

Remote life-detection criteria, habitable zone boundaries, and the frequency of Earth-like planets around M and late K stars

James F. Kasting¹, Ravikumar Kopparapu, Ramses M. Ramirez, and Chester E. Harman

Department of Geosciences, Pennsylvania State University, University Park, PA 16802

Edited by Adam S. Burrows, Princeton University, Princeton, NJ, and accepted by the Editorial Board October 31, 2013 (received for review May 13, 2013)

The habitable zone (HZ) around a star is typically defined as the region where a rocky planet can maintain liquid water on its surface. That definition is appropriate, because this allows for the possibility that carbon-based, photosynthetic life exists on the planet in sufficient abundance to modify the planet's atmosphere in a way that might be remotely detected. Exactly what conditions are needed, however, to maintain liquid water remains a topic for debate. In the past, modelers have restricted themselves to water-rich planets with CO₂ and H₂O as the only important greenhouse gases. More recently, some researchers have suggested broadening the definition to include arid, "Dune" planets on the inner edge and planets with captured H₂ atmospheres on the outer edge, thereby greatly increasing the HZ width. Such planets could exist, but we demonstrate that an inner edge limit of 0.59 AU or less is physically unrealistic. We further argue that conservative HZ definitions should be used for designing future space-based telescopes, but that optimistic definitions may be useful in interpreting the data from such missions. In terms of effective solar flux, S_{eff} , the recently recalculated HZ boundaries are: recent Venus—1.78; runaway greenhouse—1.04; moist greenhouse—1.01; maximum greenhouse—0.35; and early Mars—0.32. Based on a combination of different HZ definitions, the frequency of potentially Earth-like planets around late K and M stars observed by Kepler is in the range of 0.4–0.5.

As the appearance of this special issue of PNAS confirms, the search for exoplanets is by now well underway; indeed, it is one of the hottest research areas in all of astronomy. At the time of this writing, 428 exoplanets have been identified by ground-based radial velocity (RV) measurements and 279 planets have been discovered by the transit technique (1) and confirmed by various other methods, including RV. In addition, more than 3,000 "planet candidates" have been identified by the Kepler mission (2). Most of these planets are either too large or too close to their parent star to have any chance of harboring life. However, both astronomers and the general public are interested in identifying potentially habitable planets and in searching their atmospheres spectroscopically for evidence of life. This leads immediately to the question of how life can be recognized remotely, along with the related question of what are the conditions needed to support it. Liquid water is often mentioned as a prerequisite for life, and we will argue below that this restriction is appropriate for the astronomical search for life. However, some researchers have questioned this assumption (3), and so we begin by explaining why the presence of liquid water is so important in the search for life on planets around other stars.

Criteria for Remote Life Detection: Why Liquid Water Matters

Biologists sometimes debate what constitutes life, but one definition that most of them accept is the following (4, 5): "Life is a self-sustained chemical system capable of undergoing Darwinian evolution." That definition appeals to biologists because it is general and because it can be tested in the laboratory. Astronomers, however, are interested in looking for life on planets

around other stars by performing remote sensing of the planets' atmospheres, so to them the biologists' definition of life is not particularly useful. Instead, what they need is a way to recognize life from a great distance. It was realized many years ago that the best way to do this is by looking for the byproducts of metabolism. As early as 1965, Lederberg (6) suggested that the best remote signature of life was evidence for extreme thermodynamic disequilibrium in a planet's atmosphere (but see criticism of that idea below). In that same year, Lovelock (7) put forward a more specific remote signature, namely, the simultaneous presence of O₂ and a reduced gas such as CH₄ or N₂O in a planet's atmosphere. On Earth, all three of these gases are produced mostly by organisms, and they are many orders of magnitude out of equilibrium with each other, so the fact that they coexist is indeed strong evidence that our planet is inhabited.

This life-detection criterion, however, is much less general than is often supposed, as demonstrated by this counterexample. Suppose that a planet's atmosphere was found to have high concentrations of both CO and H₂. These gases are quite far out of thermodynamic equilibrium at room temperature, as free energy considerations strongly favor the reaction: CO + 3 H₂ → CH₄ + H₂O. Thus, a naive interpretation might be that the observed disequilibrium was evidence for life. However, CO-rich, CH₄-poor atmospheres can be produced by impacts (8) or by photolysis of CO₂ in cold, dry low-O₂ atmospheres (9). The presence of life would likely destabilize such a CO-rich atmosphere, as either methanogens would consume it (10), or alternatively acetogens would use it to produce acetate (11). So, the criterion of extreme thermodynamic equilibrium as a biomarker is directly contradicted.

Significance

National Aeronautics and Space Administration's (NASA) Kepler Space Telescope has detected over 3,000 planet candidates, about a dozen of which are probably rocky planets within the liquid-water habitable zones of their parent stars. Climate-modeling calculations discussed here shed light on the width of that zone. Within the next several years, NASA may obtain spectra of nearby transiting Earth-sized planets around M stars, using its James Webb Space Telescope. NASA hopes to build an even more capable space telescope to perform direct imaging of Earth-like exoplanets and take spectra of their atmospheres. Once data are obtained from either of these missions, correct interpretation of possible biomarker gases will become critical. We discuss here how those interpretations might be made.

Author contributions: J.F.K. designed research; R.K. and R.M.R. performed research; C.E.H. analyzed data; and J.F.K. and R.K. wrote the paper.

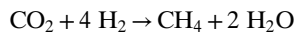
The authors declare no conflict of interest.

This article is a PNAS Direct Submission. A.S.B. is a guest editor invited by the Editorial Board.

¹To whom correspondence should be addressed. E-mail: jfk4@psu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1309107110/-DCSupplemental.

It is also easy to demonstrate that the converse of this statement is false: the absence of extreme thermodynamic disequilibrium should not be taken as evidence that life does not exist on the planet. To illustrate why this is so, consider the Earth's atmosphere and biosphere before the origin of oxygenic photosynthesis. Models suggest that organisms in such an anaerobic biosphere would combine CO₂ and H₂ to make CH₄, either directly (by methanogenesis) or indirectly (by anoxygenic photosynthesis, followed by fermentation and methanogenesis) (11). In either case, the early biosphere would have been driving the atmosphere toward thermodynamic equilibrium, not away from it, because the reaction



is energetically downhill at room temperature.

Recently, Seager et al. (12, 13) have proposed a systematic classification of potential biomarker gases, which is useful. They divide them into three categories: type I biomarkers are redox gradient byproducts, such as CH₄ in the case just described. They are produced when organisms take advantage of existing thermodynamic free energy gradients to power their metabolisms. These biomarkers are inherently equivocal, because abiotic processes might also synthesize them, provided that suitable kinetic pathways exist. CH₄, for example, can be formed by serpentinization of ultramafic rocks by water containing dissolved CO₂ (14). Similarly, Seager et al. suggest that NH₃ might be a type I biomarker on a cold "Haber world" with a dense N₂-H₂ atmosphere, because it is difficult to produce NH₃ by gas-phase photochemistry. However, NH₃ can also be produced by photolytically catalyzed reaction of N₂ with H₂O on surfaces containing TiO₂ (15), and NH₃ might be synthesized from dissolved N₂ and H₂ within submarine hydrothermal systems on a planet with a highly reduced mantle, like Mars (16). Type II biomarkers are biomass-building byproducts, such as the O₂ produced from oxygenic photosynthesis. However, this could conceivably be the only useful type II biomarker, as anoxygenic photosynthesis based on H₂S or Fe²⁺ yields the nongaseous byproducts, elemental sulfur and Fe₂O₃, respectively; and other forms of anoxygenic photosynthesis yield highly soluble byproducts such as sulfate and nitrate, or H₂O itself (12). Type III biomarkers are secondary metabolic products, such as DMS (dimethyl sulfide), OCS, CS₂, CH₃Cl, and higher hydrocarbons such as isoprene. These gases, however, are produced in sufficiently small amounts and/or are photolyzed sufficiently rapidly that they are not expected to build up to observable concentrations in a planet's atmosphere. So, the number of potentially detectable biomarker molecules may be relatively limited.

Although the Seager et al. scheme is a nice way of classifying biomarker gases, it is not clear that it really changes the fundamental nature of remote life detection. One still needs to observe something in a planet's atmosphere that cannot be explained by abiotic processes. It is significant to note that nearly all of the hypothetical planets that these authors discuss possess liquid water. They do discuss Titan, but they argue that it requires unphysically high biomass amounts based on terrestrial analogies. We would go further than this. Searching for life on our own solar system's Titan is a fascinating endeavor, because perhaps we will be surprised and find that life exists there. However, remote sensing of Titan-like planets around other stars is not likely to convince many biologists that life exists there, unless life has already been discovered on Titan itself.

One final thought about liquid water planets should be noted. Many authors have suggested that life could conceivably be present on a planet with a subsurface liquid water reservoir. These conjectures can eventually be tested by in situ measurements on bodies within our solar system, such as Mars or Jupiter's moon Europa. However, for exoplanets, in situ sampling remains

indefinitely out of reach, and the ability of a subsurface biosphere to modify a planet's atmosphere in a detectable way remains to be demonstrated. Thus, in looking for signs of life on exoplanets, it makes sense to search for planets like Earth that have abundant liquid water at their surfaces. In other words, the habitable zone (HZ), as conventionally defined, is precisely where we should focus our attention.

Habitable Zone Boundary Definitions

Inner Edge. Two different types of climate catastrophe can be triggered if an Earth-like planet's surface becomes too warm. If the surface temperature exceeds the critical temperature for water (647 K for pure H₂O), the entire ocean should evaporate. This is termed a runaway greenhouse. Alternatively, Goldblatt et al. (17) define a runaway greenhouse as an atmosphere in which the absorbed solar flux exceeds the outgoing thermal-IR flux. However, these two definitions are equivalent, in practice, because an atmosphere that is out of flux balance in this way "runs away" to supercritical surface temperatures. A runaway greenhouse could occur at a surface temperature less than 647 K on a planet that had less water than Earth; the critical temperature applies on Earth because the pressure of a fully vaporized ocean (270 bar) exceeds the critical pressure for water (220.6 bar). Runaway greenhouse atmospheres have finite lifetimes, because once the upper atmosphere becomes water-rich, the water can be lost in a few tens of millions of years by photodissociation followed by escape of hydrogen to space (18, 19).

However, a planet can also lose its water to space at much lower temperatures. Once the surface temperature exceeds ~340 K for a 1-bar Earth-like atmosphere, the H₂O saturation mixing ratio at the surface exceeds 0.2. This leads to a large increase in tropopause height and an accompanying large increase in stratospheric H₂O (18, 19). The presence of gaseous H₂O above the altitude at which it can condense then overcomes the diffusion limit on hydrogen escape (20), so the water can be rapidly photodissociated and lost. This is termed the moist greenhouse limit. Finally, a third, empirical limit can be defined based on the observation that Venus has no liquid water, even though its initial water endowment may have been high, based on the planet's high D/H ratio (21). Indeed, Venus had apparently already lost any surface water it might have possessed by 1 billion years ago (bya) (the approximate age of its most recent resurfacing; ref. 22), at which time the Sun was ~8% less bright than today (23). This third estimate is termed the recent Venus limit. It implies that the inner edge of the HZ is at least 4% farther out than Venus' orbital distance of 0.72 AU.

According to Kasting et al. (23), the three estimates for the effective solar flux, S_{eff} , at the HZ inner edge for a G2 star like the Sun were: moist greenhouse, or "water loss"—1.10; runaway greenhouse—1.41; and recent Venus—1.76. In terms of orbital distance, these correspond to values of 0.95, 0.84, and 0.75 AU, respectively, for the present solar luminosity. Note that the HZ boundaries are not defined in terms of the planet's equilibrium temperature, T_{eq} , even though that concept is sometimes erroneously applied in the astrophysics literature. To calculate an equilibrium temperature, one must have knowledge of a planet's albedo. The value of 0.31 for present Earth (24), however, is not applicable to a planet near the inner or outer edge of the HZ. In the model of ref. 23, planets near the moist greenhouse limit were predicted to have albedos of ~0.2, and planets near the maximum greenhouse limit had predicted albedos of ~0.4. So, instead, the preferred way to define the HZ boundaries is in terms of the incident stellar flux at the top of the planet's atmosphere. These boundaries change for stars of different spectral types, because the calculated albedos of moist greenhouse and maximum greenhouse planets become lower (higher) as the star's radiation is shifted toward the red (blue). The change in albedo is caused partly by Rayleigh scattering, which increases

strongly at shorter wavelengths, and partly to the increase in near-IR absorption by H₂O and CO₂ as the star's spectral peak shifts to these wavelengths.

Outer Edge. Following ref. 23, Earth-like planets are assumed to be volcanically active and have substantial supplies of carbon in the form of CO₂ and carbonate rocks. CO₂ is removed from the atmosphere by silicate weathering, followed by deposition of carbonate sediments on the seafloor. This process depends on the amount of rainfall, which slows as the climate cools (25); hence, atmospheric CO₂ should build up on planets that receive less stellar insolation than does Earth. At some point, however, CO₂ begins to condense out of the atmosphere, as it does today on Mars. This led, historically, to two theoretical estimates for the HZ outer edge. The first condensation limit is where the process of CO₂ condensation first begins. That limit, however, is no longer considered valid, because CO₂ ice clouds are now thought to generally warm a planet's surface (26, 27), and so we ignore it from this point on. The still-accepted estimate for the location of the outer edge is termed the maximum greenhouse limit (but see arguments below for how this distance might be extended). As its name implies, this is the distance at which the warming by CO₂ maximizes, so the solar flux required to sustain a 273-K surface temperature is at a minimum. CO₂ partial pressures above the value at this limit lead to surface cooling, because the increased albedo caused by Rayleigh scattering from the CO₂ outstrips the increase in the CO₂ greenhouse effect. Finally, an empirical outer edge for the HZ can be defined based on the observation that Mars appears to have had liquid water flowing on its surface at or before 3.8 bya, when the solar flux was some 25% lower (23). Some authors have argued that this does not imply a long-lived warm climate, because the fluvial features can be produced in other ways, and we return to this question below. For now, we simply assume that early Mars was within the HZ. According to ref. 23, the three estimates for S_{eff} at the HZ outer edge around a G star were: first CO₂ condensation—0.53; maximum greenhouse—0.36; and early Mars—0.32. The corresponding orbital distances were 1.37, 1.67, and 1.77 AU, respectively, for present solar luminosity.

Other authors have proposed modifications to these HZ boundaries. Selsis et al. (28) included clouds in the model of ref. 23 and calculated boundaries for 50 and 100% cloud cover. This procedure is somewhat questionable, as the model of ref. 23 already included clouds implicitly as part of the surface albedo. Not surprisingly, for 100% cloud cover, the inner edge for a Sun-like star moves in to ~0.5 AU, or $S_{eff} = 4$. However, this result is unrealistic, because condensation clouds are not expected to have such high fractional cloud cover and because it is in direct conflict with the existence of a desiccated Venus at 0.72 AU ($S_{eff} = 1.9$). The 50% cloud cover model produces more believable results, as it predicts an inner edge at 0.76 AU—very close to the recent Venus limit. However, the cloud layer in this model was placed at an arbitrary height, with an arbitrary optical depth, and the influence of the cloud on the planet's greenhouse effect was neglected. So, it is not obvious that this estimate represents any real improvement over the empirical recent Venus limit.

More recently, other authors have proposed modifications to the HZ boundaries based on suggestions of different types of habitable planets and/or the presence of additional greenhouse gases (see ref. 29 for a recent review). Specifically, Zsom et al. (30) have used 1D climate model calculations to argue that the inner edge of the HZ could be as far in as 0.5 AU around a Sun-like star for planets with low relative humidities (RH). For example, their nominal, low-RH planet has a constant tropospheric RH of 0.01. By comparison, Earth's RH varies from ~0.8 at the surface to ~0.1 in the upper troposphere (31). Such a low-RH planet could, in principle, avoid the strong water vapor feedback that makes the runaway greenhouse happen. However, this model

fails to consider energy balance at the planet's surface. The net absorbed (solar + thermal-IR) flux at the surface must equal the convective (latent + sensible) heat flux (32). Because the latent heat flux is a strong function of RH, it can be readily shown that liquid water would be unstable on the surface of a warm low-RH planet if the surface is all at the same temperature. A detailed explanation is given in the *SI Appendix*. Zsom et al. argue that there would be cold parts of the planet where water would condense, but then their 1D model is not really appropriate. In consequence, their predictions about the location of the HZ inner edge are suspect.

Climate on low-RH planets has been studied previously using 3D models. Abe et al. (33) used such a model to show that hot, rocky planets with small water endowments and low obliquities could conceivably remain habitable in their polar regions. Such planets would resemble the planet Dune in Frank Herbert's famous science fiction novel by that title; hence, the name "Dune-like" planet has stuck. Abe et al. placed the inner edge of the HZ at $S_{eff} = 1.7$, well inside the runaway or moist greenhouse limit for a water-rich planet, but still outside the recent Venus limit. Such planets would be consistent with surface energy balance, but they might be unrealistic for another reason. Some of the moisture originating from polar oases on Dune-like planets would presumably fall as rain on drier portions of the planet's surface. Once there, it might well become incorporated into minerals, e.g., clays, as water of hydration. More water could conceivably be tied up in seafloor basalts, as happens on Earth today during circulation of seawater through midocean ridges. Earth's seafloor is eventually subducted, so this water is carried down toward the mantle. This process would have removed about half an ocean of water in 2.5 billion years, were water not being continually returned to the surface by outgassing (34). A Dune-like planet would presumably have a much smaller water inventory, and so the removal time might be considerably shorter. In any case, even if Dune-like planets do exist, the inner edge of the HZ is still well outside our empirical recent Venus limit.

The outer edge of the HZ could conceivably be farther out than the estimates mentioned above as a consequence of warming by CO₂ ice clouds (26) or by additional greenhouse gases. CO₂ ice clouds apparently have only a modest warming effect on climate, based on new 3D studies of early Mars (35). However, the addition of significant concentrations of molecular hydrogen, H₂, can extend the outer edge of the HZ dramatically (36, 37). We have suggested that early Mars could have been warmed by this mechanism (38). H₂ is perfect for this task, for two reasons: (i) Its (collision induced) absorption extends over the entire thermal-infrared spectrum (39), and (ii) it condenses only at extremely low temperatures. Realizing this, Pierrehumbert and Gaidos (37) showed that a 3-Earth-mass planet with a 40-bar captured H₂ atmosphere could conceivably remain habitable out to 10 AU around a Sun-like star. Stevenson (36) had previously demonstrated that such planets could remain habitable even if they were not attached to stars at all, provided that they had sufficient internal heat. Although such planets might exist, we should think carefully about whether this should influence the design of a future habitable-planet-seeking telescope. The contrast ratio between the Earth and the Sun is ~10¹⁰ in the visible (40), which is already considered difficult to achieve. A planet of the same size and albedo located at 10 AU would have a contrast ratio 100 times larger. Would we really want to spend the extra money required to develop a starlight suppression system with this efficiency? Our resources might be better spent ensuring that the telescope can find planets more similar to Earth.

This raises a more general question: Should estimates of HZ width be conservative or optimistic? In our view, both types of estimates are useful, but for different purposes. In designing a space telescope like National Aeronautics and Space Administration's (NASA) proposed Terrestrial Planet Finder (TPF) (also

known as New Worlds Observer), one of the key questions is how big should it be. A bigger telescope should be more capable, but it is also likely to cost more and may take longer to be approved and built. TPF telescope size is dictated primarily by two factors: (i) One needs to be able to observe enough target stars to have a good chance of finding Earth-like planets. The TPF-C Science and Technology Definition Team report (40) suggested that the planet expectation value should be 3, implying that the chances of finding at least one Earth candidate is >95%. (ii) One needs to be able to distinguish the light from the planet from that of the surrounding exozodi dust. Exozodi levels are not yet well constrained, but using our own zodi background as a proxy suggests that the minimum telescope diameter is 4 m (41).

For constraint 1 above, a key parameter is the value of η_{Earth} — the fraction of Sun-like stars that have at least one rocky planet within their HZ. If η_{Earth} is high, then fewer stars need to be searched. The target stars should therefore be closer, and so the angular separation, θ , between them and their potentially habitable planets should be larger. The theoretical diffraction limit is given by: $\theta \simeq \lambda/D$, where λ is the wavelength at which one is observing and D is the telescope diameter. Real coronagraphs operate only outside of some multiple of this angle; for example, TPF-C was designed with an inner working angle of $4 \lambda/D$ (40). Closer target stars thus translate into a smaller telescope diameter. The TPF-C committee assumed that $\eta_{Earth} = 0.1$; thus, they designed the telescope to be able to search 30 full habitable zones (or, more accurately, 60 half-habitable zones around 60 nearby target stars). To do this, they required an 8×3.5 m elliptical mirror. Recent estimates of η_{Earth} for late K and M stars suggest that η_{Earth} could be significantly higher than this (see discussion below). If Sun-like stars follow this same pattern, then perhaps a 4-m diameter TPF telescope will be sufficient.

Estimating η_{Earth} requires that one use a specific definition of the habitable zone. In doing this, we suggest that one should use a conservative HZ definition, for example, the moist greenhouse limit on the inner edge and the maximum greenhouse limit on the outer edge. That way, the η_{Earth} estimate will also be conservative, and the telescope will be at least as big as it needs to be to find the requisite number of planets. Most importantly, it would be a mistake to take the outer edge at 10 AU for a Sun-like star, because then η_{Earth} might be calculated as unity, and the telescope might be correspondingly undersized. Using a more optimistic inner edge for the HZ is less of a problem, because this would require a smaller inner working angle, which itself

would place constraints on telescope diameter. However, it might turn out that a larger and more expensive telescope would be needed to find Dune-like planets close to their stars, and one would again have to ask whether this additional expenditure would be worthwhile.

Once we have constructed such a telescope and are using it to make observations, then adopting a more liberal HZ definition makes sense. At that point, it would be prudent to examine all rocky planets within the most optimistic HZ, so that one would not accidentally overlook possible Dune planets or H_2 -rich super-Earths in systems where they could be observed. Distant H_2 -rich planets would likely be hard to see, as pointed out above, and their potential biomarkers might be difficult to interpret; however, there is no reason that one could not look for them with TPF, even if it were designed to search for more conventional Earth-like planets.

Revised HZ Boundaries and Calculations of η_{Earth}

Recently, our group has rederived HZ boundaries using a revised and improved 1D climate model. Our new model (42) includes updated absorption coefficients for both CO_2 and H_2O , with the latter being based on line parameters from the relatively new HITEMP database. This database includes significantly more weak H_2O lines than does the older HITRAN database, which is used in most current climate models. This change is particularly important in the visible/near-UV spectral region, where our model predicts much stronger absorption of stellar radiation, thereby lowering planetary albedos and moving the inner edge of the HZ outward.

Fig. 1 shows the new HZ boundaries around stars of different spectral types in terms of effective stellar flux, S_{eff} , which is the way that we recommend that they be implemented in practice. Parametric formulae for these limits are given by equation 2 of ref. 42, with coefficients listed in their table 3. The formulae are expressed in terms of the parameter $T_* = T_{eff} - 5,780$ K, where T_{eff} is the star's effective radiating temperature. For a star like the Sun, with $T_{eff} = 5,780$ K, the new HZ boundaries are given in units of S_{eff} as: recent Venus—1.78; runaway greenhouse—1.04; moist greenhouse—1.01; maximum greenhouse—0.35; and early Mars—0.32. The recent Venus and early Mars limits are the same as in ref. 23, because the planets are still at the same orbital radii and because solar evolution models have not changed significantly since that time. We note that the new theoretical inner edge limits are perilously close to Earth's present orbit; however,

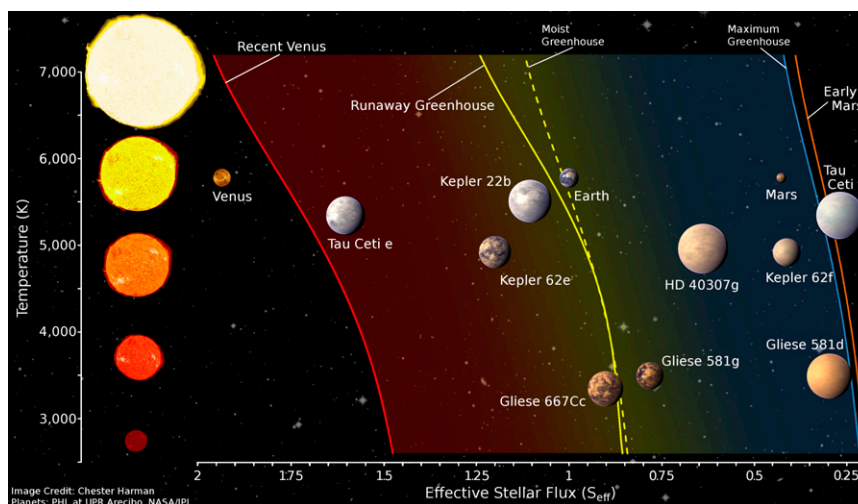


Fig. 1. Diagram showing different HZ boundaries for stars ranging in spectral type from F0 to M7. Various planets within our solar system are shown, along with selected exoplanets.

these limits are almost certainly overly pessimistic and can be improved on by using 3D climate models, as discussed below. Fig. 2 shows the new HZ boundaries in terms of distance. For simplicity, only the moist greenhouse and maximum greenhouse limits are shown. This diagram was calculated for stars at the beginning of their main-sequence lifetimes, and it is only valid at that time, because all main-sequence stars brighten as they age. Fig. 2 shows that on a log distance scale the HZ is slightly broader for late-type stars than for Sun-like stars. The reason is that the planet's albedo depends on the spectrum of its parent star, and this dependence is stronger at the HZ outer edge than at the inner edge because Rayleigh scattering is so important in dense CO₂ atmospheres. So, if planets around late-type stars are distributed geometrically in orbital distance, as they are in our own solar system, one might expect η_{Earth} to be $\sim 25\%$ larger for M stars than for G stars.

Geometrical spacing of planets is predicted by some theories of planetary formation. Dynamical analyses (43–45) have identified a well-defined stability boundary in multiplanet systems. This result led to the packed planetary systems hypothesis (46–48), which proposes that planetary systems form in such a way that they are filled to capacity: adding even one additional planet would create an unstable system. Although this is a plausible hypothesis, it will have to be verified by future observations.

Differences in HZ boundary definitions account for a large part of the differences in recent estimates for η_{Earth} from both Kepler and RV data. Dressing and Charbonneau (49) calculated $\eta_{Earth} = 0.15^{+0.13}_{-0.06}$ for M stars, based on the first six quarters of Kepler data. Their analysis used the moist greenhouse and CO₂ first condensation limits from ref. 23. We pointed out earlier that the latter limit is no longer considered valid. Furthermore, they only included planets up to 1.4 Earth radii (R_{Earth}). However, planets up to $\sim 2 R_{Earth}$ should have masses $< 10 M_{Earth}$ and thus could be rocky, although some planets could have significantly more water ice than does Earth (50). Kopparapu (51) repeated Dressing and Charbonneau's analysis using their same dataset but applying the new moist greenhouse and maximum greenhouse limits from ref. 42. For planets up to 1.4 R_{Earth} in size, he derived $\eta_{Earth} = 0.48^{+0.12}_{-0.24}$, rising to $0.51^{+0.1}_{-0.2}$ when he included planets up to $2 R_{Earth}$. He then repeated the analysis again using the more optimistic recent Venus and early Mars limits and

calculated $\eta_{Earth} = 0.61^{+0.07}_{-0.15}$. More recently, Gaidos (52), using the 50% cloud cover HZ boundaries in ref. 28 and considering planets up to $2 R_{Earth}$, derived an independent estimate of $\eta_{Earth} = 0.46^{+0.18}_{-0.15}$ based on the first eight quarters of Kepler data. These boundaries are roughly comparable to the recent Venus and early Mars limits mentioned above; hence, his estimate for η_{Earth} ought to be higher than Kopparapu's conservative value, but it is not. Gaidos' dataset is 30% longer and includes both M and late K stars, so perhaps this represents a trend toward lower η_{Earth} around more massive stars. Finally, Bonfils et al. (53) derived $\eta_{Earth} = 0.41^{+0.54}_{-0.13}$ based on a 6-y radial velocity study of 102 nearby southern M dwarfs, using the optimistic recent Venus/early Mars HZ limits. Thus, estimates for η_{Earth} for stars significantly smaller than the Sun seem to be converging at values of the order of 0.4–0.5, although the error bars remain large and the different estimates are not strictly comparable because of different assumptions regarding planet size and HZ boundaries. Because the HZ for G stars is 25% narrower than for M stars, η_{Earth} for G stars would be of the order of 0.3–0.35 if other factors, e.g., efficiency of planetary formation and planetary spacing, are equal.

Future Work

Theoretical studies of HZ boundaries and observational determinations of η_{Earth} are still both in their infancy. The most important theoretical question concerns the location of the HZ inner edge. The red area in Fig. 1 shows the large amount of orbital real estate between the conservative moist greenhouse limit and the optimistic recent Venus limit. The actual inner edge probably lies in between these two estimates, but where? This problem can potentially be resolved by more complicated 3D climate modeling. One-dimensional models are incapable of providing reliable answers to this question, for two reasons: (i) the relative humidity distribution is uncertain, and (ii) cloud feedback is impossible to gauge effectively. The moist greenhouse calculation, as performed in refs. 23 and 42, assumes a fully saturated troposphere. This is overly pessimistic, as Earth's present troposphere is significantly undersaturated, on average, in its upper regions. The reasons have to do with cumulus convection, especially in the tropics, where the air remains saturated within convective plumes, but then the relative

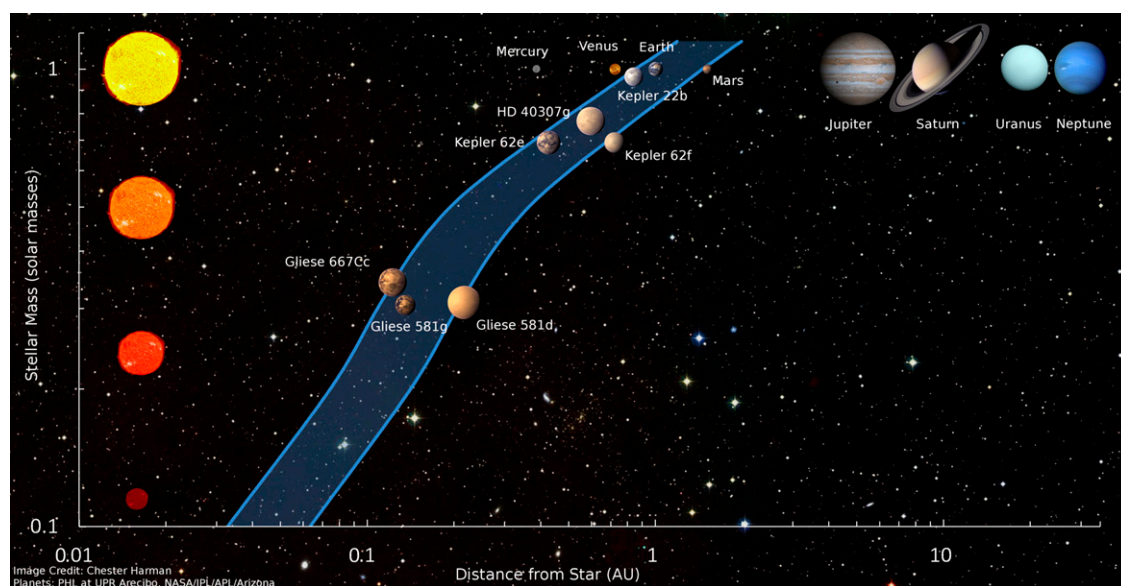


Fig. 2. Diagram showing the HZ boundaries in terms of log distance for zero-age main-sequence stars. The moist greenhouse limit is used for the inner edge and the maximum greenhouse limit for the outer edge. Stellar luminosities and effective temperatures were taken from refs. 56, 57.

humidity drops as the air spreads out at the top of the plume and descends more slowly over wider areas. The dynamics of this process can be captured in 3D models but not in 1D. Similarly, clouds form mostly on updrafts, where the air is cooling, and they are largely absent in regions where the air is subsiding, such as over the Sahara desert. Again, 3D models can capture such dynamics, although even with them it is difficult because most clouds are sub-grid-scale in size. Cloud feedback has been predicted to be negative for warm H₂O-rich atmospheres (19, 28) because their contribution to the planetary albedo is likely to be higher than their contribution to the greenhouse effect. Strong negative cloud feedback is exhibited in a recent 3D climate calculation of synchronously rotating M-star planets (54). If this result is correct, this would move the HZ inner edge for M stars to $S_{eff} \approx 2$. However, cloud feedback is actually calculated to be positive in many existing 3D climate models of present Earth, because high cirrus clouds, which tend to warm the climate, are

predicted to increase faster with increasing surface temperature than lower stratus clouds, which tend to cool it (55). Pushing suitably equipped 3D climate models up into the warm, moist greenhouse regime could help elucidate how clouds and relative humidity change with surface temperature and, thus, where the inner edge of the HZ actually lies.

In summary, much work remains to be done both to calculate reliable HZ boundaries and then to find and characterize rocky planets within them. However, this work is well worthwhile, as it could eventually lead to the detection of habitable exoplanets, and perhaps to evidence of extraterrestrial life.

ACKNOWLEDGMENTS. We thank two anonymous reviewers for constructive comments that helped us frame our arguments more precisely. Funding for this work was provided by the NASA Astrobiology Institute's Virtual Planetary Laboratory and the NASA Exobiology and Evolutionary Biology Research Program. The authors are affiliated with Pennsylvania State University's Center for Exoplanets and Habitable Worlds.

1. The Exoplanet Data Explorer [Internet]. Berkeley (CA): California Planet Survey. Available at www.exoplanets.org. Accessed November 11, 2013.
2. Table of Confirmed Planets [Internet]. Washington, DC: NASA. Available at <http://kepler.nasa.gov/Mission/discoveries/>. Accessed November 11, 2013.
3. Baross J, et al. (2007) *The Limits of Organic Life in Planetary Systems* (The National Academies Press, Washington, DC).
4. Joyce GF (1994) *Forward. Origins of Life: The Central Concepts*, eds Deamer DW, Fleischaker GR (Jones & Bartlett, Boston), pp xi–xii.
5. Joyce GF (1994) The RNA world: Life before DNA and protein. *Extraterrestrials—Where Are They? II*, eds Zuckerman B, Hart MH (Cambridge Univ Press, Cambridge).
6. Lederberg J (1965) Signs of life. Criterion-system of exobiology. *Nature* 207(992):9–13.
7. Lovelock JE (1965) A physical basis for life detection experiments. *Nature* 207(997):568–570.
8. Kasting JF (1992) Bolide impacts and the oxidation state of carbon in the Earth's early atmosphere. *Orig Life Evol Biosph* 20:199–231.
9. Zahnle K, Haberle RM, Catling DC, Kasting JF (2008) Photochemical instability of the ancient Martian atmosphere. *J Geophys Res Planets* 113:E11004.
10. Zahnle K, Freedman RS, Catling DC (2011) Is there methane on Mars? *Icarus* 212:493–503.
11. Kharecha P, Kasting JF, Siefert JL (2005) A coupled atmosphere-ecosystem model of the early Archean Earth. *Geobiology* 3:53–76.
12. Seager S, Schrenk M, Bains W (2012) An astrophysical view of Earth-based metabolic biosignature gases. *Astrobiology* 12(1):61–82.
13. Seager S, Bains W, Hu R (2013) A biomass-based model to estimate the plausibility of exoplanet biosignature gases. *Astrophys J* 775:104.
14. Berndt ME, Allen DE, Seyfried WE (1996) Reduction of CO₂ during serpentinization of olivine at 300°C and 500 bar. *Geology* 24:351–354.
15. Schrauzer GN, Guth TD (1977) Photolysis of water and photoreduction of nitrogen on titanium dioxide. *J Am Chem Soc* 99:7189–7193.
16. Grott M, Morschhauser A, Breuer D, Hauber E (2011) Volcanic outgassing of CO₂ and H₂O on Mars. *Earth Planet Sci Lett* 308:391–400.
17. Goldblatt C, Watson AJ (2012) The runaway greenhouse: Implications for future climate change, geoengineering and planetary atmospheres. *Philos Trans A Math Phys Eng Sci* 370(1974):4197–4216.
18. Ingersoll AP (1969) The runaway greenhouse: A history of water on Venus. *J Atmos Sci* 26:1191–1198.
19. Kasting JF (1988) Runaway and moist greenhouse atmospheres and the evolution of Earth and Venus. *Icarus* 74:472–494.
20. Walker JCG, Hays PB, Kasting JF (1977) *Evolution of the Atmosphere* (Macmillan, New York).
21. Donahue TM, Hoffman JH, Hodges RR, Jr., Watson AJ (1982) Venus was wet: A measurement of the ratio of deuterium to hydrogen. *Science* 216(4546):630–633.
22. Solomon SC, Head JW (1991) Fundamental issues in the geology and geophysics of Venus. *Science* 252(5003):252–260.
23. Kasting JF, Whitmire DP, Reynolds RT (1993) Habitable zones around main sequence stars. *Icarus* 101(1):108–128.
24. Goode PR, et al. (2001) Earthshine observations of the Earth's reflectance. *Geophys Res Lett* 28:1671–1674.
25. Walker JCG, Hays PB, Kasting JF (1981) A negative feedback mechanism for the long-term stabilization of Earth's surface temperature. *J Geophys Res* 86:9776–9782.
26. Forget F, Pierrehumbert RT (1997) Warming early Mars with carbon dioxide clouds that scatter infrared radiation. *Science* 278(5341):1273–1276.
27. Mischna MA, Kasting JF, Pavlov AA, Freedman R (2000) Influence of carbon dioxide clouds on early Martian climate. *Icarus* 145(2):546–554.
28. Selsis F, Kasting JF, Paillet J, Levraud B, Delfosse X (2007) Habitable planets around the star Gl581. *Astron Astrophys* 476:1373–1387.
29. Seager S (2013) Exoplanet habitability. *Science* 340(6132):577–581.
30. Zsom A, Seager S, de Wit J (2013) Towards the minimum inner edge distance of the habitable zone. *Astrophys J*, in press.
31. Manabe S, Wetherald RT (1967) Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *J Atmos Sci* 24:241–259.
32. Ramanathan V (1981) The role of ocean-atmosphere interactions in the CO₂ climate problem. *J Atmos Sci* 38:918–930.
33. Abe Y, Abe-Ouchi A, Sleep NH, Zahnle KJ (2011) Habitable zone limits for dry planets. *Astrobiology* 11(5):443–460.
34. Kasting JF, Holm NG (1992) What determines the volume of the oceans? *Earth Planet Sci Lett* 109(3–4):507–515.
35. Forget F, et al. (2013) 3D modelling of the early Martian climate under a denser CO₂ atmosphere: Temperatures and CO₂ ice clouds. *Icarus* 222:81–99.
36. Stevenson DJ (1999) Life-sustaining planets in interstellar space? *Nature* 400(6739):32.
37. Pierrehumbert R, Gaidos E (2011) Hydrogen greenhouse planets beyond the habitable zone. *Astrophys J Lett* 734.
38. Ramirez RM, et al. (2013) A CO₂-H₂O greenhouse for early Mars. *Nat Geosci*, in press.
39. Wordsworth R, Pierrehumbert R (2013) Hydrogen-nitrogen greenhouse warming in Earth's early atmosphere. *Science* 339(6115):64–67.
40. Levine M, Shaklan S, Traub W, Kasting J, Dressler A (2006) *TPF-C Science and Technology Definition Team Final Report*, available at http://exep.jpl.nasa.gov/TPF-C/tpf-c_index.cfm.
41. Roberge A, et al. (2012) The exozodiacal dust problem for direct observations of exo-Earths. *Publ Astron Soc Pac* 124:799–808.
42. Kopparapu R K, Ramirez R, Kasting J F, Eymet V, Robinson T D, Mahadevan S, Terrien R C, Domagal-Goldman S, Meadows V, Deshpande R (2013) *Habitable zones around main sequence stars: New estimates*. *Astrophys J* 765 ID 131.
43. Barnes R, Quinn T (2004) The (in)stability of planetary systems. *Astrophys J* 611:494–516.
44. Kopparapu R, Raymond SN, Barnes R (2009) Stability of additional planets in and around the habitable zone of the HD 47186 planetary system. *Astrophys J Lett* 695:181–184.
45. Kopparapu R, Barnes R (2010) Stability analysis of single-planet systems and their habitable zones. *Astrophys J* 716:1336–1344.
46. Barnes R, Raymond SN (2004) Predicting planets in known extrasolar planetary systems. I. Test particle simulations. *Astrophys J* 617:569–574.
47. Raymond SN, Barnes R (2005) Predicting planets in known extrasolar planetary systems. II. Testing for Saturn mass planets. *Astrophys J* 619:549–557.
48. Raymond SN, Barnes R, Kaib NA (2006) Predicting planets in known extrasolar planetary systems. III. Forming terrestrial planets. *Astrophys J* 644:1223–1231.
49. Dressing CD, Charbonneau D (2013) The occurrence rate of small planets around small stars. arXiv:1302.1647v2.
50. Seager S, Kuchner M, Hier-Majumder CA, Militzer B (2007) Mass-radius relationships for solid exoplanets. *Astrophys J* 669:1279–1297.
51. Kopparapu RK (2013) A revised estimate of the occurrence rate of terrestrial planets in the habitable zones around Kepler M-dwarfs. *Astrophys J Lett* 767.
52. Gaidos E (2013) Candidate planets in the habitable zones of Kepler stars. *Astrophys J* 770:.
53. Bonfils X, et al. (2013) The HARPS search for southern extra-solar planets XXXI. The M-dwarf sample. arXiv:1111.5019v2.
54. Yang J, Cowan NB, Abbot DS (2013) Stabilizing cloud feedback dramatically expands the habitable zone of tidally locked planets. *Astrophys J Lett* 771:L45.
55. Solomon S, et al. (2007) *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, 2007* (Cambridge Univ. Press, Cambridge).
56. Baraffe I, Chabrier G, Allard F, Hauschildt PH (1998) Evolutionary models for solar metallicity low-mass stars: Mass-magnitude relationships and color-magnitude diagrams. *Astron Astrophys* 337:403–412.
57. Baraffe O, Chabrier G, Allard F, Hauschildt PH (2002) Evolutionary models for low-mass stars and brown dwarfs: Uncertainties and limits at very young ages. *Astron Astrophys* 382:563–572.

Supporting Information Appendix

Inner Edge of the Habitable Zone for Dry Planets

A climate system adjusts to a new equilibrium state in response to a change in the amount of energy it receives. Energy balance must be achieved at both the top of the atmosphere and at the surface. Surface energy balance dictates that the net radiative flux into the surface (F_{net}) must be equal to the sum of the latent (F_L) and sensible (F_S) heat fluxes, which represent the energy carried upwards by convection [1,2,3,4,5,6]. F_{net} is the sum of the net absorbed solar, F_{sol} , plus net outgoing infrared, F_{IR} , surface fluxes: $F_{net} = F_{sol} + F_{IR}$. Therefore,

$$F_{sol} + F_{IR} = F_L + F_S \quad (1)$$

It should be remembered that the latent heat is associated with the phase change of a substance, for example, from liquid to vapor state, whereas sensible heat is associated with the change in the temperature of the substance with no phase change. This temperature change can either arise from conduction (through contact) or absorption of solar radiation by the surface. We can write the sum of the latent and sensible heat fluxes as the total convective heat flux: $F_C = F_L + F_S = F_L(1 + B)$. Here, B is the Bowen ratio, defined as F_S/F_L . An estimation for how B varies with surface temperature is given by [7], and is reproduced in Fig.1. Following Eq.(1), the total convective heat flux is equal to F_{net} . Thus, we can write:

$$F_C = F_{net} = F_L(1 + B) \quad (2)$$

The latent heat flux, F_L , appearing in Eq. (2) depends on the surface relative humidity, RH. Our surface RH parameterization was derived starting with the following commonly used expression, ([8], Eq.(4.27)):

$$F_L = LC_D u \rho (q_0 - q_1) \quad (3)$$

Here, L is the latent heat of vaporization of water (2.5×10^6 J/kg), C_D is the drag coefficient at layer 1 just above the surface, u is the mean horizontal wind speed near the surface in m/s, ρ is the atmospheric sea level mass density in kg/m^3 , q_0 is the surface H_2O saturation mixing ratio, and q_1 is the H_2O mixing ratio at layer 1. Layer 1 is then chosen to be at an infinitesimally small distance above the surface such that $q_1 = R_{surf} * q_0$, where R_{surf} is the surface RH. Substituting this expression into Eq.(3) yields:

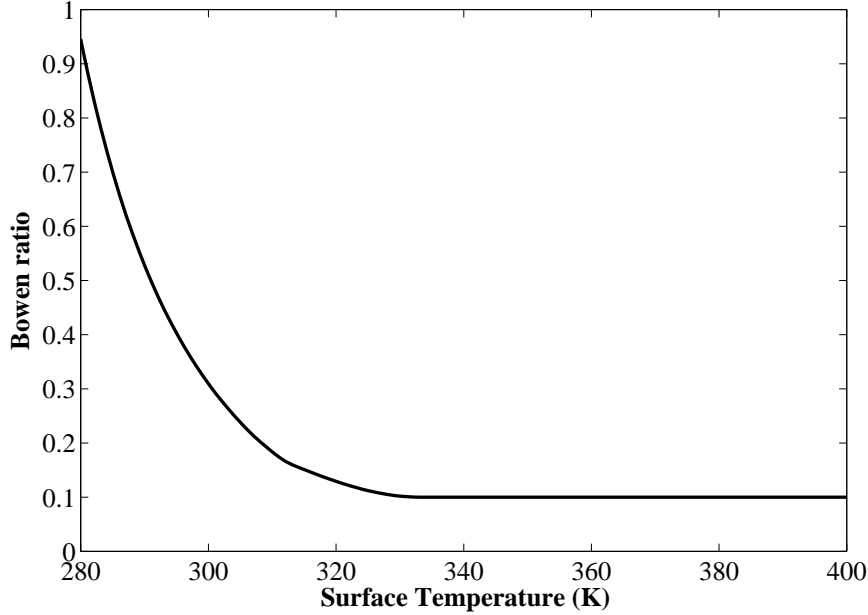


Fig. 1.— Bowen’s ratio (F_S/F_L) as a function of surface temperature [7].

$$F_L = LC_D u \rho q_0 (1 - R_{surf}) \quad (4)$$

Thus, as R_{surf} increases, so does q_1 , moderating the increase of the latent heat flux, F_L , which would otherwise increase exponentially because of the temperature dependence of q_0 . To calculate R_{surf} from the surface energy balance, we need the Bowen ratio, B .

Substituting Eq.(4) in Eq.(2) gives

$$F_{net} = LC_D u \rho q_0 (1 - R_{surf}) (1 + B) \quad (5)$$

The drag coefficient C_D and wind speed u are both unknowns in our expression and cannot be calculated self-consistently in our 1-D model. Thus, we make the assumption that the product $C_D \cdot u$ remains constant at all surface temperatures. This assumption is reasonable for a 1-D climate model, as C_D does not vary strongly over water surfaces [9]. To improve on this assumption, one would need to move to a 3-D model that included horizontal winds and boundary layer physics.

From Eq.(5), one can see the problem if the surface RH is fixed while the surface temperature is varied. As the surface temperature rises, q_0 rises exponentially, following the Clausius-Clapeyron equation, leading to a large increase in the right-hand side that cannot

be sustained, and which cannot remain in balance with the net absorbed flux (F_{net}). To moderate this increase in F_L , R_{surf} must increase as surface temperature increases¹.

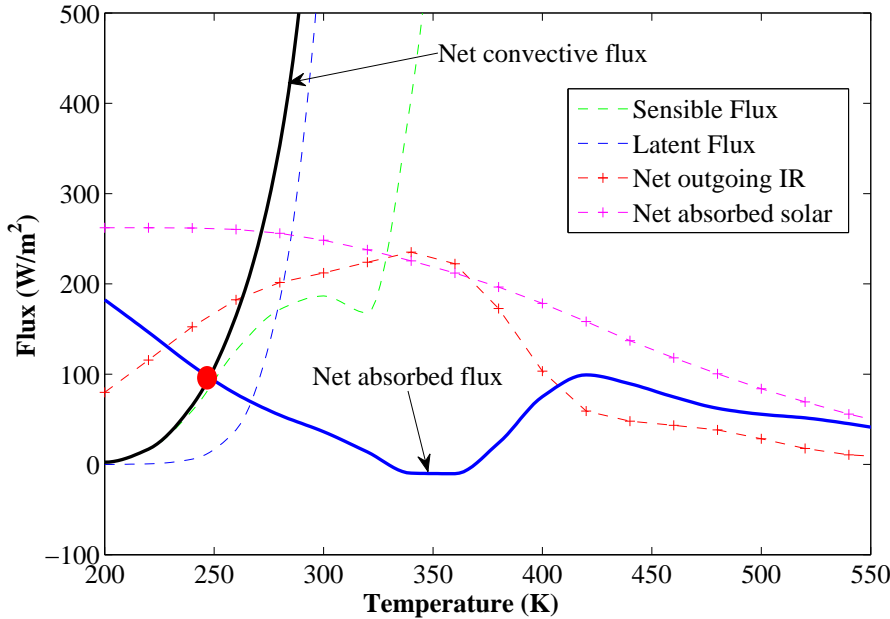


Fig. 2.— Assuming a fixed surface relative humidity, R_{surf} , of 1%, this figure shows the large increase in the net convective heat flux (black solid curve), which is dominated by the latent heat flux (dashed gray curve). Surface energy balance is maintained between the net convective heat flux and the net absorbed radiative flux (solid blue curve) at only one temperature (red filled circle) that is much colder than the typical temperatures encountered at the inner edge of the HZ. Beyond this point, any increase in temperature leads to an unphysical situation in which the net convective flux is greater than the net absorbed flux.

Fig. 2 illustrates this problem by showing various fluxes as a function of surface temperature, assuming R_{surf} is fixed at 1%. There is only one temperature (red filled circle) at which the net absorbed flux at the surface (blue solid curve) equals the net convective flux (black solid curve). Beyond this point (i.e., at higher surface temperatures) the net convective flux increases exponentially, chiefly due to the increase in the latent heat flux. Surface energy balance cannot be achieved at these temperatures if liquid water is present. Hence, the only physically possible solution is for the surface to be dry. But this, then, is not a

¹ B decreases with an increase in temperature, while ρ first decreases marginally, then increases beyond 350 K, which exacerbates the increase in F_L .

habitable planet. Surface liquid water might be stable in cooler regions, as in the "Dune" planets discussed in the main text, but such planets cannot be simulated self-consistently in a 1-D climate model.

Another way to look at this problem is to calculate the lifetime, t_{life} , of liquid water against evaporation assuming some water content on the surface of the planet:

$$t_{life} = \frac{L\rho\Delta z}{F_L} \quad (6)$$

where L and F_L are defined in Eq. (3), ρ is the density of water and Δz is the thickness of the layer of water spread evenly on the planet's surface. For the case of $R_{surf} = 1\%$ (Fig. 2), at a surface temperature of 300 K, $F_L = 600 \text{ Wm}^{-2} = 6 \times 10^5 \text{ erg cm}^{-2} \text{ s}^{-1}$. Assuming the same amount of water content as Earth's oceans ($1.4 \times 10^{21} \text{ kg}$, which is unphysical for $R_{surf} = 1\%$, but we are aiming here for the longest possible time scale), $\Delta z \sim 3 \times 10^5 \text{ cm}$. Substituting these values in Eq.(6), $t_{life} = 1.25 \times 10^{10} \text{ seconds} \sim 400 \text{ years}$. The authors of [10] estimate that hot desert worlds have 100 times smaller liquid water content than Earth. If that is true, then $t_{life} \sim 4 \text{ years}$. Therefore, even with the most optimistic estimates of water content on the surface, the lifetime of the liquid water against evaporation is very small on a planet with low relative humidity.

REFERENCES

- [1] Ramanathan, V., *The Role of Ocean-Atmosphere Interactions in the CO₂ Climate Problem*, Journal of the Atmospheric Sciences, 38(1981), 918-930
- [2] Boer, G. J., *Climate Change and the Regulation of the Surface Moisture and Energy Budgets*, Climate Dynamics, 8(1993), 225-239
- [3] Trenberth, K. E., *Atmospheric Moisture Residence Times and CyclingL Implications for Rainfall Rates and Climate Change*, Climate Change, 39(1998), 667-694
- [4] Allen, M. R., & Ingram, W. J., *Constraints on Future Changes in Climate and the Hydrologic Cycle*, Nature, 419(2002), 224-232
- [4] Held, I. M., & Soden, B. J., *Robust Responses of the Hydrological Cycle to Global Warming*, American Meteorological Society, 19(2006), 5686-5699
- [5] Soden, B. J., & Held, I. M., *An Assessment of Climate Feedbacks in Coupled Ocean-Atmosphere Models*, Journal of Climate, 19(2006), 3354-3360

- [6] Davies-Jones, R., *An Efficient and Accurate Method for Computing the Wet-Bulb Temperature along Pseudoadiabats*, Monthly Weather Review, 136(2008), 2764-2785
- [7] Garratt., *The atmospheric boundary layer*. Cambridge University Press., 1992, 316 pp.
- [8] Hartmann, D. L., *Global Physical Climatology*. Academic Press, 1994, 411 pp.
- [9] Arya, P. 2001. *Introduction to micrometeorology*, 79, Academic Press, second edition
- [10] Zsom, A., Seager, S., De Wit, J., & StamenKovic, V., *Towards the Minimum Inner Edge Distance of the Habitable Zone*, Accepted to Astrophysical Journal, 2013.