The Goldilocks Planet?
How Silicate Weathering Maintains Earth “Just Right”

James F. Kasting

Earth’s climate is buffered over long timescales by a negative feedback between atmospheric CO$_2$ level and surface temperature. The rate of silicate weathering slows as the climate cools, causing CO$_2$ to increase and warming the surface through the greenhouse effect. This buffering system has kept liquid water stable at Earth’s surface, except perhaps during certain ‘Snowball Earth’ episodes at the beginning and end of the Proterozoic. A similar stabilizing feedback is predicted to occur on rocky planets orbiting other stars if they share analogous properties with Earth, most importantly an adequate (but not overly large) abundance of water and a mechanism for recycling carbonate rocks into CO$_2$. Periodic oscillations between globally glaciated and ice-free climates may occur on planets with weak stellar insolation and/or slow volcanic outgassing rates. Most silicate weathering is thought to occur on the continents today, but seafloor weathering (and reverse weathering) may have been equally important earlier in Earth’s history.

INTRODUCTION

Earth has been continuously inhabited throughout most of its history, starting from at least 3.5 Ga (Catling and Kasting 2017). Part of the evidence for life’s antiquity comes from stromatolites, the remains of photosynthetic, mat-forming organisms. These early photosynthesizers probably did not produce O$_2$; however, like modern cyanobacteria and algae, they required surface (or near surface) liquid-water environments. But surface liquid water is not present on other planets within our solar system, and it might not have been present on Earth, either, had our planet not been endowed with a mechanism for stabilizing its climate. Here, I review how that mechanism works on Earth and what it might mean for the possibility of finding Earth-like planets, and life, around other nearby stars.

THE FAINT YOUNG SUN PROBLEM

Astronomers have known since the 1950s that main sequence (hydrogen burning) stars like the Sun brighten as they age. This brightening occurs because the conversion of hydrogen to helium in the star’s core causes it to become denser and hotter, thereby accelerating the rate of nuclear fusion. The Sun itself is thought to have been some 30% dimmer early in its history, and it has been increasing in luminosity steadily since that time (Gough 1981). This poses a problem for the early climates of the terrestrial planets (Sagan and Mullen 1972). Both Earth and Mars appear to have been warm during their early histories, despite the lower solar luminosity. Therefore, both planets must have had a bigger greenhouse effect back in the past to prevent their surfaces from freezing. How warm early Mars was is still debated, but most geologists agree that the early Earth was generally above freezing. The climate record from those times is poor, but glaciation was rare during the Archean Eon (Catling and Kasting 2017), and there is no good reason for believing that the preceding Hadean Eon was any colder.

Sagan and Mullen (1972) favored higher concentrations of ammonia and methane as the solution to the faint young Sun problem. Ammonia, in particular, is a strong greenhouse gas that absorbs radiation across the thermal-infrared spectrum (as does H$_2$O). But photochemical modeling has indicated that ammonia was photochemically unstable and, thus, unlikely to have been abundant. Methane by itself is not as strong a greenhouse gas as is commonly thought, plus it photolyzes to produce an organic haze which would have cooled the climate, as it does on Saturn’s moon Titan. However, methane in Earth’s atmosphere today is a potent greenhouse gas on a molecule-per-molecule basis because it is present in low concentrations and because it absorbs radiation in the atmospheric ‘window’ region, between 8 and 12 µm, where absorption by CO$_2$ and H$_2$O is weak.

A better candidate for producing a strong greenhouse effect on the early Earth is CO$_2$. This was recognized by Hart (1978) in a model of the circumstellar habitable zone—a concept discussed below. Carbon dioxide, along with H$_2$O, is one of two greenhouse gases that are responsible for most of Earth’s present 33°C greenhouse effect. Detailed climate model calculations showed that CO$_2$ concentrations of a few tenths of a bar, ~1,000 times the present level, could have kept the early Earth from freezing (Owen et al. 1979). But were such high CO$_2$ concentrations to have been expected? To answer this question, we must think about the processes that control atmospheric CO$_2$.

CONTROLS ON ATMOSPHERIC CO$_2$

Atmospheric CO$_2$ is regulated by the geochemical carbon cycle, which consists of two main parts: the organic carbon cycle and the carbonate–silicate cycle.

Keywords: carbonate–silicate cycle, silicate weathering, habitable zone, climate limit cycling, Snowball Earth, seafloor weathering, reverse weathering
The Organic Carbon Cycle

The organic carbon cycle begins with oxygenic photosynthesis and converts CO$_2$ and H$_2$O into organic matter, CH$_2$O, and O$_2$. This process is reversed by respiration, carried out by both animals and plants, and by aerobic decay, performed by bacteria. This cycle is rapid enough to recycle all the atmospheric CO$_2$ in ~10 years. But it is also in approximate balance, except for a small leak of organic matter into marine sediments. Burial of this organic matter forms a net source for O$_2$ and a net sink for CO$_2$. The O$_2$ that is generated eventually reacts with fossil organic matter on the continents in a process termed oxidative weathering, and this closes off the long-term cycle.

British scientist James Lovelock suggested that the organic carbon cycle kept atmospheric CO$_2$ concentrations high on the early Earth (Lovelock 1979). He dubbed his theory the Gaia hypothesis, named after the Greek goddess of Mother Earth. In Lovelock’s view, photosynthetic organisms pull CO$_2$ out of the atmosphere at just the right rate to compensate for gradually increasing solar luminosity. But this is unlikely to have been the case. The organic carbon cycle, which controls atmospheric O$_2$, cannot also be the primary control on atmospheric CO$_2$ because the feedbacks which regulate O$_2$ are quite different from those that affect climate. O$_2$ to say this another way, in order for atmospheric O$_2$ to remain constant, organic carbon burial must be balanced by organic carbon weathering, so the net effect on atmospheric CO$_2$ should be zero over long timescales. To be fair, later versions of the Gaia hypothesis included the effect of land plants on silicate weathering, which is part of the inorganic carbon cycle described below. Thus, most researchers would agree that life influences atmospheric CO$_2$ levels, probably keeping them somewhat lower than they would be in its absence. But the fundamental control mechanism lies elsewhere (see Porder 2019 this issue).

The Carbonate–Silicate Cycle

Today, atmospheric CO$_2$ is (probably) primarily controlled by the inorganic carbon cycle, also known as the carbonate–silicate cycle. It is inorganic because CO$_2$ is not reduced to organic carbon, and so no O$_2$ is produced. This cycle is depicted in Figure 1.

We begin with the weathering part of the cycle. Atmospheric CO$_2$ dissolves in rainwater to produce carbonic acid (H$_2$CO$_3$). Carbonic acid is a weak acid but it is strong enough over long timescales to dissolve continental silicate rocks by the silicate weathering process. Carbonate rocks dissolve as well—indeed, they dissolve faster than do silicates—but this has little effect on long-term CO$_2$ concentrations because no net carbon exchange takes place between the atmosphere–ocean system and sediments. On short timescales, weathering of carbonates transfers CO$_2$ from the atmosphere to the oceans, but that added carbon is eventually removed by carbonate precipitation. The by-products of silicate weathering include calcium and magnesium ions (Ca$^{2+}$ and Mg$^{2+}$), bicarbonate ions (HCO$_3^-$), and dissolved silica (SiO$_2$). These dissolved products are carried by streams and rivers down to the ocean where various organisms use them to make shells of calcium carbonate (CaCO$_3$) or silica. Today, much of this carbonate precipitation is carried out by organisms that live in the surface ocean, such as the planktonic foraminifera. During the Precambrian, this function was performed primarily by benthic, mat-forming organisms, creating stromatolites. But carbonate would precipitate anyway, even on an abiotic planet, as the products of weathering—specifically, alkalinity $\equiv [\text{HCO}_3^-]+2[\text{CO}_3^{2-}]$—accumulated in the ocean.

When organisms such as foraminifera die, they sink into the deep ocean. The deep ocean is slightly more acidic than the surface ocean, and so most of the carbonate redissolves. A portion of it is preserved, however, and forms carbonate sediments that coat parts of the seafloor. When this seafloor is subducted, some of the carbonate is scraped off, but some of it is carried down to great depths. There, the heat and pressure cause calcium and magnesium to recombine with silica (which by this time is the mineral quartz), reforming Ca/Mg silicates and releasing gaseous CO$_2$. This CO$_2$ is restored to the atmosphere by volcanism.

Ignoring Mg, the entire cycle can be represented by the reaction

$$\text{CaSiO}_3 + \text{CO}_2 \leftrightarrow \text{CaCO}_3 + \text{SiO}_2$$  \hspace{1cm} (1)

Reaction (1) running to the right represents silicate weathering plus carbonate deposition. Running to the left, it represents carbonate metamorphism, also referred to as silicate reconstitution. Written as is, with a double arrow, this reaction is sometimes termed the Urey equilibrium. The
famous geochemist Harold Urey proposed in 1952 that this equilibrium determined the concentration of CO₂ in the atmospheres of Earth, Venus, and Mars. Now, scientists consider that kinetics, rather than thermodynamics, determines where the steady state lies.

The actual carbonate–silicate cycle is more complicated because some silicate rocks weather faster than others and because volcanism is not confined to subduction zones beneath continents. For example, an ophiolite—a piece of seafloor that has been obducted, or thrust, onto the edge of a continent—contains mafic and ultramafic rocks that weather rapidly compared to more felsic rocks. Outgassing of CO₂ occurs at oceanic arcs, midocean ridges, and hotspots like Hawai’i. Some of these processes play important roles in modern models of the feedback between weathering and climate, described below. Carbon dioxide can also be recycled by metamorphism, a process that occurs when rocks are heated, but not melted. This can happen, for example, during orogenies created by the collision of continental plates.

The carbonate–silicate cycle is slow compared to the organic carbon cycle—so slow that one can consider the entire atmosphere + ocean + biosphere to be one big reservoir for CO₂. The carbonate–silicate cycle recycles all the organic carbon cycle—so slow that one can consider the atmosphere + ocean + biosphere to be one big reservoir for CO₂. The carbonate–silicate cycle recycles all the CO₂ in this combined reservoir on a timescale of about half a million years. This is too slow to balance fossil fuel emissions and to save us from global warming, but it is fast compared to the timescale for solar evolution. The Sun is currently brightening by about 1% every hundred million years (Gough 1981). On this very long timescale, the carbonate–silicate cycle should remain in approximate steady state.

The Negative Feedback Loop

If the oceans had ever frozen over, silicate weathering should have slowed, or ceased entirely, and volcanic CO₂ should have accumulated in the atmosphere. Eventually, the greenhouse effect of the CO₂ would have become very large, melting the ice, and restoring balance in the carbonate–silicate cycle (Walker et al. 1981). But the oceans do not have to actually freeze for the feedback to operate because the silicate weathering rate should slow as the climate becomes colder. This negative feedback could, in the late 1970s, be parameterized using data from laboratory experiments on silicate dissolution and from 3-D climate models that were then just beginning to produce credible results.

Since the late 1970s, the crude parameterizations of the carbonate–silicate cycle have been refined numerous times (e.g., Berner 2006 and references therein). Berner’s models included additional chemistry and accounted, to first order, for changes in land area and seafloor spreading rates. The temperature dependence of silicate weathering was estimated from measured concentrations of bicarbonate and other ions in rivers at different latitudes. Berner showed that the pCO₂ dependence of the silicate weathering rate could vary from strong to weak, depending on the exact mineral being dissolved and on the complex influence of land plants. But these models always assumed that silicate weathering rates were positively linked to surface temperature, ensuring negative feedback between atmospheric CO₂ and climate.

Not all researchers agreed that this was the case. Maureen Raymo and William Ruddiman proposed that the uplift of the Himalayas and Tibetan Plateau starting at about 40 Ma, and, when combined with heavy monsoonal rainfall induced by the plateau, caused a large speedup in silicate weathering that drew down atmospheric CO₂ and led to Cenozoic cooling (Raymo and Ruddiman 1992). They argued, based partly on Sr isotope ratios, that the silicate weathering rate increased by as much as 40% in response to this forcing, thereby overwhelming any negative feedback from declining surface temperatures. The CO₂ needed to balance this weathering rate could have been “borrowed” from the organic carbon cycle. But Kump and Arthur (1997) pointed out that other geochemical indicators, such as carbon isotopes, did not support this explanation. Furthermore, because the carbonate–silicate cycle is relatively fast, only a few percent increase in global weathering rates is needed to account for CO₂ drawdown over the tens of millions of years of the Cenozoic (Kump and Arthur 1997; Frings 2019 this issue). Instead, Kump and Arthur proposed that the Himalayan uplift increased surface weatherability, thereby strengthening the link between CO₂ and climate.

This point of view is largely borne out by more recent treatments of silicate weathering, although the details vary from one model to another. Maher and Chamberlain (2014) showed that elevated topography generally increases weatherability, but also that new mountain belts are more weatherable than old ones because more fresh rock is exposed. Caves et al. (2016) argued that Cenozoic cooling has been caused by increased exposure of reactive lithologies caused by uplift and glaciation as that era progressed. If so, there may be positive feedback loops hidden within the weathering/climate system, because glaciation is enhanced by global cooling. More recently, Macdonald et al. (2019) suggested that the main control on surface weatherability, and hence on climate, comes from tectonic events that occur in the tropics, as in modern-day Indonesia. Low-latitude arc-continent collisions expose highly reactive mafic and ultramafic rocks that are rapidly weathered due to high rainfall and warm temperatures. Macdonald et al. (2019) suggested that all three major episodes of Phanerozoic glaciation show evidence of having been triggered by this mechanism.

Clearly, the details of plate tectonics and the interaction of the surface features that it creates with climate and rainfall may be positive feedback loops hidden within the weathering/climate system, because glaciation is enhanced by global cooling. More recently, Macdonald et al. (2019) suggested that the main control on surface weatherability, and hence on climate, comes from tectonic events that occur in the tropics, as in modern-day Indonesia. Low-latitude arc-continent collisions expose highly reactive mafic and ultramafic rocks that are rapidly weathered due to high rainfall and warm temperatures. Macdonald et al. (2019) suggested that all three major episodes of Phanerozoic glaciation show evidence of having been triggered by this mechanism.

THE HABITABLE ZONE AROUND THE SUN AND OTHER STARS

In addition to helping solve the faint young Sun problem, the CO₂–climate feedback loop is also fundamental to estimating the width of the habitable zone (HZ) around the Sun and other stars. The habitable zone—also known as the goldilocks zone—is the region around a star within which a planet can maintain liquid water on its surface. That allows abundant photosynthetic life to be (potentially) present on the planet’s surface where it can modify the planet’s atmosphere in a way that might be detected remotely. Hart (1978) discussed calculations from a computer model that he had constructed to study the coupled evolution of the Earth’s atmosphere and climate. He concluded that the continuously habitable zone (CHZ) around the Sun—the region that remains habitable throughout the 4.6 billion years of solar system history—was quite narrow, only 0.06 astronomical units (AU; one AU is the mean Earth–Sun distance.) By comparison, the mean spacing between the terrestrial planets in our own solar system is about 0.3 AU. Hart (1979) later concluded that the CHZ around other main sequence stars was even narrower, or in some cases nonexistent.
Why did Hart get such a pessimistic result? His computer model did include a crude carbonate–silicate cycle, which was based on departures from the Urey equilibrium (eqn. 1). However, Hart’s computer model did not incorporate a physically based loss process for CO₂, so the feedback between CO₂ and climate was relatively weak. Furthermore, his estimate for the magnitude of the CO₂ greenhouse effect was too small, and so a planet that became globally glaciated in his model could never thaw out. That kept the outer edge of the HZ relatively close in. Assuming a strong CO₂ feedback, and using a better climate model, Kasting et al. (1993) updated Hart’s calculations (Fig. 2).

**Stagnant Lid Planets**

The theory of the habitable zone depends on the negative feedback between CO₂ and surface temperature. But on Earth that feedback loop relies heavily on plate tectonics—a process that may or may not occur on exoplanets. A rocky planet that lacks plate tectonics is sometimes referred to as a *stagnant lid* planet. Would such a planet still experience a stabilizing negative feedback?

Foley (2015) found that some stagnant lid planets may successfully transition to a plate tectonic regime in which the climate is stable. However, other planets may find themselves stuck in a *supply limited* weathering regime, in which the rate of silicate weathering is controlled by the supply of fresh rock at the surface instead of by temperature and rainfall. On such a world, the negative feedback between CO₂ and climate would vanish, allowing it to evolve into a hot, dry, Venus-like condition. Just because a planet is rocky and is within the habitable zone of its parent star does not guarantee that it will be habitable. The nature of the parent star matters, too. Dim, red, M stars are poor hosts for habitable planets for a variety of reasons, including their flare activity and active stellar winds, which could strip off planetary atmospheres.

**Waterworlds**

Models of planetary accretion predict that some exoplanets might form with much more water than is present on Earth, leading to oceans that are tens, or even hundreds, of kilometers deep. Such hypothetical planets are termed *waterworlds*. On such a planet, continents would not be subaerially exposed, and silicate weathering could only occur on the seafloor. Kite and Ford (2018) suggested that seafloor weathering is unlikely to occur on such worlds because of the extremely high pressures at the bottom of the oceans. Nonetheless, atmospheric CO₂ in their models are coincidentally in the range 0.2–20 bar for which the habitable zone is as wide as possible. If this idea is correct, then such planets could conceivably remain habitable for billions of years, even without a stabilizing negative feedback between CO₂ and surface temperature.

Silicate weathering does occur on Earth’s seafloor today, and some authors have argued that it could have been even more important in the distant past when the deep oceans were much warmer (see last section).
Climate Limit Cycling

Even planets that have plate tectonics and Earth-like water endowments might not have stable climates. If a planet’s volcanic outgassing rate is too low, or if the amount of starlight it receives is significantly less than that received by Earth today, then the planet’s climate can oscillate periodically between global glaciation and warm, ice-free conditions (Tajika 2007) (Fig. 3). This happens partly because of ice–albedo feedback: climatic cooling causes the polar ice caps to grow, increasing the planet’s albedo and making it still colder. This behavior can be simulated with an energy-balance climate model that calculates the waxing and waning of the polar ice caps. This phenomenon may be important for planets near the outer edge of their parent star’s habitable zone (Menou 2015). It probably did not happen on Earth (Haqq-Misra et al. 2016), but it may have occurred on Mars during the early part of that planet’s history (Batalkha et al. 2016). Additional greenhouse warming from H2 is required in this case because early Mars was outside of the conventional CO2–H2O habitable zone.

Snowball Earth Episodes

Although Earth probably never experienced climate limit cycles, it may well have experienced global glaciation. The existence of low-latitude glacial deposits during the late Proterozoic strongly suggests that the oceans froze over several times. The idea of global glaciation remained speculative until Hoffman et al. (1998) showed that certain Neoproterozoic glacial diamictites were topped by cap carbonates, deposited in the aftermath of the glaciations. These authors were the first to point out that the existence of cap carbonates was entirely consistent with the buildup of atmospheric CO2 during the glaciations, followed by rapid weathering of both silicate and carbonate rocks on the continents once the glaciers melted. This implies that Earth could have entered a global glaciation phase, without experiencing full climate limit cycling, if that glaciation was triggered by some process other than the carbonate–silicate cycle, e.g., a short-lived CO2 drawdown caused by organic carbon burial (Hoffman et al. 1998) or, alternatively, a rapid decrease in methane or some other greenhouse gas.

A remaining mystery about the Snowball Earth glaciations is how photosynthetic life managed to live through them. One hypothesis is that open water continued to exist in a relatively narrow band centered on the equator (Abbot et al. 2011). This happens, according to some climate models, because the sea ice that forms at low latitudes is snow-free and not as highly reflective as the ice that forms closer to the poles. We are fortunate that such Snowball Earth episodes were confined to the Precambrian, as they would have been devastating to the higher plants and animals that evolved during the Phanerozoic.

Recent Advances in Our Knowledge of Silicate Weathering

Seafloor Weathering and Continental Weatherability

Most models of the carbonate–silicate cycle have focused on continental weathering as the loss process for atmospheric CO2. But silicate weathering can also occur on the seafloor, particularly in the vicinity of hydrothermal vents, where warm seawater percolates through the highly fractured young seafloor. Calcium ions are leached out of silicate minerals in the rock, and veins of calcium carbonate are deposited within the fractures. This process is thought to play a minor role (<10%) in CO2 removal today, and it was initially assumed to be independent of ocean temperature (because it depended on the temperature of vent fluids). However, most of the seafloor weathering occurs within the flanks of the ridges, and the weathering rate, therefore, does depend on the deep ocean temperature (e.g., Coogan and Gillis 2013). This process could, thus, have been significantly faster during intervals such as the Mesozoic Era when both the poles and the deep ocean were much warmer. Deep-ocean temperatures are linked to polar temperatures because most deep water forms at high latitudes.

A model of the carbonate–silicate cycle that includes temperature-dependent seafloor weathering was published by Krissansen-Totton et al. (2017). This model simulated variations in CO2 and climate over the past 100 million years. Several interesting results emerged. The temperature dependence of continental weathering was determined to be relatively weak, based on thousands of model simulations for which results were compared to proxies for CO2, ocean pH, and surface temperature. To compensate for this weak temperature dependence, continental weatherability was required to increase by a factor of 1.7 to 3.3 over the 100 My time interval. This conclusion is in accord with the MacDonald et al. (2019) paper discussed earlier, in which the change in weatherability is attributed to arc–continent collisions in the tropics. But the balance between seafloor weathering and continental weathering, and how that balance changes over time, is only beginning to be understood.

Reverse Weathering

The silicate weathering story now gets more complicated. The seafloor is a complex chemical environment. In some locations, silicate minerals (clay minerals) can actually be formed from reactions between dissolved silica and alkali metal cations, such as Ca2+, Mg2+, and K+ (Sillén 1961). This process, which releases H+ ions rather than consuming them, is termed reverse weathering. Most researchers have concluded that it plays only a minor role in the chemistry of modern seawater. But the role of reverse weathering could have been greatly enhanced in the distant past, prior to the origin of silica-precipitating organisms such as diatoms, radiolarians, and sponges (Lisson and Planavsky...
SUMMARY

The carbonate–silicate cycle has played an important role in regulating Earth’s climate over long timescales. Climate stability is enhanced by a negative feedback between atmospheric CO₂ and the global average silicate weathering rate. The stabilization mechanism is not perfect, and so our planet has experienced multiple (but rare) Snowball Earth episodes. Recovery from these episodes is, nonetheless, ensured by an inevitable buildup of volcanic CO₂ as part of the carbonate–silicate cycle. Lesser perturbations to the system, such as the Himalayan uplift or various tropical arc–continent collisions, have modulated the climate, causing cooling in these examples, but do not destabilize it to the point of threatening the continued existence of life (at least not until we humans appeared). Finally, the same negative feedback that operates on Earth may help to stabilize the climates on rocky planets around other stars. Future telescopic observations may allow us to determine whether other such “Goldilocks planets” exist.

ACKNOWLEDGMENTS

The author thanks the editors for many suggestions that have improved both the content and readability of this paper.

REFERENCES


Coogan LA, Gillis KM (2013) Evidence that low-temperature oceanic hydrothermal systems play an important role in the carbonate–silicate weathering cycle and long-term climate regulation. Geochemistry, Geophysics, Geosystems 14: 1771-1786


Hart MH (1979) Habitable zones about main sequence stars. Icarus 37: 351-357


Poder S (2019) How plants enhance weathering and how weathering is important to plants. Elements 15: 241-246


Of these, the sponges were the earliest to evolve, appearing in the early Cambrian. The presence of such a biotic Si cycle draws down dissolved silica concentrations, slowing the formation of authigenic clay minerals. During the Precambrian, dissolved silica concentrations in the oceans could have been ten times higher than today, making reverse weathering much more efficient (Isson and Planavsky 2018).

The Isson and Planavsky (2018) argument, if correct, can help explain one of the most puzzling aspects of long-term climate history. The Mid-Proterozoic period, from approximately 1.8 Ga to 0.8 Ga, is characterized by remarkable stability in geochemical tracers such as carbon isotopes and by an almost complete lack of evidence for glaciation. By contrast, during the ensuing 543 My of the Phanerozoic Eon, Earth experienced three major glaciations lasting up to tens of millions of years each. For some reason, the climate was warm throughout the Mid-Proterozoic despite the fact that the Sun was dimmer at that time. Isson and Planavsky (2018) suggest that reverse weathering pumped up atmospheric CO₂ to approximately one-hundred times the present level during that time, well above what was needed to offset the faint young Sun. The advent of siliceous organisms brought this period of climate stability to an end, paving the way for our more variable Phanerozoic climate.