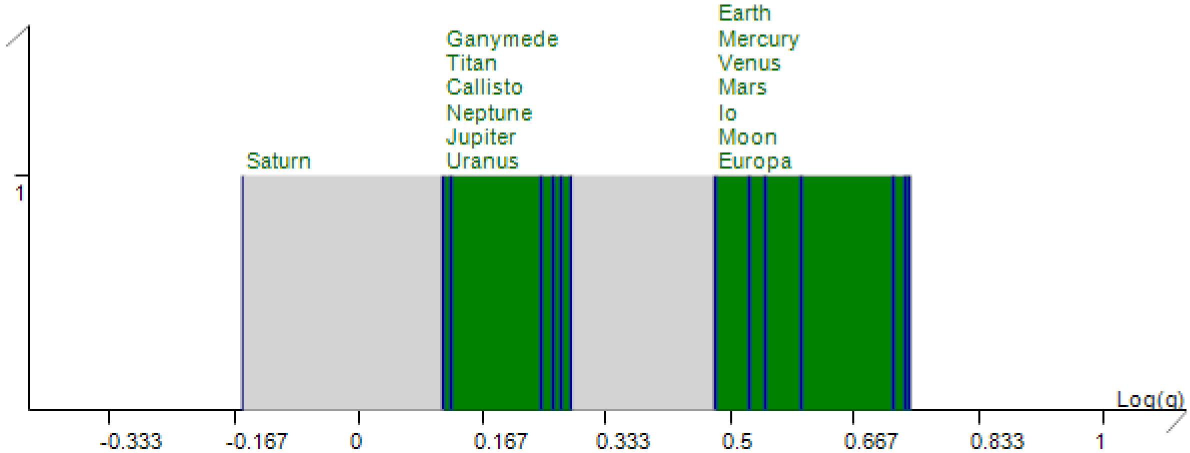
Solar System non-gravity parameters

Table 2. Probabilities of randomly generating parameters that produce clustering better than observed in the Solar System

Clusters count K	Distance to the Sun, BT	Distance to the Sun, GA	Radius, BT	Radius, GA	Mass, BT	Mass, GA	Density, BT	Density, GA
2	0.43472 ± 0.00369	0.24633 ± 0.00321	0.24000 ± 0.00318	0.06611 ± 0.00185	0.64378 ± 0.00357	0.38217 ± 0.00362	0.64606 ± 0.00356	0.17439 ± 0.00283
3	0.86911 ± 0.00251	0.31561 ± 0.00346	0.13439 ± 0.00254	**0.02922 ± 0.00126**	**0.24522 ± 0.00321**	0.31078 ± 0.00345	**0.05289 ± 0.00167**	0.10500 ± 0.00228
4	0.43328 ± 0.00369	0.27333 ± 0.00332	**0.11128 ± 0.00234**	0.01022 ± 0.00075	0.28378 ± 0.00336	0.19900 ± 0.00298	0.06789 ± 0.00187	0.08594 ± 0.00209
5	0.24611 ± 0.00321	**0.13689 ± 0.00256**	0.87300 ± 0.00248	0.01583 ± 0.00093	0.37739 ± 0.00361	0.15367 ± 0.00269	0.05350 ± 0.00168	**0.02633 ± 0.00119**
6	0.17311 ± 0.00282	0.05883 ± 0.00175	0.91200 ± 0.00211	0.01094 ± 0.00078	0.81156 ± 0.00291	0.16550 ± 0.00277	0.96667 ± 0.00134	0.04039 ± 0.00147
7	**0.14183 ± 0.00260**	0.00528 ± 0.00054	0.93611 ± 0.00182	0.01050 ± 0.00076	0.65778 ± 0.00354	**0.15000 ± 0.00266**	0.87217 ± 0.00249	0.04178 ± 0.00149

Immediate interpretation:

1. Weak clustering (**p** = 0.03) is seen for radii separating large planets from the rest of the bodies and satellites
2. No clustering of masses is observed (**p** = 0.245), implicitly supporting the expectation of continuous mass distribution in the Solar System.
3. Three weak density clusters (**p** = 0.05) are found, separating rocky compositions from gaseous/icy ones, and from Saturn:



Overall, these outcomes meet the expectations with respect to Solar System bodies' radii, masses, densities, and solar distances distribution.

Gravities within Extrasolar Planetary Systems

Table 3. Probabilities of generating random gravities with clustering better than observed in reality

Clusters count K	Kepler-11, BT	Kepler-11, GA	Kepler-20, BT	Kepler-20, GA	Kepler-62, BT	Kepler-62, GA	Kepler-102, BT	Kepler-102, GA
2	**0.05400 ± 0.00168**	**0.00833 ± 0.00068**	0.51111 ± 0.00373	0.43317 ± 0.00369	**0.19950 ± 0.00298**	**0.20761 ± 0.00302**	0.42144 ± 0.00368	0.19806 ± 0.00297
3	0.97928 ± 0.00106	0.03167 ± 0.00131	0.38478 ± 0.00363	0.40617 ± 0.00366	0.74483 ± 0.00325	0.31911 ± 0.00347	**0.33267 ± 0.00351**	0.01000 ± 0.00074
4	0.96356 ± 0.00140	0.07778 ± 0.00200	**0.37833 ± 0.00361**	**0.13122 ± 0.00252**	0.53250 ± 0.00372	0.02561 ± 0.00118	0.99428 ± 0.00056	0.13711 ± 0.00256

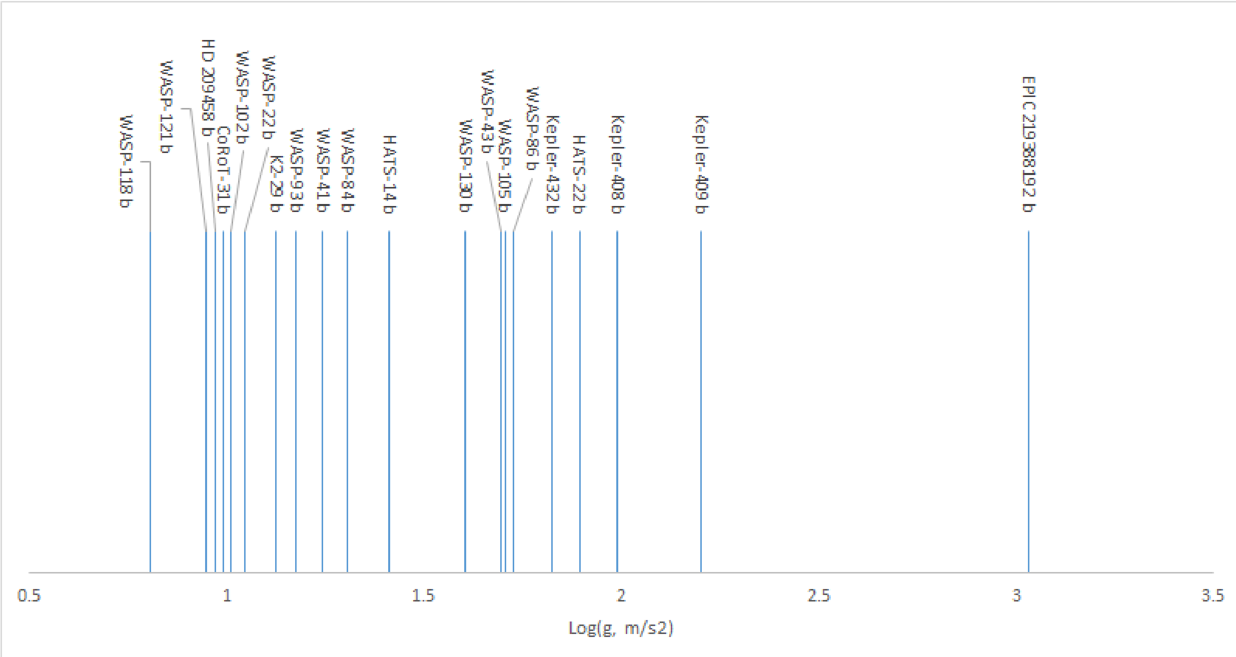
No gravity clustering comparable to that of the Solar System's is seen except for Kepler-11. However, its low p-value is due to Kepler-11 g. The mass of that planet is provided as $0.95^{+0}_{-0.95}$ M_{jup} [110], offering an upper bound estimate only, almost certainly being an overestimate. Clustering analysis has simply flagged that large uncertainty.

Extrasolar Planets with gravities known within 25% precision

Table 4. Probabilities of generating random gravities with clustering better than observed in reality for High Precision Set

Clusters count K	Blur Tolerance	Gaps Area
2	**0.15461 ± 0.00269**	**0.00378 ± 0.00046**
3	0.41833 ± 0.00368	0.01883 ± 0.00101
4	0.91256 ± 0.00211	0.03694 ± 0.00141
5	0.99506 ± 0.00052	0.07039 ± 0.00191
6	0.99756 ± 0.00037	0.14433 ± 0.00262
7	0.98817 ± 0.00081	0.21700 ± 0.00307
8	0.94272 ± 0.00173	0.29217 ± 0.00339

The presence of an obvious outlier EPIC 219388192 b causes strong discrepancy between BT and GA measures, making it difficult to reason about this data set. Visual inspection of **Log**(**g**) distribution suggests that BT figures should probably be favored, with overall conclusion of the lack of significant clustering:



Conclusions & Discussion

The probability to generate same or better gravities clustering than that of the 14 largest Solar System bodies by a pure chance is **p** = $8*10^{-4}$ (under the assumptions and simplifications made).

Neither other physical properties of the Solar System tested, nor exoplanet systems with known gravities have confidently demonstrated clustering comparable to the observed in the Solar System gravities.

Therefore, these gaps in gravities distribution are likely real and as such may deserve an attempt at explanation. What those explanations could be?

The tight grouping around ~10 m/s^2 is probably attributable to Solar System environment. It seems that within the range of the Solar System conditions any rocky body significantly heavier than Earth is likely to accrete sizable quantities of Hydrogen and Helium. That would result in density decrease and some form of a flat "shelf" on gravity vs. mass dependence [20, 25, and 80]; Uranus and Neptune seemingly reside there. Thus, rocky-to-gaseous composition transition is, at least partially, responsible for one cluster in the Solar System gravities.

What if we exclude giant planets from the analysis?

There are still two prominent gaps and three clusters in this case:

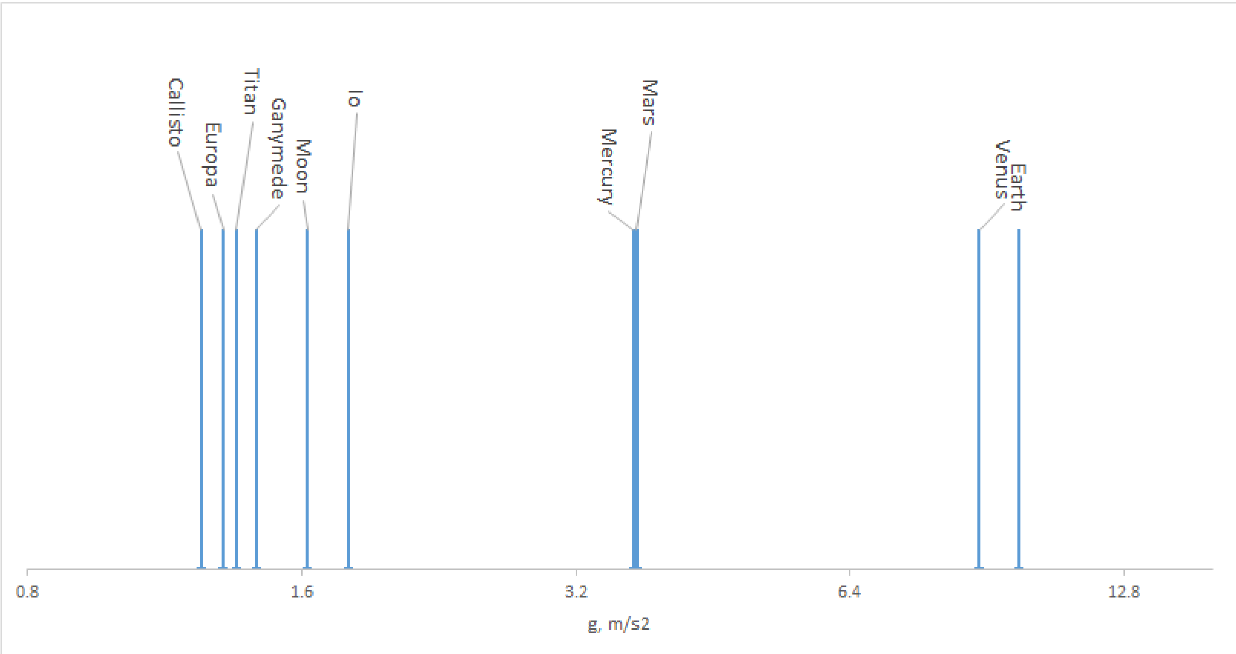


Figure 6. Gravities clustering in Solar System bodies with $g > 0.8$ m/s^2, with giant planets excluded

The same statistical testing suggests that they are possibly real, too. The probability of the opposite is **p** = (5-6)*10^{-3}.

How could that clustering be explained? Relying more on the common sense rather than significant knowledge of Planetary Science, I can put forward the following potential causes, roughly in the order of descending believability:

Stochastic

1. A purely random fluctuation. The world is full of coincidences with 1:200 chances of happening, and this could be just one of them. Under this explanation, Solar System gravities distribution could be outstanding indeed – but without any outstanding reasons behind it.
2. An implicit case of data dredging / p-hacking ([130]). Given sufficiently many parameters, it is always possible to find some that would look unusual by a pre-selected metric. While I did not look for *something* that clusters, a mere act of eying many Solar System parameters could have drawn my attention to what appeared as the most interesting of them.

Assumptions or simplifications that could have affected the results:

1. Drawing test gravities from a log-uniform distribution
2. The choice of clustering & linkage criteria
3. The choice of clustering quality definitions (less likely as two different measures mostly agree on the conclusions)
4. Averaging of polar and equatorial gravities to assign values for rapidly rotating bodies
5. Somewhat arbitrary cutoff choice of **g** > 0.8 m/s^2 for bodies to be considered

Physical causes

1. A manifestation of some relation between planetary density **ρ** and its radius **R**. Surface gravity is proportional to their product **ρR**. The pressure at the center of a body is (neglecting compression)

proportional to **(ρR)**^{2}, or $\mathbf{g^2}$. So it seems that the product **ρR** is an important variable at least in some contexts. If there is a hidden non-trivial relationship between **ρ** and **R**, it could have manifested via clustering of their product distribution.

2. Early Solar System conditions could've prevented formation of bodies with size/mass intermediate between those of Mars and Earth, or between Galilean satellites and Mars. Such bodies would have naturally filled the gaps between the gravity clusters. However, I am not aware of physical processes that would deterministically produce such a separation.
3. Observation bias. Perhaps, bodies with intermediate gravities do exist in the Solar System – but we just have not discovered them yet. This hypothesis does not contradict gravity ranges derived from masses and radius values presented in [70] for putative Planet Nine. However, relatively low statistical confidence of the result (**p** ≈ 0.5%) precludes any serious discussion of this opportunity.

Appendix

For verification purposes, extrasolar planets surface gravities derived from [110] are listed here.

Gravities of Kepler-11 system [110]

Planet	g, m/s2
Kepler-11 b	5.88259646
Kepler-11 c	3.53178836
Kepler-11 d	7.52270685
Kepler-11 e	5.44253482
Kepler-11 f	3.23587523
Kepler-11 g	224.899982

Gravities of Kepler-20 system [110]

Planet	g, m/s2
Kepler-20 b	27.8517062
Kepler-20 c	13.78114129
Kepler-20 d	13.78114129
Kepler-20 e	41.30097452
Kepler-20 f	142.6210825

Gravities of Kepler-62 system [110]

Planet	g, m/s2
Kepler-62 b	51.90503324
Kepler-62 c	138.775
Kepler-62 d	36.87884793
Kepler-62 e	138.2854938
Kepler-62 f	175.8226253

Gravities of Kepler-102 system [110]

Planet	g, m/s2
Kepler-102 b	18.70113379
Kepler-102 c	84.46153846
Kepler-102 d	19.23846154
Kepler-102 e	18.188593
Kepler-102 f	8.132030123

Gravities of 20 extrasolar planets known with 25% or better precision [110]

Planet	g, m/s2	g relative uncertainty
CoRoT-31 b	9.792	0
EPIC 219388192 b	1077.57791	0.184979608
HATS-14 b	25.71540526	0.141480233
HATS-22 b	78.1987406	0.236453945
HD 209458 b	9.391304348	0.10149219
K2-29 b	13.36176824	0.176946836
Kepler-408 b	97.27903922	0.016438911
Kepler-409 b	159.6354575	0.113389123
Kepler-432 b	66.69545779	0.197982202
WASP-102 b	10.20394541	0.19515109
WASP-105 b	50.625	0.236681661
WASP-118 b	6.425	0.178092195
WASP-121 b	8.82325899	0.20375246
WASP-130 b	40.24946345	0.200401153
WASP-22 b	11.12453316	0.150385338
WASP-41 b	17.49841999	0.208415995
WASP-43 b	49.55561187	0.125543177
WASP-84 b	20.2718163	0.174344049
WASP-86 b	53.27751963	0.221707977
WASP-93 b	14.93972143	0.233493476

The data was filtered according to the following criteria:

- publication_status == “Published in a refereed paper” | “Announced on a professional conference”;
- planet_status == “Confirmed”;
- detection_type == “Imaging” | ”Primary Transit”;
- mass_detection_type == “Radial Velocity” | ”Spectrum”;
- radius_detection_type == “Primary Transit” | “Flux”;
- mass != “”;
- radius != “”;

- $(g_{max} - g_{min})/g < 0.25$, where $g_{max} = G*(m + mass\_error\_max)/(r - radius\_error\_min)^2$, $g_{min} = G*(m - mass\_error\_min)/(r + radius\_error\_max)^2$.

References

[10] Katharina Lodders, Bruce Fegley, Jr. *The Planetary Scientist's Companion*. New York, Oxford, Oxford University Press, 1998.

[20] Diana Valencia, Dimitar D. Sasselov, Richard J. O'Connell. *DETAILED MODELS OF SUPER-EARTHS: HOW WELL CAN WE INFER BULK PROPERTIES?* The Astrophysical Journal, 665:1413–1420, 2007 August 20, http://iopscience.iop.org/article/10.1086/519554

[25] Leslie A. Rogers. *MOST 1.6 EARTH-RADIUS PLANETS ARE NOT ROCKY*. Accepted to ApJ, in press as of 05/12/14, http://arxiv.org/abs/1407.4457

[30] George H. A. Cole, Michael M. Woolfson. *Planetary Science (subtitle: The Science of Planets around Stars).* IOP Publishing Ltd 2002.

[40] https://en.wikipedia.org/wiki/Hierarchical_clustering#Linkage_criteria

[50] Jiawei Han, Micheline Kamber, Jian Pei. *Data Mining Concepts and Techniques, Third Edition.* Morgan Kaufman Publishers, 2012.

[60] https://en.wikipedia.org/wiki/List_of_Solar_System_objects_by_size#Larger_than_400_km (essentially a public data)

[70] Jonathan J. Fortney, Mark S. Marley, Gregory Laughlin, Nadine Nettelmann, Caroline V. Morley, Roxana E. Lupus, Channon Visscher, Pavle Jeremic, Wade G. Khadder, Mason Hargrave. *THE HUNT FOR PLANET NINE: ATMOSPHERE, SPECTRA, EVOLUTION, AND DETECTABILITY.* https://arxiv.org/abs/1604.07424

[80] Fernando J. Ballesteros, B. Luque *Walking on exoplanets: Is Star Wars right?* https://arxiv.org/abs/1604.07725

[110] The Extrasolar Planets Encyclopaedia, http://exoplanet.eu/

[130] Data dredging / p-hacking in Wikipedia: https://en.wikipedia.org/wiki/Data_dredging

[140] Some well-known measures of clustering quality on Wikipedia: https://en.wikipedia.org/wiki/Cluster_analysis#Evaluation_and_assessment

[150] Kernel density estimation overview: https://en.wikipedia.org/wiki/Kernel_density_estimation

[160] Yanchi Liu, Zhongmou Li, Hui Xiong, Xuedong Gao, Junjie Wu. *Understanding of Internal Clustering Validation Measures*, 2010 IEEE International Conference on Data Mining