PROCEEDINGS OF SCIENCE

Stellar, Galactic, and Super-Galactic Habitable Zones

Paul A. Mason[a,b,*] **and Peter L. Biermann**[c,d]

[a]New Mexico State University,
MSC 3DA, Las Cruces, NM 88003, USA
[b]Picture Rocks Observatory and Astrobiology Research Center,
1025 S. Solano, Suite D, Las Cruces, NM 88003, USA
[c]Max Planck Institute for Radio Astronomy, Bonn, Germany
[d]Univ. of Alabama, Tuscaloosa, AL, USA

E-mail: `pmason@nmsu.edu`

The concept of habitable zones for planets in astrobiology is briefly reviewed. Progress has been significant in the nearly century-long effort to theoretically characterize habitable planets, The Liquid Water Belt of Shapley has been revised to include a wide variety of modern concepts. These efforts are all related to the location of planets in both space and time, which might be expected to be capable of supporting both simple and complex life as we know it. In particular, circumstellar habitable zones around single and binary stars have been proposed and depend mainly on the luminosity of the host star(s). Ultraviolet (UV) habitable zones have also been proposed to locate orbital distances and spectral types with sufficient UV to promote the origin of life and with UV below destructive levels. Circumbinary habitable zones are quite varied, with some very harsh environments and others with sedate conditions. Some binaries experience tidal torques that reduce stellar rotation and magnetic activity, the so-called Binary Habitability Mechanism, and thereby possess "better than Earth" habitability conditions. Across the spectrum of possibilities, it is important to distinguish between locations allowing for the origin of life and those providing long-term viability of life. Galactic habitable zones are defined according to the availability of elements for making potentially habitable planets as well as non-habitable zones characterized by frequent threats to complex life such as a nearby supernova, gamma-ray bursts, accreting black holes, frequent asteroid or comet impacts, and high levels of Galactic cosmic rays. Usually, the Galactic Habitable Zone is represented by an annulus in the Galactic disk, but much depends on the orbital history of the planetary system and the merger and star-formation history of the host galaxy. Avoiding star formation regions for extended periods in a low star formation rate disk galaxy can be of great benefit, conducive to the development of complex life. Furthermore, we suggest that the concept may be extended to the Super-Galactic Habitable Zone (SGHZ), defined as hospitable regions within clusters of galaxies, bounded by inhospitable active galaxies and low-metallicity dwarf galaxies in the outskirts of galaxy clusters. The SGHZ - if defined just for survival and not including the birth of complex life - is a large network of connected local habitable zones. The birth places for complex life form a smaller disconnected set of islands.

Multifrequency Behaviour of High Energy Cosmic Sources XIV (MULTIF2023)
12-17 June 2023
Palermo, Italy

---

[*] Speaker

PoS(MULTIF2023)081

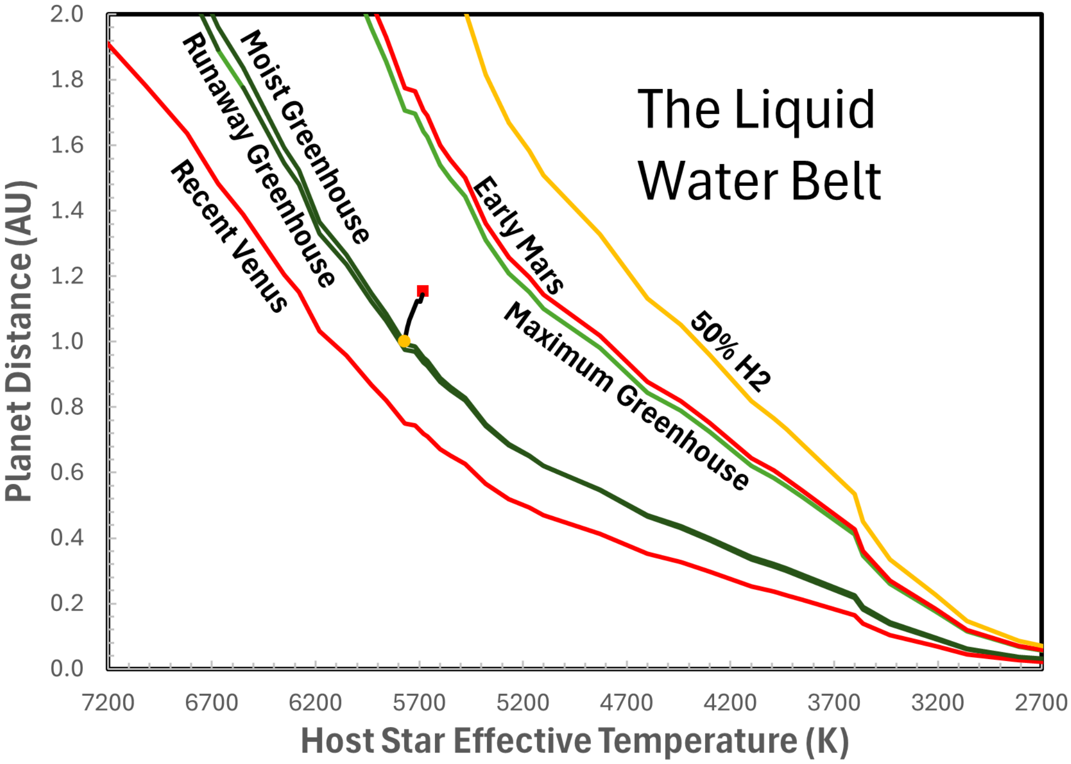

Figure 1: The liquid water belt, usually called the Circumstellar Habitable Zone (CHZ) is shown. The lines represent the inner and outer edges of habitability based on different assumptions. The optimistic CHZ is shown by the red lines giving the Recent Venus and Early Mars empirical limits. The greenhouse limits are shown as green lines. The volcanic habitable zone's extreme outer limit (orange line) is based on a 50% hydrogen (H_2) atmosphere. The yellow dot is the current position of the Earth and the red square is the position of the Archean Earth, 3.5 Gyr ago, when life first appeared and the Solar luminosity was 75% of the present value. A Solar evolution model is used to calculate L and T_{eff} for the Sun to produce the path of the Earth for the last 3.5 Gyr.

1. Introduction: The Liquid Water Belt

The quest for locating life in the universe has long included the concept of a habitable zone. Edward Maunder [1] appears to be the first to have used the term "Habitable Zone" in his 1913 book "Are the Planets Inhabited?". Harlow Shapley coined the term "Liquid Water Belt", equivalent to the Habitable Zone term as it is used today. The term ecosphere, for ecology sphere, was also often used. This idea is focused on the recognition that the presence of liquid water is fundamental to life as we know it. It is based on the amount of stellar flux incident at the top of the planet's atmosphere. Implying that the luminosity, temperature, and main sequence lifetime of the planet's host star are the major determining factors for the limiting distances allowing surface water and thereby the potential for life. In addition to the luminosity of the host star, the surface temperature of a planet depends on the planet's albedo and the presence and quantity of greenhouse gasses, CO_2,

NH_3, and H_20 vapor.

The quantification of the Circumstellar Habitable Zone (CHZ) [2,3], replaced the more accurately descriptive Liquid Water Belt, with the same goal, namely that of estimating the minimum and maximum distance from the Sun that would allow water on the surface of an Earth-like planet and then extending the result to other main sequence stars. The planet's distance for various habitable zone limits versus the effective temperature of the host star is shown in Figure 1. Atmospheric models allowed the Earth-like planet to have a different atmosphere than Earth. The original models included the greenhouse effect but did not include the carbonate-silicate cycle which regulates Earth's temperature through negative feedback [4], this extended the outer limit to greater distances than first derived [3] and provides a natural stabilizing effect, or thermostat, on the surface temperature [4]. This was done first with the Sun and then extrapolated to other main-sequence stars. Various authors have modified the CHZ limits over the years, with some shifting back and forth and the inclusion of atmospheric compositions that are different than Earth's. See Appendix A for a list of published Solar habitable zone limits.

2. The Circumstellar Habitable Zone (CHZ)

The liquid water belt or more commonly the circumstellar habitable zone (CHZ) is the annulus surrounding a star (or close binary) in which liquid water should be stable on a planetary surface given some atmosphere composition, most importantly greenhouse gasses, and pressure. In Figure 1, the liquid water belt is shown using several definitions [5,6,7] for the circumstellar habitable zone. The optimistic or empirical limits on the CHZ are shown in red. The outer optimistic or empirical limit is called the early Mars limit and is based on the apparent presence of water on Mars's surface before 3.8 Gyr ago [6]. The assumption is made that Mars had surface water some 3.8 Gyr ago based on Martian surface features showing flowing water that were formed when the Solar luminosity was about 75% of its present value [8]. Early Mars flux is about 0.32 of the flux received at Earth today. The inner optimistic or empirical habitable zone limit is called the present Venus limit and is based on Venus not having surface water for the past 1 Gyr and it assumes that before 1 Gyr ago Venus did have water on its surface, when the Solar luminosity was 92% of the present value [6] or 1.76 times the flux received at Earth today [8].

At the inner edge of the CHZ, like the present-day position of Earth, a planet must have very little CO_2 in its atmosphere, because if there is too much Greenhouse Effect, then runaway greenhouse conditions will cause super-heating of the atmosphere and subsequent water loss. At the outer CHZ edge, a planet must have $\sim$8 bars of CO_2 to maintain surface water in liquid form, compared to a pressure of $\sim$ 1 bar on Earth, depending on elevation. A planet without enough greenhouse gasses, CO_2 or NH_3, to heat its atmosphere sufficiently for its position in the CHZ, will freeze, producing global glaciation, at least until sufficient CO_2 is supplied by volcanoes to the atmosphere to cause additional greenhouse warming and to allow temperate conditions. The maximum greenhouse heating limit can be extended even further if the planet's atmosphere also contains molecular hydrogen from, for example, sustained high-level volcanic activity. The outer (extreme) volcanic outer limit (Figures 1-3: orange line) is based on an atmosphere including 50% molecular hydrogen [7] that would be lost to space if not resupplied by volcanoes. Solutions with less H_2 are correspondingly closer to the maximum greenhouse curve of Figure 1, like the 5% H_2

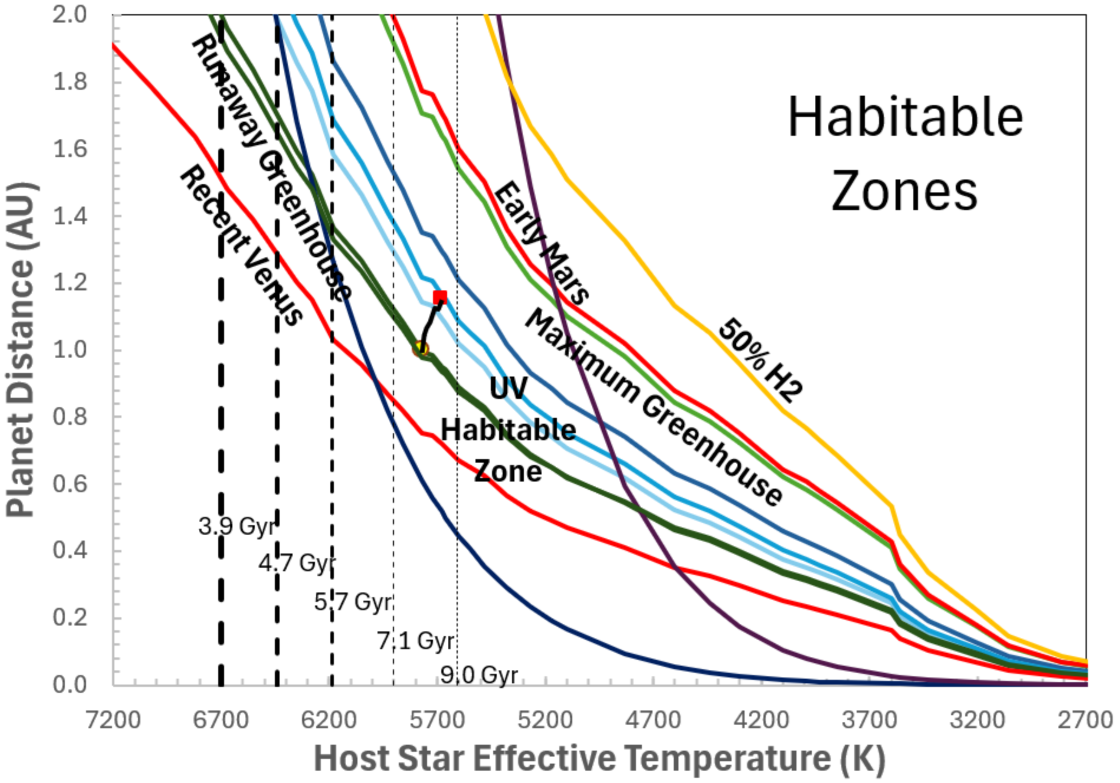

Figure 2: Circumstellar Habitable Zones (CHZ), including the optimistic (red lines) and conservative (green lines) CHZ are shown based on the luminosity and temperature of the host-star. The outer (extreme) volcanic outer limit (orange line) is based on an atmosphere including 50% molecular hydrogen (H_2). The Ultraviolet Habitable Zone (UVHZ) limits are shown as purple lines. The yellow dot shows the present position of the Sun. To show the evolution of the habitable zones over time, the Earth moves along the black line in this diagram, while in reality, the Earth remains at 1 AU and the habitable zone limits move outwards in the Solar System with time. The red square is the Earth's position at the time of earliest photosynthetic life, 3.5 Gyr ago. The vertical dashed lines show the CHZ lifetime of planets, ~ 90% of the stellar main sequence lifetime, as a function of host-star effective temperature and correspond to 1.4 $M_{\odot}$, 1.3 $M_{\odot}$, 1.2 $_{\odot}$, and 1.1 $M_{\odot}$ stars from left to right.

outer CHZ limit shown in Figure 3. The host-star temperature, T_{eff} , is related to the luminosity L, by an empirical main-sequence L-T_{eff} relation, and the planet distance is related to the luminosity by the inverse square law. A Solar evolution model [8] is used to calculate L and T_{eff} to produce the path of the Sun in Figures 1, 2, and 3.

A one-dimensional cloud-free radiative-convective atmospheric model, developed over many years [4,5, and references therein], was used to determine the globally averaged atmospheric T-P relation. These models include N_2, CO_2, and H_2O vapor atmospheres, like Earth. Note that, today, the main constituent (78%) of Earth's atmosphere is N_2, which like 0_2 (21%) is not a greenhouse gas. The inner habitable zone limit is the Runaway Greenhouse Limit of an atmosphere with CO_2

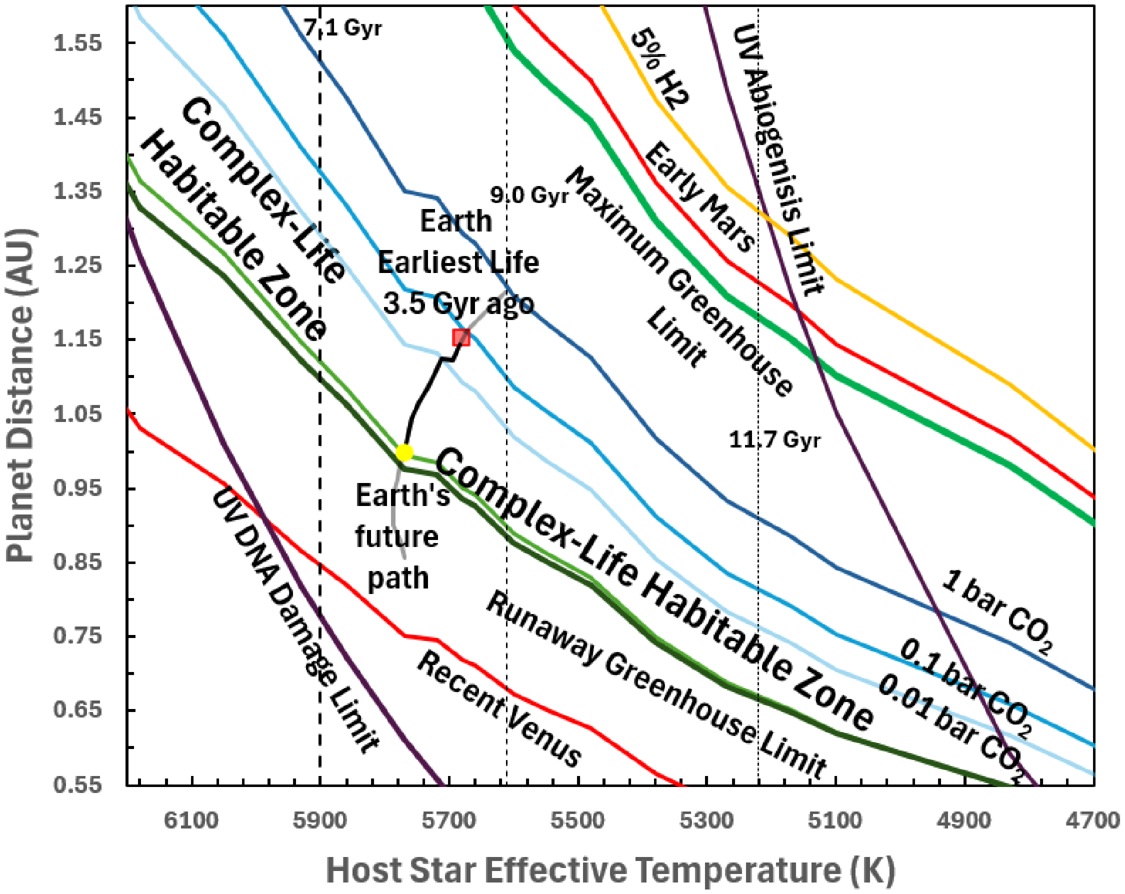

Figure 3: The Complex Life Habitable Zones (CLHZ). See Figure 2 caption, except that here the 5% H_2 atmosphere outward extension to the Maximum Greenhouse Limit is shown as the orange line. The black line shows the evolution of the Solar flux received at Earth in this diagram, based on the evolution of the Sun for the last 3.5 Gyr, while the faint gray line shows the Earth's path for 0-8 Gyr of the Sun's evolution. The left edge at 6200 K corresponds to a 1.2 $M_\odot$ star with a 5.7 Gyr habitability lifetime, while the dashed lines correspond to a 1.1 $M_\odot$, 1.0 $M_\odot$ and 0.9 $M_\odot$ star respectively with lifetimes labeled in the figure.

and H_2O vapor. The Moist Greenhouse Limit is the inner CHZ limit when the atmosphere is saturated by water vapor. Today, the Earth, which is shown as the yellow circle at 1 AU, in Figures 1, 2, and 3, lies very close to the Moist Greenhouse Limit at 0.99 AU. This result is somewhat deceptive as the Moist Greenhouse Limit is based on a saturated atmosphere, which may not apply to Earth's future. The outer CHZ limit is the Maximum Greenhouse Limit and is defined as the distance for which CO_2 condensation and CO_2 scattering dominate over greenhouse warming. The Runaway Greenhouse Effect is defined as when surface temperatures exceed the critical point of H_2O at about 647K and 220 bar and H_2O loss from the atmosphere to space will be rapid [5]. For comparison, the surface temperature of Venus is 737 K.

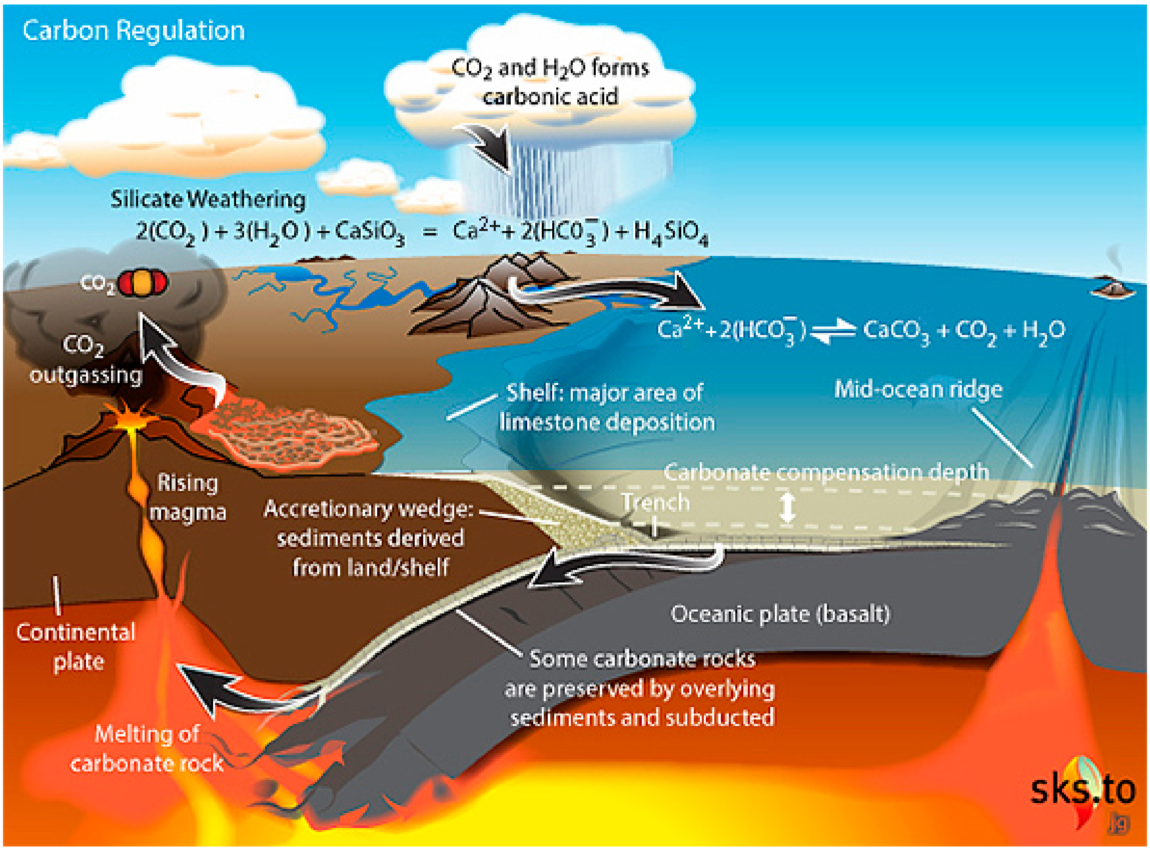

Figure 4: A schematic depiction of the carbonate-silicate cycle. Atmospheric carbon dioxide (CO_2) interacts with rock (containing the elements Mg, Si, Ca, and so on). Carbon is trapped in the rocks and moves as sediment by rivers to the oceans, becoming part of the oceanic crust. Plate tectonics, caused by the up-welling of magma from the very hot mantle, pushes the plates together causing the denser oceanic plate to subduct below a continental plate, carrying carbon with it into the mantle. After a million or so years the carbon is returned to the atmosphere by volcanoes as CO_2. Credit: Wikipedia free use.

2.0.1 The Carbon Cycle

The carbon cycle as it operates on Earth consists of two parts. These are the carbonate-silicate cycle, accounting for 80% of carbon cycling, and the organic carbon cycle, accounting for the other 20%. The carbonate-silicate cycle simply involves the movement of carbon between the atmosphere, oceans, surface crust, and mantle, see Figure 4. The process is straightforward. Water in the ocean absorbs energy from the Sun and evaporates to form clouds. Next, rain from the clouds falls from the sky, converting its gravitational potential energy into the kinetic energy of the raindrops. The drops crash into the rock causing the erosion of the rock and a process commonly called chemical weathering. Carbonic acid, H_2CO_3, is formed when CO_2 combines with H_2O. Carbonic acid then breaks down minerals in the rock, for example, carbonic acid in a reaction with calcite produces calcium plus bicarbonate, $H_2CO_3 + CaCO_3 \rightarrow Ca_2^+ + 2HCO_3^-$. So, by this and other reactions, see Figure 4, carbon is removed from the atmosphere. From there, rivers bring the water and sediment down to the ocean. The carbon often ends up at the bottom of the ocean. Eventually, because of plate tectonics, the oceanic crust is subducted below a crustal plate, sending carbon into the mantle. Water acts as a lubricant between the plates during subduction. The mantle contains an enormous quantity of CO_2 and H_2O which are slowly released back into the atmosphere in volcanic eruptions thus completing the carbonate-silicate cycle,

The organic carbon cycle involves photosynthesizing organisms, that span the full range of sizes for living organisms, from tiny bacteria to enormous trees. If present, they also pull carbon from the atmosphere by converting CO_2 into O_2 using water as a catalyst. Photons from the Sun supply the essential energy source. Negative feedback in the organic carbon cycle can maintain a temperate Earth and potentially some exoplanets as well. If there's an excess of greenhouse gases in the atmosphere, then the planet's surface will heat due to the Greenhouse Effect. The heating of the planet's surface causes more evaporation of water from the oceans. This causes more rain, which then results in more rock erosion and chemical weathering. Warmer conditions can also cause plant life to thrive, resulting in the increased removal of CO_2 from the atmosphere. The reduction of CO_2 in the atmosphere causes the atmosphere to have a weaker Greenhouse Effect and therefore it cools. A cooler surface will reduce photosynthetic productivity and thereby CO_2 will gradually accumulate in the atmosphere, thus completing the negative feedback loop.

If too much cooling occurs, then the reverse happens. Less evaporation means less rain and therefore less weathering, which means more cooling. A positive feedback cycle may ensue. In addition, colder conditions for plant life compromise their growth, O_2 production, and removal of CO_2 from the atmosphere. An interesting complication is that volcanoes also inject pulverized rock into the atmosphere producing a haze that reflects sunlight into space and cools the atmosphere. So the effect of volcanoes is first to cool the atmosphere by the injection of aerosol haze and then to heat it, after the dust settles, due to increased atmospheric CO_2.

If the carbonate-silicate cycle is too efficient then over time CO_2 is removed from the atmosphere faster than it can be transferred from the mantle back into the atmosphere through volcanoes. Then the surface of the Earth will gradually cool. Less evaporation and less rain will occur. The result is that the amount of rock weathering is reduced unless atmospheric CO_2 is replaced by volcanic out-gassing. Likewise, the colder environment means less plant activity, which also reduces the rate of CO_2 extraction from the atmosphere. This finally reduces the greenhouse effect, thereby cooling the surface. The ocean plays an important role as a buffer in all of this. If CO_2 is increased in the atmosphere, then about half of this CO_2 is absorbed by the oceans. Likewise, when CO_2 is removed from the atmosphere, some of it is replaced by out-gassing from the oceans.

The presence of sufficient amounts of exposed crust as well as sufficient surface water are essential components of a planet habitable for complex life. For example, the more exposed rock the planet has, the more weathering is possible. Weathering of rock is impossible if not for rain. The biosphere benefits from weathering because sediments provide nutrients like phosphorus that would otherwise remain locked up in rocks. Life thrives at the intersection of sunlight, water, and rock. However, a mixed land and water planet is not as likely as either a dry planet or a full water world. Small land masses tend to erode fairly quickly back into the ocean, take Hawaii as an example. The older islands are quickly eroding into the ocean because there's no longer a volcano making them grow. Only the Big Island of Hawaii has active volcanoes as the hot spot at the top of the mantle slowly moves with respect to the island chain. It is growing faster than erosion is wearing it away, while the other Hawaiian Islands are rapidly eroding back into the Pacific Ocean.

On the other hand, a mostly land planet with small oceans or lakes is susceptible to drying out further because a smaller ocean is not efficient at removing CO_2 from the atmosphere.

The negative feedback of the carbonate-silicate cycle works well on a mixed ocean plus continent world. However, a positive feedback loop causing global glaciation may occur when the planet's surface is too cold and ice, especially sea ice, builds up at the poles. The increased ice is efficient at reflecting sunlight and therefore increases the albedo, which is defined as the total reflectivity of the planet. The increased albedo causes increased atmospheric cooling and more conversion of the very dark ocean surfaces to bright ice and with that the falling of sea levels. This runaway glaciation results in a predominantly ice-covered planet. Such a snowball Earth has occurred several times in Earth's past although liquid water may have always remained near the equator. The positive feedback causing the snowball condition is reversed if and only when volcanic out-gassing increases the quantity of greenhouse gases in the atmosphere is sufficient to begin to warm the planet again. Unlike snowball Earth, the reverse process of global heating causing the evaporation of the oceans in the runaway greenhouse effect can result in the removal of all water from the atmosphere. Unlike snowball Earth, the runaway greenhouse effect is non-recoverable.

Recent work [9] uses simulations of Earth's climate across geological time to derive a present-day runaway greenhouse effect for an Earth-like planet in the Solar System inside 0.982 AU, which is in between the Moist Greenhouse Limit (presently at 0.9935 AU) and the Runaway Greenhouse Limit (presently at 0.9760 AU) shown in Figures 1,2, and 3 [6]. Beyond 1.18 AU the climate would undergo limit-cycling between global glaciation and temperate climates with polar caps [9]. This work places Earth in the center of a narrow Solar CHZ. Accordingly, with this intermediate determination of the inner CHZ limit, the evolution of the Sun [8] will cause the revised inner limit [9] to move past the Earth's orbit 1.0 Gyr from now when the Earth is 5.5 Gyr old. See Appendix A for a list of most of the published habitable limits, spanning more than 60 years.

2.0.2 The Ultraviolet Habitable Zone (UVHZ)

The principles of the ultraviolet habitable zone (UVHZ) are as follows. Too much UV radiation on the surface of a planet, especially if not protected by atmospheric ozone (O_3), is damaging to DNA and deadly to living cells. Too little UV radiation and life may not originate at all, as UV radiation is potentially essential for the origin of life. For this reason, the UVHZ concept is related to the abiogenesis (origin of life) habitable zone. Being in the CHZ is a prerequisite for presence in the UVHZ.

Buccino et al. [9] defined the boundaries of the UVHZ, see Figure 3, by considering the intensity of Solar UV radiation that is incident on the Earth during the Archean. The inner boundary of the UVHZ, labeled the DNA Damage Limit, is the distance corresponding to the maximum UV-radiation dose tolerable for biological systems, taken to be twice Archean levels. The outer boundary of the UVHZ, labeled the Abiogenesis Limit is defined based on the minimum UV-flux needed for the chemical synthesis of complex molecules, like amino acids and lipids, taken to be 50% of the Archean Earth level [10]. Studies of the UVHZ have focused on the UV portion of the blackbody radiation emitted by a star and not on the radiation associated with magnetic activity, such as flares and coronal mass ejections. Host-star magnetic activity decreases with age, so the UVHZ becomes important at later times. The UVHZ favors the hotter main sequence hosts, namely the F, G, and K-types over the cooler M-type stellar hosts because the latter have insufficient UV radiation at all points within the CHZ (see Figures 2 and 3).

2.0.3 The Complex Life Habitable Zone (CLHZ)

A more restrictive Complex Life Habitable Zone (CLHZ) is a subset of the CHZ with the additional requirement for low levels of CO and CO_2. The CLHZ outer limits are shown in Figures 2 and 3, as blue lines of increasing CO_2 partial pressures of 0.01 bar, 0.1 bar, and 1 bar respectively [11], moving outwards in steps from the Moist Greenhouse Limit. For reference, at sea level on Earth, the partial pressure of CO_2 is 4.15 x 10^{-4} bar. The CLHZ corresponds to the inner region of the CHZ as the outer regions require a CO_2 partial pressure that is up to ~8 times the total pressure of Earth's atmosphere (1 bar). For the middle and outer regions of the CHZ, high CO_2 levels are toxic for life as we know it. Low pH of the shallow oceans and other surface water would result. High CO_2 concentrations are harmful to respiration in air-breathing animals, especially at low O_2 levels. The high atmospheric pressure of several bars of CO_2 would compromise lung function. It is worth noting that the pressure at the surface of Venus is about 92 times that at the surface of Earth. The atmosphere of Venus consists of 96.5% CO_2 and 3% N_2. The intersection of the CLHZ and the UVHZ is the ideal region for the search for extraterrestrial life.

Carbon monoxide (CO) also presents a problem for complex life on planets with K and M-type host stars. Photochemical reactions expected in the atmospheres of planets orbiting K and M-type stars produce a high concentration of CO. Partly because it is similar to O_2, CO is extremely toxic for life as we know it. Because we have only one example namely Earth life, however compelling, the validity of the CLHZ is uncertain given life's ability to adapt. The strict CLHZ as adopted here is bounded on the inside by the Moist or Runaway Greenhouse Limit and the first carbon dioxide limit of 0.01 bar of CO_2, the first of the blue lines shown in Figures 2 and 3.

2.1 The Continuous Circumstellar Habitable Zone

Further consideration proves that the habitability question is significantly more complicated than stated. First of all, stars evolve. Stellar lifetimes range over many orders of magnitude. It is immediately clear that if the lifetime of the planet-hosting star is short compared to the time scale for the origin and development of life including complex life (~4 Gyr in the case of the Earth), then complex life will not be possible. So optimistically, complex life is thus prohibited for host stars with (habitability) lifetimes less than ~ 3 Gyr. This excludes O, B, A, and most F-type stars. Clearly, G, K, and M-type main sequence stars have sufficiently long lifetimes for life as we know it, based on our example of the Solar System. Combining the continuous CHZ with the UHZ, for the short-lived F-types yields a very narrow continuous UHZ, allowing for only relatively Solar-like stars; see the habitability lifetimes shown in Figures 2 and 3. The increase in the luminosity of the Sun has been significant. At the formation of life 3.5 Gyr ago the Sun was only 75% as luminous as it is today. We note that the longer-lived K and M-type stars also increase in luminosity and change color somewhat as they age, but on timescales much longer than the development of complex life on Earth, which is about 4 Gyr, when the Cambrian Explosion occurred.

To accommodate the effects of stellar evolution on the CHZ, the concept of the continuous CHZ was developed and refined [e.g. 3]. Stars increase in luminosity gradually as they age, see the Solar evolution curve in Figures 1, 2, and 3. The temperature and luminosity of the Sun changed over time and the CHZ distance limits changed accordingly. This means that the location of the formation of the planet within the CHZ is of critical importance to its habitability.

2.1.1 Photosynthetic Biomass Production

If formed close to the CHZ inner edge, then the potential for photosynthetic production is maximized. This results from the inverse square law of light. However, as the host star brightens with age, the Runaway Greenhouse Limit will move past the planet, triggering the end of habitability. For Solar-like G-type stars this will occur in a time short compared with biological timescales. So, while life may originate and thrive for some time on planets formed orbiting near the CHZ inner edge, there might not be enough time for photosynthetic life to change the atmosphere sufficiently, by the production of O_2 and the removal of CO_2 from the atmosphere, to develop complex life. While liquid surface water is the key ingredient for life, water and CO_2 are key to photosynthetic life, and water, O_2, and nutrients like N_2 and phosphorus (P) are key to complex life as we know it on Earth.

On the other hand, a planet formed at the outer edge of the habitable zone benefits from having the maximum time in the continuous CHZ for any particular host-star spectral type. However, its photosynthetic productivity is compromised, and a significant quantity of greenhouse gas must be present in the atmosphere to maintain liquid water. The amount of time needed for photosynthetic life to replace several bars of atmospheric CO_2 pressure with O_2 is burdened with 3 to 4 times less photosynthetic production than Earth, as it is significantly further from the star, is extended to at least 10-15 Gyr. Even under the most optimistic assumptions, the photosynthetic biosphere-driven cooling that has taken place, ever since the Great Oxygenation Event, 2.5 Gyr, took about 2 Gyr after the formation of Earth to occur. This effect is mitigated, if the host star is a longer-lived K-type star for example. However, the geological lifetime for a planet to maintain active volcanoes, plate tectonics, and a magnetic dynamo would likely be the limiting factor for life on planets orbiting long-lived stars. This means that planets would cool down to inactivity before photosynthetic life could oxygenate the atmosphere unless the planet is near the inner edge of the CHZ for all spectral types.

2.1.2 Solar System Examples

Mars is the Solar System example of a planet near the outer habitable zone edge, however, Mars has the confounding issue of its small mass, $M_{Mars} = 0.11 \ M_{Earth}$, such that it was not able to maintain surface water due to atmospheric escape. It is important to keep in mind that atmospheric escape, either hydrodynamically or by photolysis, which is the dissociation of H_2O by UV light, followed by hydrogen loss decreases for larger planet distances. So the outer CHZ edge provides minimized harm from stellar winds and flares, but a significant price is paid in the reduced potential for photosynthetic biomass production at distant locations from the host-star.

Paradoxically, the Earth's atmosphere was hotter when life first appeared than it is today, despite the Sun being significantly fainter. This is the "faint young Sun" paradox or problem. The Earth's atmosphere seems to have been warm enough to support life very early on, so the composition of the atmosphere must have played a major role. The point is that the habitable zone limits also depend on the atmospheric composition of the planet in question and so do not form precise distance limits. The faint young Sun problem is solved by introducing the heating effect of greenhouse gasses in the atmosphere and their reduction by the carbon cycle. Remarkably, as the Earth was formed well away from the runaway greenhouse (inner) CHZ limit when the Sun was faint (See Figure 3), Earth's atmosphere contained significant greenhouse gasses, mostly CO_2, keeping early life warm. The slow rise in Solar luminosity was completely compensated by the reduction of CO_2 in the atmosphere by a combination of surface weathering (rock erosion) and photosynthetic life itself for billions of years. This directly permitted the continuous maintenance of life through chemical weathering and photosynthetic cooling enough to compensate for the increase in heating from the slow and steady increase in the luminosity of the Sun. So, since the continuous CHZ as it is usually formulated does not take into account potential changes in the planet's atmosphere, it must be applied with caution. Rather, it is defined only by the greenhouse limits, which as described above require different atmospheric compositions at the inner and outer edges.

Back to our example. Today, the Earth contains much less than 1% CO_2, and is teetering near the inner edge of the CHZ. Plants and the carbon-silicate cycle maintain a temperate climate on Earth. Being at the inner CHZ edge, photosynthetic production is maximized as is photosynthetic cooling by CO_2 reduction. Without plants and rock weathering, the end of the habitability of Earth would have been hastened. To our great benefit, the reduction of greenhouse gasses in the atmosphere of Earth compensated for the slow rise in the luminosity of the Sun. Likewise, if in the future, humans increase greenhouse gasses like CO_2 and NH_3 beyond the limit, the Earth will move out of the Solar habitable zone. Even with human interventions to minimize Earth's greenhouse gasses, the rise in Solar luminosity will eventually result in the evaporation of the oceans and cause the inner CHZ limit to move past the Earth.

2.2 Circumbinary Habitable Zones and Niches

The potential habitability of planets in binary star systems has been investigated [e.g 12,13]. Such planets come in two basic varieties, the p-type (planetary-type) where the planet orbits a close stellar binary system, and the s-type (stellar-type) where the planet orbits one of the stars in a wide binary. Habitable zones may exist in both cases [14, 15]. In the case of the p-type planets, several of which were discovered using the Kepler Telescope, enhanced habitability conditions for habitable zone planets may result from tidal braking of the rotation of one or both of the stars in the binary [14]. Some circumbinary habitable niches were investigated [16] and several Kepler-discovered circumbinary planets were found to have better than Earth habitability conditions [17].

Figure 4 shows two examples of circumbinary habitable niches [16]. On the left, a pair of 1.0 $M_\odot$ stars are shown with a binary orbital period of 15 days and an eccentricity of e = 0.1. On the right, a 1.0 $M_\odot$ primary and a 0.55 $M_\odot$ companion with an orbital period of 12 days and an eccentricity of e = 0.1 is shown. In the top panels of Figure 4, the empirical circumbinary habitable zone, between recent Venus and early Mars limits, is shown as a function of time as the shaded area. The circumbinary continuous CHZ is shown as the darkest region. It can be seen that this is at a rather large distance from the binary and there is a fairly wide continuous CHZ, exceeding 0.6 AU in the Solar twin case and about 0 4 AU in the case shown in the right. The dashed lines in both cases represent the same habitable zone limits in the case of a single star with the mass of the primary, which in both cases shown is 1.0 $M_\odot$.

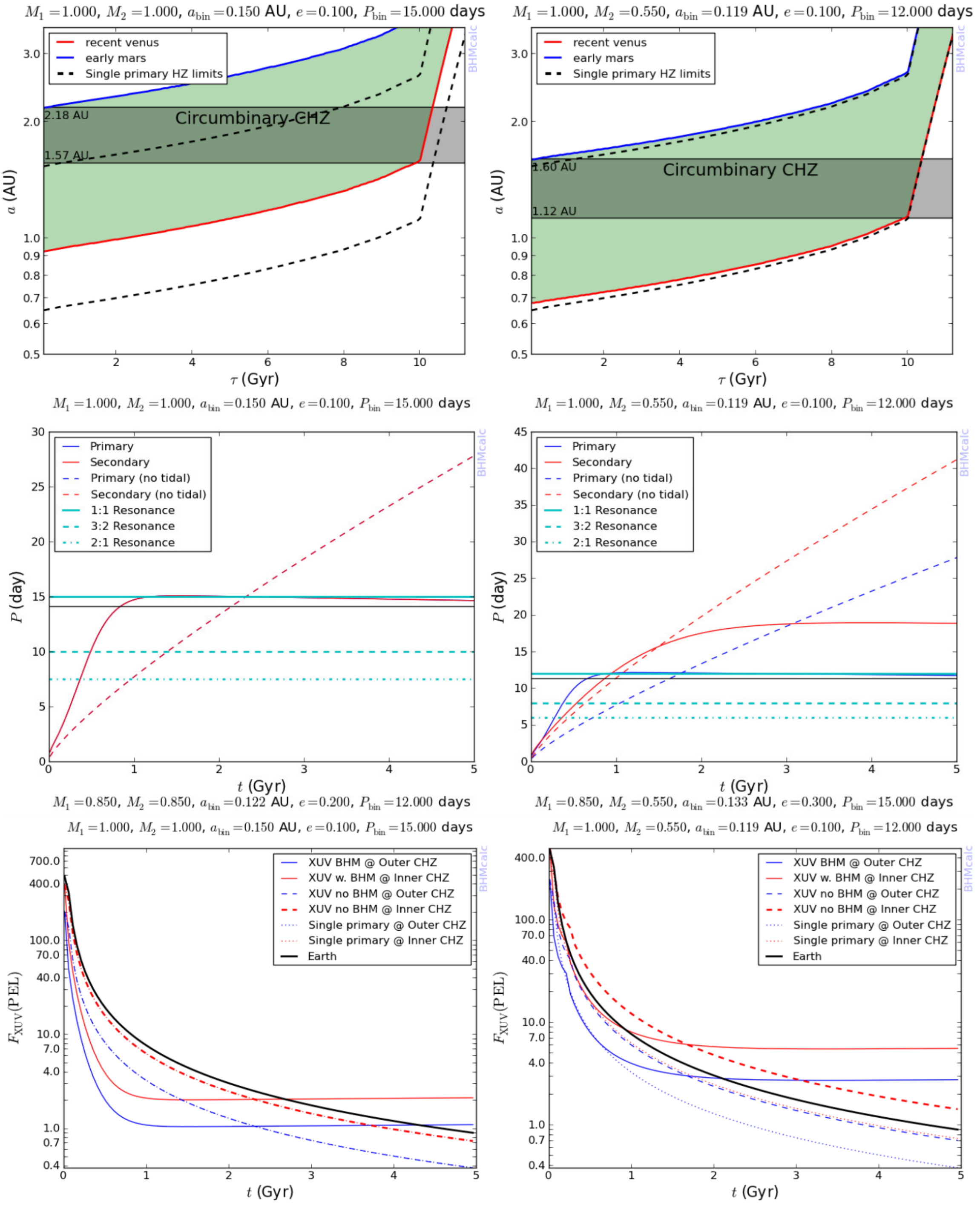

Figure 5: Circumbinary Habitable Zones (top), rotational evolution (middle), and XUV fluxes (bottom) for two Habitable Niches. Details are given in the text. Both binaries experience the Binary Habitability Mechanism (BHM). Credit: Mason et al. [15].

2.2.1 The Binary Habitability Mechanism

The Binary Habitability Mechanism (BHM) is based on the mutual tides between stars [14, 16-19] in a moderately close binary and is illustrated in the second row of Figure 5 panels, which show the evolution of the rotation of both stars. The primary is shown in blue and the secondary in red. In both cases gravitational tidal effects, analogous to the mutual tides between the Earth and Moon, operate to cause both stars to slow their spins and evolve towards longer rotational periods during early phases of evolution ($\sim$1 Gyr) than does the isolated (primary or secondary) star alone. This is especially important in the early phases of planets and the potential formation of life on CHZ planets in the first 1 - 2 Gyr when the majority of host-star winds and magnetic activity occurs. In the Solar twin's case, in the middle left panel of Figure 5, the stars become synchronized with the binary period of 15 days. In the unequal mass case, the primary evolves to become synchronous with the 12-day orbital period, in less than one Gyr and remains so indefinitely. The secondary on the other hand evolves to a significantly longer rotational period of about 19 days.

The way this translates into the habitability of the planet, which is first of all assumed to be in the circumbinary habitable zone is shown in the bottom panels of Figure 4. The combined X-ray plus ultraviolet (XUV) flux incident on a planet at the inner and outer boundaries of the continuously habitable zone, with the tidal effects of BHM included, is shown as solid lines. Dash lines show the inner and outer habitable zone limits of the same binary without considering BHM. The solid black line is the case of Earth. One can readily see that a planet anywhere in the continuous habitable zone experiences significantly reduced XUV flux early, as compared with Earth. It will then maintain a steady level of XUV flux at late times. Potentially, a circumbinary planet can maintain a presence in both the CHZ and the UVHZ, given various combinations of stellar masses, orbital period, and eccentricity. Better than Earth conditions could result from a subset of the wide variety of binary habitable zones and the conditions within. The results of the work on BHM and the habitable niches that they offer suggest that the habitability of a circumbinary planet orbiting whatever primary that happens to exist in the binary may be enhanced or at least retain more H_2O compared to its single star case if the orbital period is between about 10 and 40 days [14, 16-19] (see Figure 4).

The orbital eccentricity plays a major role in the time scale for synchronization to an equilibrium configuration. Often, in this orbital period range, habitability is not hindered by the presence of a companion, unless the eccentricity is very high. In the Solar twin binary case (bottom left panel of Figure 4, we can see that a planet anywhere in the continuous CHZ of the binary will continue to provide one to two times the present level of Earth XUV flux, long after the Earth itself is moved out of the UVHZ. On the other hand, Solar activity continues to decrease even as Solar luminosity increases. With a G2-type main sequence star, like the Sun, the planet cannot continuously maintain its presence in both the continuous habitable zone and the UV habitable zone. In the bottom right case, with the solar mass primary and 0.55 solar mass, the binary continuous habitable zone maintains three to six times the XUV flux of present-day Earth. The constant level of UV flux that is reached asymptotically with time depends mainly on the binary orbital period. The same stars with a 15 or 20-day orbital period, and a suitable eccentricity, will provide a circumbinary CHZ with long-term XUV flux that is a bit lower than that shown in the bottom right panel, matching the Solar level at about 30 days. With binaries, Nature experiments broadly with the same habitability factors that we know maintain life on Earth. Most of the results are likely to be harmful to the long-term formation and maintenance of complex life, while some circumbinary planets will find themselves in optimal conditions.

3. Beyond Circumstellar Habitable Zones

It is clear from the Solar System examples of the Moon and Mars that the term habitable zone is a bit of a misnomer. Shapley's term liquid water belt more correctly describes the concept as used today. Being in the liquid water belt / CHZ is necessary, but not sufficient for being habitable. For one thing, the type of planet in the CHZ is critical to its habitability. While there are further complications to consider such as its interior and atmosphere compositions, a mass range that includes Earth and Venus (M_{Venus} = 0.82 M_{Earth}), but not Mars is adopted. A lower mass limit of roughly 0.50 M_{Earth} is reasonable. Estimating the maximum mass of a habitable planet is more difficult as there is no Solar System object to compare, but a limit of 2-3 M_{Earth} has been proposed to allow for the operation of plate tectonics and to avoid water worlds and mini-Neptune conditions.

In addition to having a mass in the acceptable range, planetary composition is key to forming a habitable planet. Water is especially critical. A potentially habitable planet must either have native water that survived the early erosion of the young host star or have water delivered by wet asteroids or comets. The presence or absence of giant planets in planetary systems is expected to have a profound impact on the balance of early comets to deliver water and late comets causing mass extinction events. Detailed impact history will allow or constrain the development of complex life on rocky planets.

4. The Galactic Habitable Zone (GHZ)

The Galactic Habitable Zone was introduced by Gonzalez et al. [20] and elaborated by others [21, 22, 23] to illustrate that the environment favoring the development of habitable planets, especially the metallicity, decreases as a function of distance from the Galactic center. While, on the other hand, both the star formation rate and the stellar density are higher moving towards the Galactic center. So, while there are likely many more rocky planets in the inner disk of the Galaxy, compared to the outer disk, the threats to life on planets are increased from a nearby supernova, gamma-ray burst, or astrosphere collapse [24] during passage through a giant molecular cloud, or even a close passage by a massive star resulting in a flood of short-period comets from an Oort cloud disruption. All of these threats increase sharply towards the Galactic center. In the outer regions of the Galactic disk, the metallicity is sub-solar, so Earth-like planets are currently rare and most likely young. The star formation rate and stellar density are lower than in the inner Galaxy, so the threats to habitability are few and far between.

So, given these constraints, the GHZ seems to be an annulus, of some thickness, in the disk of the Milky Way, as depicted in Figure 6, and other star-forming disk galaxies. The GHZ has also been studied in other galaxies [25, 26]. Within the GHZ, planets are formed with sufficiently high metallicity to form Earth-like planets but without exposure to a prohibitively high frequency of mass extinctions due to cosmic mass-extinction level events. There is no strong consensus to date concerning the validity or position of the GHZ. Various authors give inner and outer GHZ limits, similar to those shown in Figure 6, while others argue that most or all of the Galaxy is habitable. Just like the CHZ, the GHZ concept is a useful starting point, but not fully satisfactory in predicting the habitability of any particular planet. For example, a planet-hosting star may be on a fairly circular around the Galactic center, like the Sun. Many other stars are on more eccentric orbits. This may have significant consequences as the Galactic center distance is not the fundamental issue. The amount of time a planetary system spends near or crossing an active star formation region within one of the 4 spiral arms of the Milky Way (see Figure 6) is the likely habitability deciding factor, which depends sensitively on the orbital details. The Earth is likely in the Galactic spiral-arm co-rotation region which dynamically traps stars, especially if born within. If so, this has kept the Earth away from the worst of the damage associated with supernovae and gamma-ray bursts. Notice that Figure 6 shows that the Earth is in both the GHZ annulus and a co-rotation region along with the Local Arm. Another "island of Galactic habitability" likely exists on the other side of the Galaxy.

The GHZ concept is analogous to the CHZ as it is an annular region lying in the plane of the Galactic disk with two opposing habitability factors. Proto-planetary disks must possess elements, especially iron, necessary for terrestrial planet formation which favors the inner part of the Galactic disk. The inner disk of the Galaxy builds up metal content early and rapidly because of its high star formation rate, while the more distant regions, with low star formation rates, even now, remain deficient in the elements required for terrestrial planets [25]. However, in the inner regions of the Galactic disk, high star formation rates mean high supernova and gamma-ray burst rates, potentially causing mass extinctions, as well as correspondingly high stellar densities, as star formation in the inner disk has gone on for a very long time. We know that asteroid and comet impacts have had a dramatic effect on life, including the extinction of the dinosaurs only 65 Myr ago. A higher stellar density increases the risk of frequent encounters with passing stars and free-floating (rogue) planets, resulting in Oort cloud disruption or worse. Thus, the requirement of billions of years without too many or too devastating mass extinctions is of the essence of the GHZ. The GHZ is inherently time-dependent, with habitability moving outward over time, which is also analogous to the CHZ.

5. The Super-Galactic Habitable Zone (SGHZ)

In our review [27], we introduced the Super-Galactic Habitable Zone (SGHZ) to take into account regions in the local Universe and beyond that are well-suited for the development of complex life [28]. Namely, regions with an availability of the chemical elements essential for life, and the presence of water in liquid form on the surface of planets are within the SGHZ. However, the SGHZ excludes entire galaxies with accreting supermassive black holes and galaxies that undergo frequent mergers, as well as those that are deficient in the elements necessary to build habitable planets.

The SGHZ is an extension of the GHZ with the recognition that many entire galaxies will be uninhabitable. Similar to the CHZ and the GHZ, the SGHZ splits the Universe into three regions, with a Goldilocks region between two extremes. The Local Super Cluster of galaxies has the massive Virgo Cluster of galaxies at its center. Galaxies are merging frequently in such rich clusters. Planets in rich clusters and superclusters will likely be exposed to high frequencies of mass extinction and sterilizing events such as gamma-ray bursts, supernovae, and comet impacts. These high-density regions of the Universe are inhospitable to life, especially complex life (Figures 6, 7. and 8). Beyond the other end of the SGHZ, regions of the Universe that do not have sufficient

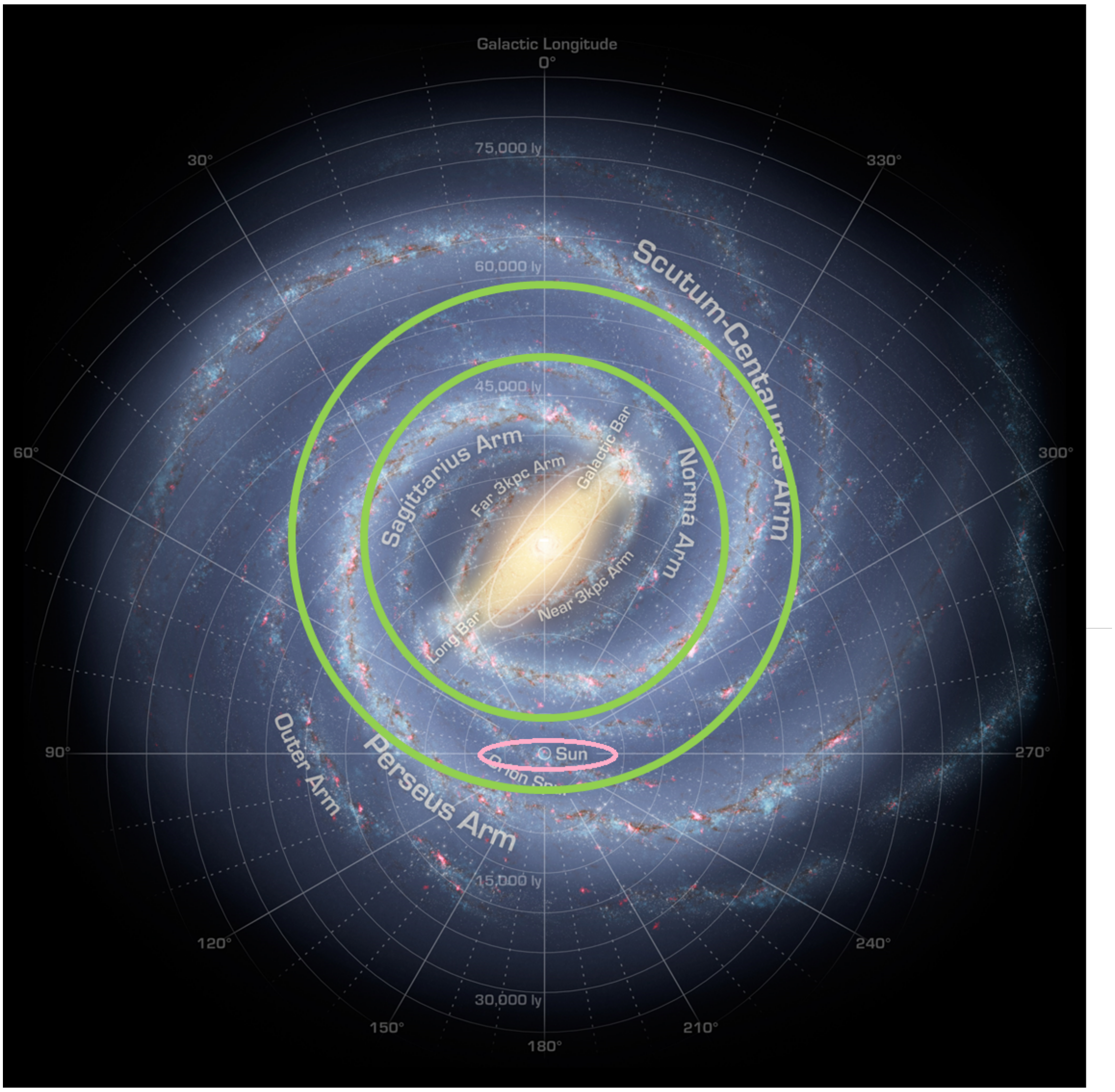

Figure 6: The Galactic Habitable Zone and the Local Co-rotation Region. The spiral arms are the location of current star formation in the Milky Way and the spiral-arm structure orbits the Galactic center with a constant pattern velocity, such that its spiral structure remains largely intact. Stars that are not in the spiral arms orbit the Galactic center following the rotation curve of the Galaxy. A tiny fraction of stars, like the Sun, with orbits that allow them to stay within special regions in co-rotation with the spiral arm structure, like the Local Co-rotation Region depicted as the region within the pink oval, are "islands of habitability" for complex life. Adapted from: Image Credit: NASA.

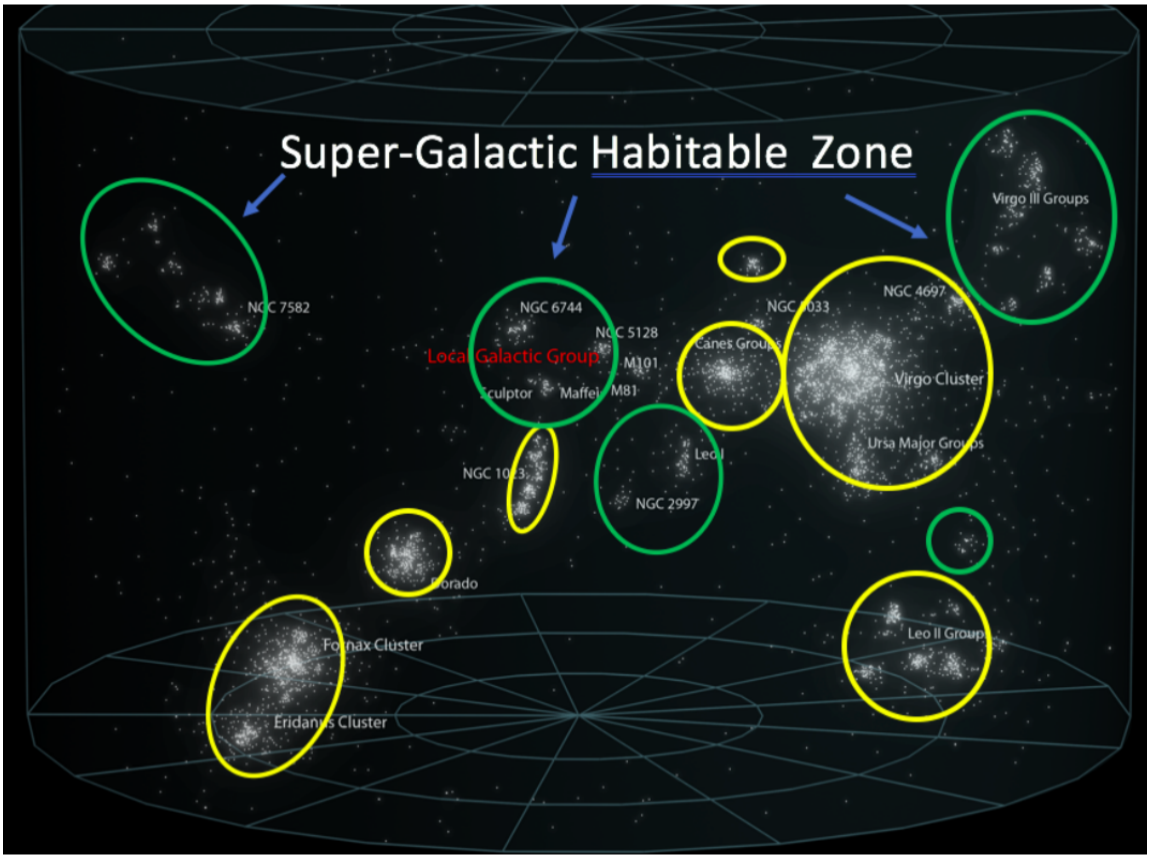

Figure 7: The Local Super-Galactic Habitable Zone (SGHZ) is shown. The highest probability for the existence of complex life in the Local Supercluster of galaxies is in regions of intermediate density, as shown within the green bubbles. The higher-density galaxy clusters, shown within the yellow bubbles, are inhospitable regions of high star-formation rates and frequent galaxy and black hole mergers. The vast low-density (and low metallicity) volume outside of all of the ovals is unlikely to form many Earth-like planets and so is beyond the limits of the SGHZ. Credit: Figure, adapted from an image of Andrew Z. Colvin, See Mason and Biermann [25] Figure 7.

density, and hence star formation rates, like dwarf galaxies, to have produced the chemical elements necessary to build rocky planets, with iron cores, oceans, and so on are excluded from the SGZH, see Figures 7 and 8. This includes vast regions between major superclusters containing many dwarf galaxies and inter-galactic space. Some galaxy groups, like the Local Group of galaxies and many others, are in the SGHZ, as they are situated far from rich galaxy clusters and superclusters. For specific examples, the aforementioned Virgo Cluster, containing M87 and other giant elliptical galaxies undergoing frequent mergers are not in the SGHZ. Even closer galaxies might also be outside the SGHZ, like M82 which is a starburst galaxy, with accreting black holes and precessing jets, deadly for life on Earth-like planets [29 and references therein].

Unlike the circumstellar habitable zone (CHZ) and the Galactic habitable zone (GHZ) which have been described as regions in the plane of the solar system and the Galaxy respectively, the Super-Galactic Habitable Zone (SGHZ) is a complex volume in three-dimensional space containing galaxies and parts of galaxies that are amenable to complex life as we know it based on the same

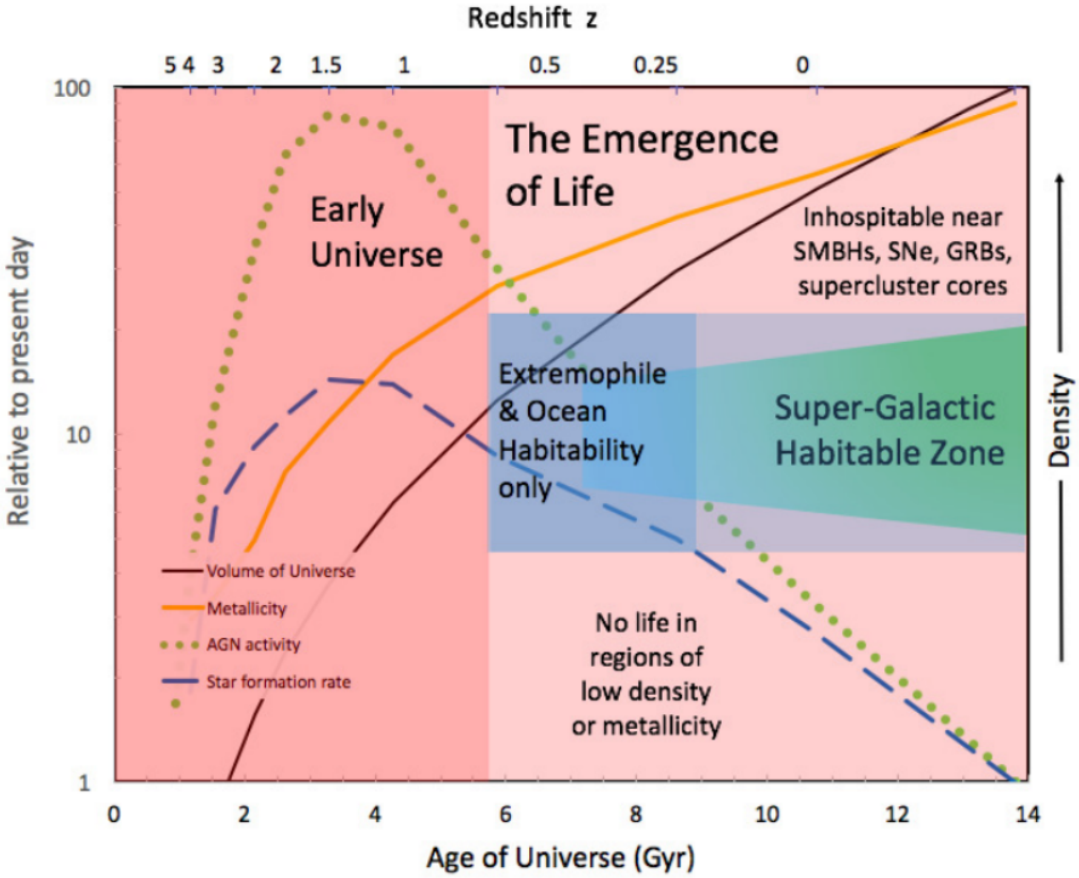

Figure 8: A schematic depiction of the emergence of life in the Universe and the Super Galactic Habitable Zone is shown. The time evolution curves of star formation rate (SFR) (a proxy for supernovae (SNe) and gamma-ray burst (GRB) rates), Active galactic nuclei (AGN) activity, metallicity, and volume of the Universe are shown. The elements essential for life slowly build up with time, while habitability threats decrease. On the right side, there is a schematic view of the development of a Super-Galactic Habitable Zone (SGHZ). Life is compromised near the center of rich superclusters and merging galaxies as depicted in Figure 7. Earth-like planets cannot form in low-density regions of the Universe. Far from dense galaxy concentrations habitability is possible in regions with sufficient metallicity. The shaded region around 8 Gyr and extending to later times in certain areas correspond to extremophile habitability without complex life on land as we know it. Figure credit: Mason and Biermann [25].

principles discussed previously concerning the GHZ. The topology of the SGHZ is fairly complex. From far away, the SGHZ might look somewhat like simulations of the large-scale structure of the Universe, with interconnecting filaments surrounded by voids. Closer in, the topology is more like Swiss cheese containing holes with inhospitable regions. The SGHZ - if defined just for survival and not including the birth of complex life - is a large network of connected local habitable zones. The birth places for complex life form a smaller disconnected set of islands. Just like the CHZ and the GHZ, the SGHZ changes dramatically with time. For further discussion of the SGHZ, see our work on challenges facing complex life in the universe [29], a summary of the benefits of living on Earth [30], the development of intelligent life "From Big Bang to Big Brains" [28], as well as our review [27].

References

[1] E. W. Maunder, "Are the planets Inhabited?", (London: Harper and Brothers) (1913).

[2] S. Huang, American Scientist, 47, 3 (1953).

[3] M. Hart, Icarus, 37, 351 (1979).

[4] J. C. G. Walker, P. B. Hays, and J. F. Kasting, Journal of Geophysical Research, 86, 9776 (1981).

[5] J.F. Kasting, D.P. Whitmire, and R.T. Reynolds, Icarus, 101, 108 (1993).

[6] R. Kopparapu et al., Astrophysical Journal, 765, 131 (2013).

[7] R.M. Ramirez, L. Kaltenegger, Astrophysical Journal. Lett. 837, L4 (2017).

[8] J. N. Bahcall et al., Astrophysical Journal, 555, 990 (2001).

[9] A.P. Buccino, G.A. Lemarchand, and P.J.D. Mauas, Icarus, 192, 582 (2007).

[10] R. Spinelli et al, MNRAS, 522, 1411, 2023.

[11] E.W. Schwieterman, et al. Astrophysical Journal, 878, 19 (2019).

[12] P.A., Mason, J.M. Clark,. 2012, BAAS, 220, 52504. (2012)

[13] I.I. Shevchenko, Astronomical Journal, 153, 273 (2017).

[14] P. A. Mason, J.I. Zuluaga, J.M. Clark, and P.A. Cuartas-Restrepo, 2013, Astrophysical Journal (Letters), 774, 26 (2013).

[15] Z. Eggl, et al., Monthly Notices of the Royal Astronomical Society, 428, 3104 (2013).

[16] P.A. Mason, et al., Int J Astrobiology, 14, 391 (2015).

[17] J,I. Zuluaga, P. A. Mason, and P. A. Cuartas-Restrepo, ApJ, 818, 160 (2016).

[18] P. A. Mason, J. I. Zuluaga, and P. A. Cuartas-Restrepo, in Frontier Research in Astrophysics, Proceedings of Science, 237, 45. DOI: https://doi.org/10.22323/1.237.0045 (2016).

[19] C.P. Johnstone, et al, Astronomy and Astrophysics, 626, 22 (2019).

[20] G. Gonzalez, D. Brownlee, and P. Ward, Icarus. 152, 185-200 (2001).

[21] C.H. Lineweaver, Icarus, 151, 307-313 (2001).

[22] N. Prantzos, The concept of “galactic habitable zone”. In: Gargaud M, Lopez-García P, Martín H (eds) Origins and evolution of life: an astrobiological perspective. Cambridge University Press, Cambridge, pp 154–166. (2011)

[23] M.G. Gowanlock, D.R. Patton, and S.M. McConnell, Astrobiology 11, 855–873 (2011).

[24] J. Scalo, J.C. Wheeler, Astrophysical Journal, 566, 723 (2002).

[25] M. Sundin, Int J Astrobiology 5, 325 (2006).

[26] F. Suthar, C.P. McKay, Int J Astrobiology 11, 157 (2012).

[27] P.A. Mason, P.L. Biermann, "Astrophysical and Cosmological Constraints on Life", In: Gordon, R., Sharov, A.A. (Eds.), Habitability of the Universe before Earth. Elsevier B.V, Amsterdam, 89-126 (2017).

[28] P. A. Mason, Proc. "From Big Bang to Big Brains", in Frontier Research in Astrophysics III, (Proceedings of Science). DOI: 10.22323/1.331.0090 (2019).

[29] P. A. Mason, P.L. Biermann, "Challenges for Life in the Local Universe", in The Golden Age of Cataclysmic Variables and Related Objects V, (Proceedings of Science). DOI: 10.22323/1.368.0059 (2021).

[30] P. A. Mason, P.L. Biermann, "Benefits of Living on Earth", in Multifrequency Behaviour of High Energy Cosmic Sources - XIII, (Proceedings of Science), DOI: 10.22323/1.362.0076 (2020).

[31] S.H. Dole, Habitable Planets for Man, Blaisdell Publishing Company. p. 103.

[32] M.I. Budyko, Tellus. 21 611. (1969).

[33] D. Rasool, C. De Bergh, Nature, 226, 1037 (1970).

[34] M. J. Fogg, Journal of the British Interplanetary Society. 45, 3 (1992).

[35] D.S Spiegel et al., The Astrophysical Journal. 721, 1308 (2010).

[36] Y. Abe et al., Astrobiology. 11 443 (2011).

[37] R. Pierrehumbert, E. Gaidos, The Astrophysical Journal Letters. 734, L13 (2011).

[38] J. Lecante et al., Nature. 504, 268 (2013).

[39] G. Vladilo et al.The Astrophysical Journal. 65 (2013).

[40] A. Zsom et al., Astrophysical Journal, 778, 109 (2013).

[41] I. Gomez-Leal et al., Icarus. 321, 608 (2019).

[42] B.P. Levenson, Planetary and Space Science, 206, 105318 (2021).

6. Appendix A: Circumstellar Habitable Zone Limits

Year	**Inner Limit** (AU)	**Outer Limit** (AU)	**Reference and Notes**
1964	0.725	1.24	Dole[31]: optically thin atmospheres, fixed albedo
1969		1.005–1.008	Budyko[32]: ice albedo feedback global glaciation
1970	0.92–0.96		Rasool, De Bergh[33]: min distance for stable oceans
1979	0.958	1.004	Hart[3]: evolution of Earth's atmospheric composition and surface temperature.
1992		3.0	Fogg[34]: included the carbon cycle
1993	0.95	1.37	Kasting et al.[5]: greenhouse gases CO_2 and H_2O cloud cooling, carbonate–silicate expands CHZ width Their optimistic limits are 0.84–1.67 AU.
2010		2.0	Spiegel et al.[35]: seasonal surface water for high obliquity and high eccentricity orbits
2011	0.75		Abe et al.[36]: desert planets with water at the poles
2011		10	Pierrehumbert and Gaidos[37]: with thousands of bars of H_2
2013	0.99	1.67	Kopparapu et al.[4]: updated moist greenhouse water loss Their optimistic limits are 0.97–1.67 AU.
2013	0.95		Leconte et al.[38]: 3-D atmosphere models
2013	0.77–0.87	1.02–1.18	Vladilo et al.[39] wider CHZ for higher pressures narrow for pressure down to 15 mbar
2013	0.38		Zsom et al.[40]: many possible atmospheric composition combinations pressure and relative humidity
2017	0.95	2.4	Ramirez and Kaltenegger[7]: volcanic habitable zone atmospheric concentration of 50% H_2
2019	0.93–0.91		Gomez-Leal et al. [41]: Earth analog moist greenhouse limit with and without ozone, global climate model
2019		1.14	Schweiterman et al. [11]: Habitable zone for complex life, based on $CO_2 < 0.01$ bar
2021	0.982		Levenson [42] based on Earth's climate across geological time