Uncertainty and Hypothesis Testing in Planetary Thermal History Models

Thesis by
Johnny Seales

Thesis for the Degree of Doctor of Philosophy
Department of Department of Earth, Environmental and Planetary Sciences
Rice University (Houston, Texas) April, 2020
RICE UNIVERSITY
Uncertainty and Hypothesis Testing in Planetary Thermal History Models

By
Johnny D. Seales Jr.

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE
Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Adrian Lenardic
Professor
Earth, Environmental and Planetary Science

Rajdeep Dasgupta
Professor
Earth, Environmental and Planetary Science

Helge Gonnermann
Associate Professor
Earth, Environmental and Planetary Science

Pedram Hassanzadeh
Pedram Hassanzadeh
Mechanical Engineering

HOUSTON, TEXAS
April 2020
For those that were told they couldn’t do it and those that never made it home to try.
ABSTRACT

Plate tectonics convectively cools the Earth at present. Debate remains in defining plate tectonics as a dynamic system. What is the primary resistor to plate motions: plates, mantle viscosity, some mixture of the two? Here I shed some light on this debate. First, I quantify the uncertainties associated with the different hypotheses. Then, I constrain the probability that different thermal paths from each of these hypotheses match Earth’s geologic proxy data. Only a single hypothesis is true for Earth. The data, however, cannot tell us which. So, rather than using the most probable hypothesis, we can embrace this uncertainty. Each successful thermal path informs us of how other plate tectonic planetary interiors may have evolved in the past or will evolve in the future. Another question arises from inspecting Earth’s thermal history data: what is the convective mechanism for the multi-stage cooling present in some sets of thermal history data? I show that changes in the deep water cycle, namely a switch from a net dehydrating to rehydrating mantle, can act as this mechanism. If this mechanism is true, it may directly influence surface conditions. This may be important for determining whether a planet can maintain liquid water at its surface. Interior processes, then, may play a vital role in determining the inhabitance of a planet. This leaves many searches for life on exoplanets fixated on finding Earth2.0. Depending on the question we ask, however, this may be the wrong approach. I demonstrate that if we cast a broader net in our search for inhabited planets, the increased reward likely outweighs the increased cost.
Preface

The chapters in this thesis represent a cross-section of the published and pending work completed during my studies. The specific chapters relate to the following papers.


Thank you to those who helped me professionally and personally. Thank you to all in the department - faculty, staff, students and everyone else that helped things move forward. To my advisor, thank you for your patience, leadership and insight. Thank you to my friends for the support over the years. Thank you to my family, specifically Ashley, Hailee, Landen, and Riggs. All of your sacrifices helped me reach this goal, so this is our accomplishment. May these learning experiences serve us all well as we move into the next chapter.
Contents

1 Introduction 1

   2.1 Abstract ...................................................... 3
   2.2 Introduction .................................................... 4
   2.3 Methods ....................................................... 7
      2.3.1 Parameterized Thermal History Models .................. 7
      2.3.2 Perturbations ............................................. 12
      2.3.3 Coupled Thermal History and Deep Water Cycling Models ... 14
      2.3.4 Exoplanet Scaled Models (XSM) ......................... 15
   2.4 Analysis ...................................................... 17
      2.4.1 Impulse Response and Reactance Time Analysis ........... 17
      2.4.2 Stochastic Fluctuations and Perturbed Physics Analysis .... 21
      2.4.3 Thermal-Volatile Evolution Model ....................... 25
   2.5 Implications for Planetary Modeling ......................... 34
   2.6 Conclusions .................................................. 37

3 Uncertainty Quantification in Planetary Thermal History Models: Implications for Hypotheses Discrimination and Habitability Modeling 38
3.1 Abstract ................................................................. 38
3.2 Introduction .......................................................... 39
3.3 Methods ............................................................... 43
3.4 Results ................................................................. 51
3.5 Discussion ............................................................ 58
3.6 Conclusions ........................................................... 64

4 Deep Water Cycling and the Multi-Stage Cooling of The Earth 65
   4.1 Abstract ............................................................. 65
   4.2 Introduction ........................................................ 66
   4.3 Methods ............................................................. 69
      4.3.1 Coupled Deep Water Cycling and Thermal History Model .... 69
      4.3.2 Analyzing Structural Stability .................................. 79
      4.3.3 Observational Data Constraints ................................. 79
      4.3.4 First Order Model Feedbacks ................................... 80
      4.3.5 Assessing Feedback Strengths ................................. 82
   4.4 Results ............................................................... 83
      4.4.1 Structural Stability and Structural Uncertainty ............... 83
      4.4.2 Models That Match Observational Data Constraints .......... 86
      4.4.3 Feedback Analysis ............................................. 89
   4.5 Discussion .......................................................... 93
   4.6 Conclusions ........................................................ 97

5 Different is More: The Value of Finding an Inhabited Planet
   that is Far From Earth2.0 98
   5.1 Abstract ............................................................. 98
   5.2 Introduction ........................................................ 99
   5.3 Non-Earthness, Planetary Life Potential, and Search Strategies .... 101
List of Figures

2.1 Schematic showing the detailed feedback loop for the coupled thermal and deep water cycle model. ............................................ 16
2.2 Thermal history response to a single perturbation. ......................... 18
2.3 The residual between the perturbed and reference model shows how quickly each model reacts out a single perturbation. ................. 20
2.4 The reactance time of the convective system as a function of temperature and $\beta$. ................................................................. 22
2.5 Stochastic fluctuations introduced to thermal evolution models results in a cloud of evolution paths. ........................................... 23
2.6 The bound of the uncertainty cloud shows that models with lower reactance times unsurprisingly accumulate less uncertainty. ............ 24
2.7 Coupling the deep water cycle to the thermal model increased the amount of accumulated uncertainty. ................................. 26
2.8 Example use of uncertainties for hypothesis testing. ....................... 27
2.9 Competing models, both of which working for Earth, lead predict different future thermal paths. .............................................. 30
2.10 Planetary size may influence the amount of uncertainty that accumulates. ................................................................. 31
3.1 Geologic proxy data for mantle potential temperature throughout Earth’s history. ............................................................... 41
3.2 Schematic of plate resisting forces and how they relate to the choice of $\beta$. . 46
3.3 Ensemble runs for a single thermal history and how to compare it against present day constraints. .......................................................... 49
3.4 Classified mean ensemble paths for a range of models. .................. 53
3.5 Probability distribution for models that satisfy the present day temperature constraint. ............................................................ 54
3.6 Probability distribution for models that satisfy the present day Urey ratio constraint. ............................................................... 54
3.7 Probability distribution for models that satisfy the present day temperature and Urey ratio constraint. ........................................ 55
3.8 Probability distribution for models that satisfy the paleo temperature constraints. ................................................................. 56
3.9 Probability distribution for models that satisfy all observational constraints. ............................................................ 57
3.10 Mantle water volume and thermal paths that matched all constraints. ... 59

4.1 Estimates of Paleo mantle potential temperature. .......................... 67
4.2 Schematic of the coupled thermal and deep water cycle model. ........... 70
4.3 Simplified feedback loops associated with the thermal (negative feedback) and deep water cycle (positive feedback) modules of the coupled model. . 81
4.4 Structural uncertainty analysis. ..................................................... 84
4.5 Distribution of model paths that matched paleo constraints. .............. 86
4.6 Distribution of model paths that matched present day constraints. ......... 87
4.7 Distribution of model paths that matched all constraints. .................. 88
4.8 Present day surface and mantle water volumes for models that matched all constraints. ......................................................... 89
4.9 Mantle water volume and thermal paths that matched all constraints. . 90
4.10 Quantitative analysis of feedbacks. .............................................. 91
4.11 Summary of all results and uncertainties. ...................................... 94
5.1 Schematic of how different energy sources may affect life. . . . . . . . . . . 103
5.2 Hypothetical distribution of planets relative to Earth based on generic Earth-similarity parameter. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 105
5.3 Cost-benefit and accuracy assessment of different search strategies. . . . . 108
5.4 Results from numerical experiments that vary the total percentage of inhabited galactic bodies. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 110
5.5 The number of discoveries, of inhabited planets, it takes to converge to the attributes that maximize inhabitance for different priors and for different life probability potentials. . . . . . . . . . . . . . . . . . . . . . . . 112
Chapter 1

Introduction

Uncertainty introduces new complexities and computational difficulties into scientific problems. However, if considered in the proper light, properly accounting for uncertainties can shed new light onto old problems. My studies have led me to consider the uncertainties in how a planet, specifically a plate tectonic planet like Earth, cools. Understanding these uncertainties impacts our understanding of Earth’s past, present and future, as well as other planets that may have behaved like it. At the outset, I acknowledge that different models defining how plate tectonics operates dynamically exist. In this thesis I consider these different models as separate hypotheses. In the first part of my thesis, I evaluate these different hypotheses for structural stability and test how well each apply to Earth. In Chapter 2 I show how these different hypotheses respond to unmodeled effects and what this implies about their structural stability and intrinsic uncertainty. In Chapter 3, I evaluate the probability that each of these hypotheses satisfy Earth constraints. I also discuss what this means for Earth’s, and planets that follow an Earth cooling trajectory, future.

There are also more nuanced hypothesis about how Earth cooled. For instance, uncertainty remains in whether plate tectonics has operated over Earth’s entire history. Some believe thermal history data show a change in cooling efficiency in Earth’s past, and they invoke a change in convective regime to explain it. I weigh in on this debate In Chapter 4. I show that a coupled thermal and deep water cycle history model can explain Earth’s apparent multi-stage thermal history. Whether this hypothesis is correct or not, it turns out that the apparent change in convective efficiency aligns with a key atmospheric event in Earth’s history, the Great Oxidation Event. This suggests that Earth’s interior may have
helped life thrive. We are uncertain whether Earth is alone in hosting life, and if they
do exist, what they are like. Given this uncertainty, searching for planets just like Earth
may not be the best approach for looking for life elsewhere. In Chapter 5 I evaluate how
different search strategies may effects our understanding of life in the galaxy.
Chapter 2


2.1 Abstract

Thermal history models, historically used to understand Earth’s geologic history, are being coupled to climate models to map conditions that allow planets to maintain life. However, the lack of structural uncertainty assessment has blurred guidelines for how thermal history models can be used toward this end. Structural uncertainty is intrinsic to the modeling process. Model structure refers to the cause and effect relations that define a model and are assumed to adequately represent a particular real world system. Intrinsic/structural uncertainty is different from input and parameter uncertainties (which are often evaluated for thermal history models). A full uncertainty assessment requires that input/parametric and intrinsic/structural uncertainty be evaluated (one is not a substitute for the other). We quantify the intrinsic uncertainty for several parameterized thermal history models (a subclass of planetary models). We use single perturbation analysis to determine the reactance time of different models. This provides a metric for how long it takes low amplitude, unmodelled effects to decay or grow. Reactance time is shown to scale inversely with the strength of the dominant model feedback (negative or positive). A perturbed physics analysis is then used to determine uncertainty shadows for model outputs. This provides probability distributions for model predictions. It also tests the structural stability of a model (do model predictions remain qualitatively similar, and within assumed model limits, in the face of intrinsic uncertainty). Once intrinsic uncertainty is accounted for, model outputs/predictions and comparisons to observational data should
be treated in a probabilistic way.

**Plain Language Summary**

The Earth’s internal energy is the source of tectonics, volcanism, and geologic activity, all of which effect the surface conditions of our planet. Thermal history models have been used to help understand how this energy, and associated tectonic-volcanic-geologic activity, has evolved over the Earth’s history. These models are now being adopted to map conditions that allow for life on planets beyond our own. However, the uncertainty of these models has not been fully assessed which casts a cloud over how their predictions should be viewed. Here we investigate the uncertainty of thermal history models using several techniques. Our results indicate that model predictions should be viewed in a probabilistic sense which breaks from the way they have traditionally been used for Earth application. These results can also be extended to the exoplanet community. They suggest that efforts to delineate conditions that allow for planetary life will need to proceed under a probabilistic umbrella.

### 2.2 Introduction

The discovery of terrestrial exoplanets has rejuvenated interest in determining the factors that maximize the life potential of a planet [e.g., Lenardic, 2018b, Meadows and Barnes, 2018, Walker et al., 2018]. Modeling the evolution of terrestrial planets, with a focus on mapping scenarios that maintain livable surface conditions, falls under this research umbrella. Consideration of solar energy feeds into this exercise, as does a consideration of a planet’s internal energy. Internal energy drives volcanic and tectonic activity. This, in turn, cycles planetary volatiles between surface and interior envelopes and volatile cycling may be critical for maintaining clement conditions over timescales that allow life to develop and evolve [Walker et al., 1981, Kasting et al., 1993]. The internal energy of terrestrial planets decays over time and the interior of a planet, in the long term, cools until it becomes geologically inactive. Determining/mapping particular and/or potential
planetary cooling paths is the goal of thermal history modeling [Davies, 1980, Schubert et al., 1980]. Thermal history models developed for the Earth have been, and continue to be, adopted and adapted for exoplanet modeling [Kite et al., 2009, Schaefer and Sasselov, 2015]. They are also being linked to climate models so as to model the surface conditions of terrestrial planets over geological time scales [Jellinek and Jackson, 2015, Foley, 2015, Lenardic et al., 2016, Foley and Smye, 2018, Rushby et al., 2018].

Thermal history models, like all models, have uncertainties. Informed application and/or evaluation of models requires assessment of all the potential sources of uncertainty [Draper et al., 1987, Beck, 1987, Draper and Draper, 1995, Walker et al., 2003, Petersen, 2012, Bradley, 2011]. One source of uncertainty is associated with the fact that planetary models involve input parameters that are not perfectly known (e.g., material properties, initial conditions). This is referred to as input and/or parameter uncertainty. It can be assessed by varying input conditions to determine how model outputs respond [e.g., Saltelli et al., 2008, Loucks et al., 2017, chap. 9]. Another source of uncertainty comes from the fact that a model is, at some level, simpler than the phenomena it seeks to model. This introduces an intrinsic uncertainty associated with the effects of unmodelled factors (both known and unknown). This type of uncertainty is also referred to as epistemic, systemic, and structural uncertainty [Strong and Oakley, 2014, Wieder et al., 2015]. Model structure refers to the specific cause and effect relationships that define a model. We use the term intrinsic to highlight that this type of uncertainty is connected to the very definition of a model. Methods for assessing intrinsic uncertainty are model dependent but a general assumption is that the effects of model simplifications do not alter the ability to make useful comparisons of model outputs to existing and/or forthcoming observations (i.e., the intrinsic uncertainty is assumed to be small relative to the uncertainty associated with observational data). Testing the validity of this assumption requires evaluating the intrinsic uncertainty of a model.

Input/parameter uncertainty is generally evaluated in thermal history studies [e.g., Davies, 1980, Christensen, 1984, Korenaga, 2003] and in studies that couple thermal history
models to climate models [e.g., Driscoll and Bercovici, 2013, Foley and Driscoll, 2016, Rushby et al., 2018]. Intrinsic uncertainty has not, to the best of our knowledge, been evaluated for thermal history models. It is not correct to assume that an evaluation of input/parameter uncertainty alone is adequate to assess the confidence level of model outputs/predictions [Draper and Draper, 1995, Hosack et al., 2008, Curry and Webster, 2011, Taleb, 2011, Frigg et al., 2013]. Thermal history models were originally developed, and still widely used, to model the Earth. As such, there has been a focus on models that can match observations and observational constraints are often built directly into thermal history models that focus on the Earth [Christensen, 1984, Korenaga, 2003]. When observational data is available, the proof of a model’s worth is often seen exclusively as its ability to account for the data. If a model can not match the data then it is considered incorrect in terms of assumptions and/or incomplete in the sense that unmodelled factors are not negligible. The latter can be seen as an indirect assessment of intrinsic uncertainty. However, that type of approach can mix intrinsic and parameter uncertainty in a manner that masks the quantitative aspects of the former. Uncertainty in the observational data itself can add to the screening of intrinsic model uncertainty if the only assessment used is how well model results match data. If the level of intrinsic uncertainty is comparable to input/parameter uncertainty and/or data uncertainty, and the intrinsic uncertainty of a model is not directly assessed, then the confidence given to the model can be inflated [e.g., Frigg et al., 2013].

Intrinsic uncertainty assessments can take on a different level of importance when thermal history models are moved from Earth application to exoplanet modeling. Thermal history models applied to Earth data are postdictive. That is, they seek to provide explanations for existing data. Moving models to applications were observations are forthcoming, such that model predictions are now being used to guide efforts at acquiring observations and/or guide methods to evaluate future observations, changes the way model uncertainty assessments are used [e.g., Hofmann and White, 1982]. In general, this enhances the need for a complete model uncertainty analysis as model results are now being used in a
predictive mode to help guide decision making [e.g., Pindyck, 2017].

The primary goal of this work is to quantify the intrinsic uncertainty for a class of thermal history models. We restrict ourselves to providing methods to assess intrinsic uncertainty. In application to any particular planetary modeling problem, intrinsic uncertainty would be coupled to input uncertainty assessment (and potentially uncertainties due to numerical solution methods) to provide a layered uncertainty measure [Taleb, 2011]. An associated goal of this work is to assess the structural stability of thermal history models. The outputs from a structurally stable model do not change qualitatively if the model is perturbed. If, on the other hand, small amplitude perturbations cause the qualitative nature of model solutions to change (e.g., attractors in model solution space appear or disappear), then the model is structurally unstable [Guckenheimer and Holmes, 1983, George and Oxley, 1985]. The remainder of this paper develops and assesses intrinsic uncertainty measures for thermal history models that have been applied to the Earth. All of the models assessed assume that the geological manifestation of mantle convection is in the form of plate tectonics. However, the methods we employ can be applied to models that assume other forms of planetary tectonics [Lenardic, 2018a].

2.3 Methods

2.3.1 Parameterized Thermal History Models

Thermal history modeling seeks to map the cooling paths of terrestrial planets and/or moons. Sub-solidus thermal convection in the rocky interior (the mantle) of terrestrial planets is a key cooling mechanism in their evolutions [Schubert et al., 1979]. Modeling mantle convection over the lifetime of a terrestrial planet is, in principal, possible using 3-D numerical models. In practice, this approach comes with high computational and time costs that can be restrictive for mapping large regions of parameter space. For this reason, thermal evolution models that track spatially averaged quantities remain popular for Earth-focused studies [Conrad and Hager, 1999a, Korenaga, 2003, Sandu et al., 2011, Foley et al., 2012, Höink et al., 2013] and for exoplanet studies [Kite et al., 2009, Schaefer
and Sasselov, 2015, Komacek and Abbot, 2016]. The reduced models are referred to as parameterized thermal history models since they do not solve the full convection equations but use, instead, an empirical formulation that parameterizes variations in convective heat flux as a function of the physical factors that drive and resist convective motions [Schubert et al., 1979, 1980]. These models are efficient for traversing vast regions of parameter space [McNamara and Van Keken, 2000]. They also allow layers of complexity to be added to base level models in relatively simple and efficient ways - for example: deep water cycling [McGovern and Schubert, 1989, Sandu et al., 2011], planetary carbon cycling [Tajika and Matsui, 1990, 1992, Franck and Bounama, 1995, Sleep and Zahnle, 2001, Sleep et al., 2001, Abbot et al., 2012], coupled thermal evolution and climate modeling [Jellinek and Jackson, 2015, Foley, 2015, Foley and Driscoll, 2016, Lenardic et al., 2016]. They can also be scaled to different planetary mass and/or volume in ways that maintain model efficiency [Valencia et al., 2007b].

Parameterized thermal history models are based on a balance between the heat generated within the interior of a terrestrial planet’s mantle and the heat flow through the planet’s surface according to

\[
CT' = H - Q
\]

where \( C \) is the heat capacity of the mantle, \( T' \) is the change in average mantle temperature with time, \( H \) is the total amount of heat produced internally and \( Q \) is the total surface heat flow.

For most of the Earth’s history, heat is produced principally by the radiogenic decay \(^{238}\text{U}, ^{235}\text{U}, ^{232}\text{Th} \) and \(^{40}\text{K} \) isotopes. The total amount of heat produced at any time is modeled as

\[
H(t) = H_0 \sum_{n=0}^{4} h_n \exp(\lambda_n t), h_m = \frac{c_m P_m}{\sum_m c_m P_m}
\]

where \( H_0 \) is total present day heat generation, \( h_n \) is the amount of heat produced by a given isotope, \( \lambda_n \) is the decay constant for a given isotope, and \( t \) is time. Present day fractional concentrations and power production for each isotope are represented by \( c_n \).
and $p_n$, respectively. The values used for each parameter is given in Table 2.1. We assume present day proportions of $U : Th : K = 1 : 4 : (1.27 \times 10^4)$ and normalize by total U to calculate relative isotope concentrations [Turcotte, 1992]. Although we will not include them here, short-lived isotopes of Al and Fe can also make significant contributions to internal heat early in the history of rocky planets [Macpherson et al., 1995].

<table>
<thead>
<tr>
<th>Isotope</th>
<th>$P_n$ (W/kg)</th>
<th>$c_n$</th>
<th>$h_n$</th>
<th>$\lambda_n$ (1/Ga)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{238}U$</td>
<td>$9.37 \times 10^{-5}$</td>
<td>0.9927</td>
<td>0.372</td>
<td>0.155</td>
</tr>
<tr>
<td>$^{235}U$</td>
<td>$5.69 \times 10^{-4}$</td>
<td>0.0072</td>
<td>0.0164</td>
<td>0.985</td>
</tr>
<tr>
<td>$^{232}Th$</td>
<td>$2.69 \times 10^{-5}$</td>
<td>4.0</td>
<td>0.430</td>
<td>0.0495</td>
</tr>
<tr>
<td>$^{40}K$</td>
<td>$2.79 \times 10^{-5}$</td>
<td>1.6256</td>
<td>0.181</td>
<td>0.555</td>
</tr>
</tbody>
</table>

The heat flux through the surface, which depends on convective vigor in the mantle, is typically parameterized using a scaling equation given by Schubert et al. [1979, 1980]:

$$Nu = aRa^{\beta}$$

(2.3)

where $a$ is a constant, $Nu$ is the Nusselt number defined as the convective heat ($Q$) flux normalized by the amount of heat that would be conducted over the entire layer, $q = \frac{k\Delta T}{D}$ where $k$, $\Delta T$ and $D$ are the thermal conductivity, difference between surface and interior temperature and thickness of the convecting layer, respectively. The Rayleigh number ($Ra$) is defined as

$$Ra = \frac{\rho g \alpha \Delta T D^3}{\kappa \eta (T_m)}$$

(2.4)

where $\rho$, $g$, $\alpha$, $\kappa$, and $\eta(T_m)$ are density, gravity, thermal expansivity, thermal diffusivity and viscosity, respectively. The scaling parameter, $\beta$, varies for different classes of thermal history models. We will return to this issue shortly.

The temperature-dependent viscosity of the mantle is defined as

$$\eta(T_m) = \eta_0 exp \left( \frac{A}{RT_m} \right)$$

(2.5)
where $A$ is the activation energy, $R$ the universal gas constant, and $\eta_0$ a constant [Karato and Wu, 1993]. As temperature increases, the viscosity will decrease, leading to an increase in $Ra$. If we assume that all values in equation (2.4) are constant except $T_m$ and $\eta(T_m)$, then combining equations (2.3), (2.4), and the definition of $Nu$ leads to

$$Q = a' \frac{T_m^{1+\beta}}{\eta(T_m)^\beta}$$

(2.6)

$$a' = \frac{ak}{D} \left( \frac{\rho g \alpha D^3}{\kappa} \right)^\beta$$

(2.7)

where all constants have now been combined into $a'$.

Classic thermal history models, developed for Earth, set $\beta$ to a value of 0.33 based on laboratory experiments and boundary layer theory [e.g., Davies, 1980, Schubert et al., 1980]. The constant, $a'$, in equation (2.3), is determined based on experimental results or numerical simulations. The viscosity constant, $\eta_0$, can be based on experiments or set such that the viscosity at a reference present day mantle temperature matches constraints on present day mantle viscosity. With initial conditions, the model is then closed and equation (2.1) can be integrated forward in time.

A $\beta$ value of 0.33 suggests that internal mantle viscosity is the dominant resistance to the motion of tectonic plates [e.g., Davies, 1980, Schubert et al., 1980]. If plate strength offers significant resistance then the scaling constant has been argued to be 0.15 or less [Christensen, 1984, 1985, Conrad and Hager, 1999a]. Korenaga [2003] has argued that $\beta$ is negative as a result of dehydration during plate formation (plates become stronger in the Earth’s past and, as a result, plate velocity and associated mantle cooling decreases even though $Ra$ increases). To account for these different assumptions we will examine cases with $\beta$ values of 0.33, 0.15, 0.0 and -0.15. It is worth noting that variations in $\beta$ represent different assumptions regarding physical processes that are critical to planetary cooling. This distinguishes $\beta$ variations from uncertainties in material parameters (e.g., viscosity function parameters) and initial conditions. Stated another way, variable $\beta$ is associated
with competing hypotheses regarding planetary cooling.

There is a further break between models that have assumed negative or very low $\beta$ values and classic thermal history models (CTM). As noted, CTM have traditionally integrated the energy balance equation forward in time. Workers who argued for low or negative $\beta$ have taken a different approach and used the present day as the starting point and integrated backwards in time [Christensen, 1985, Korenaga, 2003]. A key thought behind that approach is the idea that present day observations, i.e. data constraints, have the lowest data uncertainty (i.e., lower than data constraints for past conditions). The constant $a'$ in equation (2.6) is set such that present day heat flux ($Q_0$) is achieved for assumed present day mantle temperature and viscosity values ($T_0$ and $\eta(T_0)$, respectively). The present day ratio of heat produced within the mantle to that which is transferred through the surface, termed the Urey ratio ($Ur$), can then used to set present day radiogenic heat production according to $H_0 = Ur * Q_0$. The thermal history model then combines equations (2.1), (2.2) and (2.6) resulting in

$$CT = UrQ_0 \sum_{n=0}^{4} h_n \exp(\lambda_n t) - Q_0 \left( \frac{T_m}{T_0} \right)^{1+\beta} \left( \frac{\eta(T_0)}{\eta(T_m)} \right)^{\beta}. \quad (2.8)$$

Equation (2.8) is then integrated backwards in time to produce the Earth’s thermal history constrained by present day values of mantle temperature, heat flux, and Urey ratio (we will refer to this as an Earth-Scaled Model (ESM) formulation). Note that, unlike CTM that integrate forward in time and only specify initial mantle temperature, approaches that integrate backwards in time often fix the initial temperature and its derivative to present day values. This leads to a strongly determined system that restricts the model’s solution domain.

To allow for an apples to apples comparison of uncertainty accumulation for a given $\beta$ model, we use the ESM formulation solved forwards in time. To accommodate for this, and to test low and negative $\beta$ models in a manner that holds to the methods of the workers who argued for these models, we first solved the energy balance equation backwards
by prescribing $T_0$, $Q_0$ and present day $U_r$. Integrating backwards to 4500 Ma, using a Fourth-Order Runge-Kutta scheme, provided the mantle temperature ($T_i$) shortly after accretion and differentiation. The value of $T_i$ was then applied as the initial condition for forwards integration for a period of 4500 Myr. Following this procedure, forward integration recovered the initial conditions (i.e., present day values) of the ESM backward integration approach. This then allows for an apples to apples comparison to evaluate how different models respond to perturbations applied forward in time.

### 2.3.2 Perturbations

A useful metric, for dynamic systems models, is their time response - the time it takes a perturbation to decay [Close et al., 2001, Seely, 1964]. This metric is referred to as the system’s reactance [Texas Instruments Incorporated. Learning Center. et al., 1978]. The reactance for variable thermal history models can be gauged by applying a single perturbation to mantle temperature and tracking the time it takes the thermal history path to damp the change in temperature. In addition to being a useful metric on its own, this will also provide insight into potential model uncertainty associated with unmodelled effects that are not singular in time. If the unmodelled effects are associated with variations that would occur on a timescale longer than the system’s reactance time, then the system has the potential to damp the variations. If the variations occur on a timescale that is shorter than the system’s reactance time, then the variations could be amplified over time and a model has the potential to lose structural stability.

We will perform single-perturbation analysis on full thermal history models and on stripped down versions that exclude the effect of decaying heat sources. The full model analysis will give metrics for full system reactance while the later will give metrics for convective reactance for different models. To isolate the influence of the convective reactance, Equation 2.8 was modified by dropping the heat source term from the right hand side. Using the modified equation, a reference solution was generated by defining an initial temperature and then allowing the mantle to evolve towards colder temperatures.
The initial temperature was then perturbed by 15 °C to generate a perturbed solution. The difference between these two cooling paths was used to evaluate the convective reactance. If the two paths were converging, the perturbation was decaying and an e-folding time was defined as the amount of time it took for the perturbation to decrease by $1/e$. A divergence between the reference and perturbed solutions indicated perturbation growth. For perturbation growth, the e-folding time was defined as the amount of time it took the perturbation to grow by a factor of $e$.

This singular in time perturbation concept can be extended by adding a low amplitude noise term to the governing equation(s) of a model. This is often referred to as a “perturbed physics” analysis. After assessing the reactance times of different thermal history models, we perform uncertainty and stability tests by adding randomized, low amplitude perturbations to each model. Physically, this additional term will mimic the chaotic nature of mantle convection - a factor that is not included in parameterized thermal history models. As well as assessing structural stability, these tests will give metrics for comparing the intrinsic uncertainty between models, i.e., models whose outputs change by a small percentage, due to low amplitude noise, will have a lower intrinsic uncertainty relative to models whose outputs change by a larger percentage.

For the randomized perturbations, the total distribution, for each model integration, was assumed to be normally distributed according to

$$ f(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} $$

where $x$ is the percentage by which the current mantle temperature is perturbed and $\mu$ and $\sigma$ the mean and standard deviation characterizing the distribution of perturbations, respectively. During each integration, a perturbation was randomly drawn from the probability density function (PDF) where $\mu = 0$ and $2\sigma = A$. The parameter $A$ is the prescribed amplitude that defines the positive and negative limits between which 95% of perturbation percentages are located. Therefore, the density of large impulses will be much less than the density of smaller impulses. The impulses are applied at a fixed time.
interval (the interval itself can be varied).

For the case of randomized perturbations, different perturbation time series are possible. For those cases, multiple integrations were performed, for each model, to simulate the manner in which differing random draws affected the uncertainty and stability of different models. For each specific thermal history model, i.e. models with different $\beta$ values, the mean and standard deviation of $T_m$ were calculated at each time step to test the reproducibility of the average mantle temperature that the thermal history models track and to assign associated model uncertainty windows.

### 2.3.3 Coupled Thermal History and Deep Water Cycling Models

An advantage of parameterized thermal history models is that added complexity can be incorporated into them relatively efficiently to build models that couple a range of planetary processes (e.g., models that couple the geologic history of a planet to its climate history). Adding complexity has the potential to transform a structurally stable model into an unstable one (or an unstable model into a stable one). Whether this will or will not be the case is not always easy to assess a priori as added complexity in planetary models can introduce multiple new feedback loops. Some of the loops may be negative (which would favor stability) but some may be positive (which could allow for instability). As an example, we will assess the intrinsic uncertainty of a model that couples deep water cycling to thermal history.

The full description of the deep water cycling model can be found in Sandu et al. [2011]. A conceptual sketch will be useful at this stage (Figure 2.1). Warm mantle rises beneath mid-ocean ridges (MOR) and is subjected to de-compressional melting if it passes the mantle solidus, which is depressed by the presence of water. In the volume of melt produced, a fraction is hydrated resulting in a net dehydration of the mantle and therefore an increase in mantle viscosity. This can lower the vigor of mantle convection which can lead to a hotter mantle and, hence, increased melting and dehydration. This allows for the potential of a positive feedback. As the melt carrying the water migrates to the surface,
some of the water is locked into the lithosphere whereas the remaining water escapes to
the hydrosphere. As the lithosphere is advected away from the MOR, it cools, thickens
and gains water via metamorphic processes. At subduction zones, the downwelling slab
is heated and releases some of its bound water back to the mantle. The re-injection of
water into the mantle lowers its viscosity. This enhances cooling and cooler conditions
favor enhanced mantle rehydration at subduction zones. This allows for a feedback that
works against the previously noted feedback. The degree to which the feedbacks are or
are not balanced, at any time in model evolution, will determine a models uncertainty
metrics. This can be assessed following the procedures outlined in the previous section.
An advantage of running an analysis on the base level thermal history model and the
water-cycling version is that the affects of different model components on the final intrinsic
uncertainty can be assessed.

2.3.4 Exoplanet Scaled Models (XSM)

The discovery of large terrestrial exoplanets (LTePs) has lead to Earth-based thermal history
models being scaled for larger mass and volume planets [Valencia et al., 2007b, Schaefer
and Sasselov, 2015]. Scaling a model has the potential to alter its uncertainty metrics. As
an example, we will scale two models of the previous sub-sections (a water-cycling and
a negative $\beta$ model) and assess the intrinsic uncertainty of the scaled models. We will
follow the scaling approach of Valencia et al. [2006]. The scaling holds core mass fraction
constant at a value of 0.3259. Planet and core radius scale as $R_i \sim R_{i,\oplus} \left( \frac{M_p}{M_\oplus} \right)^a_i$ where $\oplus$ indicates the value of Earth. If the index $i$ represents planet radius, then $a_p = 0.27$. If the
index $i$ represents core radius, then $a_c = 0.247$. From these constraints, mantle volume
and therefore average mantle density, $\langle \rho_m \rangle$ can be calculated. The acceleration due to
gravity for each planet is defined by the relationship $\frac{GM_p}{R_p^2}$. For a scaled water cycling model, ridge length ($L$) is assumed to scale as 1.5 times the planetary radius. The concentration
of mantle water is held constant. This means that total water in the mantle will not be
constant. We hold all other parameters fixed and compute thermal evolutions for 1 $M_\oplus$
Figure 2.1: Accounting for mantle melting, associated water loss, and the effect of dehydration on mantle viscosity complicates the feedback structures of thermal-tectonic evolution models. Different feedbacks can dominate the model’s overall feedback causing it to shift from positive or negative over evolution time. In the particular model diagrammed, the competition between temperature (negative feedback) and water effects (positive or negative feedback) on the mantle viscosity can lead to evolving model feedback dominance.
and $2\ M_{\oplus}$.

### 2.4 Analysis

In this section, we first assess the structural stability and intrinsic uncertainty of several thermal history models using singular in time perturbations and then low amplitude random perturbations that mimic the chaotic nature of mantle convection. Next, we assess the uncertainty of a model that adds a layer of complexity to a particular base level thermal model to evaluate the influence of competing feedbacks on the evolution of model uncertainty. At the end of this section is an uncertainty analysis of two models, both argued to be applicable to Earth, scaled for LTePs.

#### 2.4.1 Impulse Response and Reactance Time Analysis

Figure 2.2 shows the response of several thermal history models to a singular in time perturbation. Model parameters can be found in Table 2.2 and were chosen to match present day constraints. Perturbations decay most rapidly in the classic thermal history model ($\beta = 0.3$) signifying a short reactance time. These models assume the primary resistance to plate tectonic motions is mantle viscosity, which has an inversely exponential dependence on mantle temperature [e.g., Kohlstedt et al., 1995]. This dependence gives rise to the models strong negative feedback [Tozer, 1972] (e.g., if mantle temperature increases, then mantle viscosity decreases which enhances mantle cooling leading to a drop in mantle temperature that works counter to the initial temperature rise). This negative feedback leads to a relatively rapid decay of model perturbations (Figure 2.2a).

As $\beta$ was decreased, reactance time increased. Models with decreasing $\beta$ assume that plate motion is resisted by a combination of plate strength and mantle viscosity [Conrad and Hager, 1999a]. This weakens the negative feedback discussed above and, as a result, model reactance time increases (Figure 2.2b). The assumption for $\beta = 0$ models is that resistance to plate motion is dominated by plate strength which has no dependence on mantle temperature [Christensen, 1984, 1985]. This removes a negative (buffering) feedback...
Figure 2.2: Thermal history response to a single perturbation. The unperturbed reference model for each model is the solid black line. The different choice of beta for each model represents different model assumptions regarding the dominant resistance term(s) to plate motion. As $\beta$ is decreased from its classic value to zero and then into the negative domain, it takes longer for convective mantle overturn to eliminate the perturbation, if it does so at all.
Table 2.2: Convection Parameters

<table>
<thead>
<tr>
<th>Model Parameter</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta_0$</td>
<td>$2.21 \times 10^9$</td>
<td>$Pa \cdot s$</td>
</tr>
<tr>
<td>$A$</td>
<td>300</td>
<td>$kJ/mol$</td>
</tr>
<tr>
<td>$R$</td>
<td>8.314</td>
<td>$J/(Kmol)$</td>
</tr>
<tr>
<td>$Q_0$</td>
<td>36</td>
<td>$TW$</td>
</tr>
<tr>
<td>$T_0$</td>
<td>1623</td>
<td>$K$</td>
</tr>
</tbody>
</table>

Model dependent parameters

- $\beta = -0.15$
  - $U_r = 0.15$  
  - $T_i = 1956$ $K$

- $\beta = 0.00$
  - $U_r = 0.35$  
  - $T_i = 1939$ $K$

- $\beta = 0.15$
  - $U_r = 0.55$  
  - $T_i = 2182$ $K$

- $\beta = -0.15$
  - $U_r = 0.75$  
  - $T_i = 2016$ $K$

from the model system and the effects of perturbations survive to present day (Figure 2.2c). The slow convergence of perturbed model paths in Figure 2.2c is due to the decay of mantle heat sources - over a long evolution time, heat sources will be tapped and all the paths must approach a final non-convecting state (the model analog for a geologically dead planet). Models with negative $\beta$ assume that resistance to plate motion is dominated by plate strength and further assume that plate strength increases with mantle temperature [Korenaga, 2003, 2008]. This leads to a positive feedback within the model system [Moore and Lenardic, 2015]. As a result, evolution paths become highly sensitive to perturbations.

The differences in model reactance times can be highlighted by tracking the percent error between the perturbed and reference model paths (Figure 2.3). The percent error decayed quickly for models with shorter reactance times. Increasing reactance times, caused by shifting $\beta$ towards zero, led to a slower decay. For $\beta$ less than zero, the positive
Figure 2.3: The residual between the perturbed and reference model shows how quickly each model reacts out the single perturbation. (a-c) As $\beta$ is reduced, the time it takes for the perturbation to decay by $1/e$, a measure of the reactance time ($\tau$), increased. (d) When $\beta = -0.15$, the perturbation grows rather than decays.

system feedback led to error growth. For any given $\beta$, the error decay or growth rate was a function of perturbation sign. Warmer perturbations decayed faster than colder ones for all models with $\beta \geq 0$.

The asymmetry between warm and cold perturbations indicates that reactance time is a function of $\beta$ and of mantle temperature. To assess the influence of these two effects, we isolated the convective reactance time ($\tau_c$), defined as an $e$-folding timescale for perturbation growth or decay (this metric removes the effects of decaying heat sources on model reactance times). The value of $\tau_c$ was calculated for different initial temperatures and
choices of $\beta$ (Figure 2.4). For $\beta > 0$, a positive perturbation increases convective vigor and heat transport leading to a rapid perturbation decay. A negative perturbation decreases convective vigor relative to the unperturbed state. The overall feedback, however, remains negative which still leads to a perturbation decay but at a slower rate than for a positive perturbation. As $\beta$ was decreased, $\tau_c$ increased. The lengthening of $\tau_c$ continued until $\beta$ took on negative values.

When $\beta$ is reduced past zero, the overall model feedback changes from negative to positive. For negative $\beta$, $\tau_c$ is the amount of time it takes a perturbation to grow by a value of $e$. As $\beta$ became more negative, perturbations grew more rapidly (Figure 2.4). For a positive feedback, a perturbation that increases mantle temperature decreases convective heat transfer. Decreasing mantle temperature leads to the opposite effect which resulted in a greater divergence between the perturbed and the reference cooling path. The fastest perturbation growth occurred at the combination of coldest temperatures and most negative $\beta$ values (Figure 2.4).

2.4.2 Stochastic Fluctuations and Perturbed Physics Analysis

The response of thermal history models to random fluctuations, over model evolution time, is shown in Figure 2.5. Each model was subjected to 500 random perturbation sequences. During each sequence, the model was perturbed at a 10 Myr interval. We chose this timescale to assess how short time-scale fluctuations, associated with the chaotic nature of mantle convection, can influence model behavior (associated with chaotic convection are potential changes in convective wavelength which can also cause fluctuations in mantle heat loss [Grigné et al., 2005]). Each perturbation was drawn at random from a distribution defined by a mean of zero and a standard deviation of 7.5 $^\circ$C. The maximum single perturbation was then approximately 1% of the total temperature, which is in line with the natural variation due to the chaotic nature of vigorous convection [e.g., figure 4 of Weller et al., 2016]. The random draw meant that the amplitude and sign of the perturbations were stochastic and the full perturbation time series was non-periodic. Plotting all of the
Figure 2.4: The reactance time of the convective system ($\tau_c$) as a function of temperature ($T$) and the choice of $\beta$ (the perturbation amplitude is fixed at 15°C). For $\beta \geq 0$, the perturbation was damped by the model’s negative feedback. A stronger negative feedback (more positive $\beta$) as well as increased $T$ lowered $\tau_c$. For $\beta < 0$, the model’s positive feedback, which strengthened as $\beta$ decreased, amplified the perturbation.
Figure 2.5: Stochastic fluctuations introduced to thermal evolution models results in a cloud of evolution paths, the gray lines in a-d, about the mean (black lines). The blue lines in a-d are the statistical mean of the clouds and match fairly well with the reference model. As $\beta$ was decreased, the uncertainty cloud widened. Distinct time slices through this cloud – red, green and blue are 0, 1000, and 2000 Ma – result in normally distributed temperatures with increased variance as the model evolved towards present day as $\beta$ decreased.

Perturbed paths generates an uncertainty shadow (top row of Figure 2.5). The uncertainty shadow is plotted along with the mean trend from the 500 perturbed models (blue lines) and the trend from the unperturbed models (black lines). A comparison of the mean trends and the unperturbed model trends shows that all the models maintain structural stability. All the models are not, however, associated with the same uncertainty structure. For the classic thermal history model, model uncertainty tended to saturate tightly around the reference solution. As $\beta$ was decreased, model uncertainty increased.

For any specific model evolution time, a probability distribution of mantle temperatures can be compiled. Distributions for the present day, 1000 Ma and 2000 Ma are shown in the bottom row of Figure 2.5. Each distribution is normally distributed about the unperturbed path with a time-dependent standard deviation. The smaller the standard deviation, the lower the model uncertainty.
Figure 2.6: The bound of the uncertainty cloud (two standard deviations from the mean) shows that models with lower reactance times unsurprisingly accumulate less uncertainty. Over the age of the Earth, a model with a negative feedback tended to have uncertainty saturate whereas a positive feedback model accumulated uncertainty.

Rather than limit the analysis to three specific model times, a continuous uncertainty window can be tracked (Figure 2.6). For positive $\beta$, uncertainty saturates over time. The uncertainty saturation limit is proportional to $\tau_c$. A small $\tau_c$ indicates that the system reacts out perturbations quickly, resulting in a smaller accumulation of uncertainty. As $\beta$ approached zero, but remained positive, $\tau_c$ increased, which enhanced uncertainty accumulation, though it remained bounded.

Unlike the $\beta \geq 0$ models, there is no uncertainty saturation limit for the $\beta < 0$ model. Instead, as the model evolved to lower temperatures, there was an increase of model uncertainty and its accumulation rate. As the mantle secularly cooled, any positive perturbation took the perturbed model further from the reference model, reducing the rate at which it cooled, creating a greater divergence between the two models. Alternatively, any negative perturbation created more rapid cooling in the perturbed model, causing it to also diverge from the reference model.
2.4.3 Thermal-Volatile Evolution Model

Introducing new layers of complexity to a model has the potential to change its stability and uncertainty properties. As an example, we consider adding deep water cycling to a thermal history model [see Sandu et al. [2011] and the description of Figure 2.1 for details]. Mantle de- and re-hydration couple with temperature to determine mantle viscosity which, in turn, feeds into convective vigor. As a result, model reactance times, structural stability, and intrinsic uncertainty can all be altered.

Uncertainty was evaluated for the coupled model using the procedures of the previous subsection. The uncertainty shadow from the coupled model is plotted in Figure 2.7 along with the $\beta = 0.3$ reference model. Uncertainty tripled relative to the reference case at the final model evolution time. The addition of a positive feedback associated with the deep water cycle decreased the overall negative feedback of the model. Although the uncertainty is not fully bounded over the model run time shown in Figure 2.7, it is approaching a limit asymptotically (this will be made clear in the next subsection). Enhanced uncertainty in the coupled model is due to the continuous competition between thermal and water cycle effects on mantle viscosity. The competition alters the overall system feedback over model evolution time, which affects the accumulation of uncertainty. This points to the value of performing feedback analysis to map possible model trends consistent with internal feedbacks [e.g., Crowley et al., 2011, Astrom and Murray, 2008]. Feedback analysis can provide qualitative insights into uncertainty potential (e.g., does the model allow for unbounded uncertainty) without having to run the full model. For models of the type explored here, the computational time to run models is not a restriction but for more complex models it can be which adds to the value of feedback analysis as a first step in evaluating uncertainty potential.

An example of how model results, accounting for intrinsic uncertainty, can be compared with observational data is shown in Figure 2.8 (the data comes from Condie et al. [2016]). In Figure 2.8a, the typical method of presenting a model solution with data constraints is shown. Adding an uncertainty shadow (Figure 2.8b) allows for a more complete com-
Figure 2.7: The deep water cycle model (blue line) weakened the negative feedback in the model, lengthening $\tau_c$, which allowed for more uncertainty accumulation as compared to the same thermal history model without water cycling (red line).
Figure 2.8: Comparing model results to observations is a means of hypothesis testing. In (a) we plot a thermal evolution that accounts for the deep water cycle along with the data constraints [modified from Figure 4b Condie et al., 2016]. Accounting for intrinsic model uncertainty, leads to plotting an uncertainty cloud along with the mean (b). Within the uncertainty cloud, different cooling trends (dotted and dash-dotted lines) are capable of satisfying the data, limiting our ability to constrain the particular path a single planet may have taken (c).

Comparison of the model prediction with observational data. For the chosen initial conditions and parameters, the coupled model could satisfy observational constraints at a model uncertainty level comparable to that of the observational data itself. Two of the model solutions that determined the model uncertainty shadow are shown along with the mean solution in Figure 2.8c. Either path, within model uncertainty, is a viable model result. Plotting them shows the level of model deviations that can occur relative to the mean model trend (i.e., the trend from numerous models that account for stochastic fluctuations). The individual paths in Figure 2.8c evolved similarly in the initial and final stages of model evolution. However, a peaked divergence of \(~150~^\circ\text{C}\) occurred between them at around 2500 Myr of model evolution. Whether this level of model uncertainty is large enough to alter the results from more complex models that, for example, couple climate evolution to thermal evolution will depend on the dynamic properties of the climate models used. Given that many climate models allow for bi-stable behavior, i.e. multiple solutions under equivalent parameter conditions [e.g., Scheffer, 2009], this possibility can not be ruled out.
**Exoplanet Forecasting: Age Dependence**

Thermal history models are being projected in time to understand how terrestrial planets evolve outside our own solar system. Doing so introduces new forms of uncertainty. Even for the Earth, the planet with the best observational data, there is debate regarding which thermal history model best represents the Earth’s thermal evolution [e.g., Conrad and Hager, 1999a, Grigné et al., 2005, Korenaga, 2003, Grigné and Labrosse, 2001, Korenaga, 2008, Silver and Behn, 2008, Sandu et al., 2011, Moore and Webb, 2013] Model selection uncertainty is a form of structural uncertainty: competing models for the Earth’s thermal history have different mathematical structures to represent what different researchers consider to be the essential factors that have determined the Earth’s thermal evolution.

Figure 2.9 shows thermal evolution paths from two models that have been applied to the Earth. One model is the deep water cycling model of the previous section [Sandu et al., 2011]. The other is an ESM with $\beta < 0$ value [Korenaga, 2003, 2008]. Petrological data constraints are plotted with each model [Herzberg et al., 2010, Condie et al., 2016]. The models, as originally presented, did not include uncertainty shadows, which we have determined following the procedures of the previous sections. We have also extended model time beyond the age of the Earth. Although the Earth provides the best observational constraints on a thermal history model, there is still uncertainty in the data. This includes uncertainty in observational constraints on the Earth’s past thermal state [e.g., Herzberg et al., 2010, Condie et al., 2016] and uncertainty in observational data that can constrain the Earth’s present thermal state (average surface heat flux, present day radiogenic heat production, the average internal temperature of the Earth’s mantle [e.g., Jaupart et al., 2015, Sarafian et al., 2017]). Given data and model uncertainties, different models can satisfy observational constraints. Evaluating intrinsic uncertainty with other modeling uncertainties – initial condition and parametric – can broaden the viable model space. Different techniques (e.g., grid search, Monte Carlo) can be used to determine which models can match observational constraints over a larger portion of potential parameter space in an effort to gauge models that are statistically preferred [e.g., McNamara and Van Keken,
2000, Höink et al., 2013]. Such approaches can rule out some classes of models but multiple competing models can remain viable. In short, models based on different assumptions can fit Earth based data constraints within allowable uncertainty bands. The critical point for modeling exoplanets is that there are competing hypothesis for the the Earth’s thermal evolution which brings with it an uncertainty associated with model selection.

The model paths plotted in Figure 2.9 are for particular parameter and initial conditions that can match data trends within uncertainty. The differences in model structure, and associated uncertainty, becomes clear when the models are projected forward in time for 10 billion years (Gyr). The intrinsic uncertainty for the deep water cycling model flattens over time. For the ESM model, uncertainty increases over time and after 5 Gyr of evolution the temperatures on the cold side of the uncertainty shadow become so low that the cold paths rapidly decay. This leads to the mean of multiple perturbed models deviating from the unperturbed model trend (for this reason, the uncertainty shadow is cut off at this point - beyond this point, probability distributions of model outputs deviate from Gaussian and can become fat-tailed on the cold side of model evolution). The ESM model predicts that volcanic-tectonic activity should end after 5-6 Gyr (mantle temperatures become too cold to allow for continued melt generation). The deep water cycling model predicts a longer volcanic-tectonic lifetime. If model uncertainty is not taken into account this could lead to some discordant claims about the geologically active lifetime of our planet and, by association, about the habitability potential of exoplanets.

**Exoplanet Forecasting: Size Dependence**

As well as being extended in time, thermal history models can be scaled for larger planets. There is debate about which convective regime is likely to occur on larger planets. Some suggest that plate tectonics is likely [Valencia et al., 2007b, Valencia and O’Connell, 2009, Van Heck and Tackley, 2011, Tackley et al., 2013] whereas others suggest that need not be the case [O’Neill and Lenardic, 2007]. Keeping in line with the theme of this paper, we only evaluate the uncertainty associated with a plate tectonic mode, but do not dismiss
Figure 2.9: In the case of Earth, competing models – deep water cycling with $\beta = 0.33$ in (a) and the ESM with $\beta = -0.15$ in (b) – are capable of matching petrological constraints within model and data uncertainty (red dots with error bars are from Condie et al. [2016, Figure 4b] and blue dots are from Herzberg et al. [2010]). However, the models lead to different predictions for the evolution of a terrestrial planet beyond the current age of the Earth.
Figure 2.10: Scaling a model to terrestrial planets of different size ($1M_\oplus$ in blue and $2M_\oplus$ in red) may or may not affect the intrinsic uncertainty. In the deep water cycling model (a), which has $\beta$ set to 0.33, the strong negative feedback of the scaled up model keeps it near an evolution similar to the Earth sized model. In (b) the model has no water cycling and a positive feedback, $\beta = -0.15$, and therefore the larger planet model has initially warmer temperatures that persist longer, preventing the onset of structural instability until further into the model evolution, as indicated by the cutoff in the uncertainty cloud.
that other modes are possible. The simplest method for scaling a planetary model is to increase size and mass of the planet while holding all else equal – an assumption that is used herein only for demonstration. Parameters used for the scaled models can be found in Table 2.3. For the deep-water cycling models, it is assumed that each planet begins with the same mantle water concentration and the same mantle potential temperature, which is consistent with previous studies that have scaled deep water cycle models [e.g., Schaefer and Sasselov, 2015]. Similar to previous studies, we report average mantle temperature ($T_m$) rather than potential temperatures. This means we have extrapolated the potential temperature along the adiabat to mid mantle depth. We use this convention for consistency with previous exoplanet studies.

In Figure 2.10(a) we plot the reference solutions for 1$M\oplus$ and 2$M\oplus$ planets together with uncertainty shadows. Model differences that arise in the early stages of evolution result from the competition between thermal and deep-water cycle effects on mantle viscosity. For the choice of parameters, the Earth-sized model experienced a time window of little to no cooling between 1-2 Gyr. During this time, the mantle was dominantly dewatering which tended to increase mantle viscosity, lower convective vigor, and favor mantle heating. That heating tendency, due to more sluggish convection, was balanced by the tendency of temperatures to drop due to decaying heat sources. This lead to a flat line cooling trend. Following this period, water cycled back into the mantle from the surface and cooling was enhanced. Doubling the size of the planet increased the negative thermal feedback such that it outweighed the positive feedback due to mantle dewatering (that positive feedback was critical for the flat line cooling phase of the Earth sized model). This decreased $r_c$ for the scaled up model, relative to the Earth model, which lowered its relative uncertainty. This implies that, for this particular model suite, the uncertainty of a reference Earth model will not underestimate the uncertainty of a scaled up model.

The ESM model with $\beta = -0.15$ was also scaled for a larger planet (Figure 2.10(b)). After 5 Gyr of model evolution, the reference model runs away to cooler temperatures. This signals that the evolution has moved outside of the conditions that the model was designed for.
Table 2.3: Deep Water Cycle Model Parameters

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convective Model</td>
<td>$T_s$</td>
<td>Surface temperature</td>
<td>300</td>
<td>K</td>
</tr>
<tr>
<td></td>
<td>H(0)</td>
<td>Initial radiogenic heat</td>
<td>4.51</td>
<td>J/(m^3 yr)</td>
</tr>
<tr>
<td></td>
<td>Rm</td>
<td>Mantle radius</td>
<td>6271</td>
<td>km</td>
</tr>
<tr>
<td></td>
<td>Rc</td>
<td>Core radius</td>
<td>3471</td>
<td>km</td>
</tr>
<tr>
<td></td>
<td>$\rho_m$</td>
<td>Mantle density</td>
<td>3000</td>
<td>kg/m^3</td>
</tr>
<tr>
<td></td>
<td>$k_m$</td>
<td>Thermal conductivity</td>
<td>4.2</td>
<td>W/(mK)</td>
</tr>
<tr>
<td></td>
<td>$c_p$</td>
<td>Specific heat</td>
<td>1400</td>
<td>J/(kgK)</td>
</tr>
<tr>
<td></td>
<td>$\alpha$</td>
<td>Thermal expansivity</td>
<td>$3.00 \times 10^{-5}$</td>
<td>K^{-1}</td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>Convective exponent</td>
<td>0.33</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$\lambda$</td>
<td>Decay constant</td>
<td>$3.4 \times 10^{-10}$</td>
<td>yr^{-1}</td>
</tr>
<tr>
<td></td>
<td>$Ra_{cr}$</td>
<td>Critical Rayleigh number</td>
<td>1100</td>
<td>-</td>
</tr>
<tr>
<td>Water Cycling</td>
<td>$\eta_0$</td>
<td>Viscosity constant</td>
<td>$1.7 \times 10^{17}$</td>
<td>Pa·s</td>
</tr>
<tr>
<td></td>
<td>$A_{cre}$</td>
<td>Material constant</td>
<td>90</td>
<td>MPa·r/s</td>
</tr>
<tr>
<td></td>
<td>$r$</td>
<td>Fugacity exponent</td>
<td>1.2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$Q_a$</td>
<td>Creep activation energy</td>
<td>$4.8 \times 10^{5}$</td>
<td>J/mol</td>
</tr>
<tr>
<td></td>
<td>$\chi_d$</td>
<td>Degassing efficiency factor</td>
<td>0.03</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$\chi_r$</td>
<td>Regassing efficiency factor</td>
<td>0.015</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>OM(0)</td>
<td>Mass of 1 Earth ocean</td>
<td>$1.39 \times 10^{21}$</td>
<td>kg</td>
</tr>
</tbody>
</table>

from the start (models, in general, assume limits of validity and, as such, what they predict should only be physically interpreted within those limits). In this particular case, the model relies on the assumption that melting can occur within the mantle to generate a chemical lithosphere. Once mantle temperature drops so low that melting can not occur the assumed limits of model validity have been crossed and a different formulation would likely need to be implemented. This could be in the form of another plate tectonics model or a transition to a different convective regime. Doubling the mass of the planet increased the resilience against runaway by lessening the positive feedback inherent to this model (cf. Figure 2.4). As a result, uncertainty grew more slowly than for the Earth model. However, several of perturbed models, that evolved toward the colder side of the uncertainty envelope, did run away at 8 Gyr of evolution (this is why the uncertainty envelope is shown truncated at 8
2.5 Implications for Planetary Modeling

The goal of our uncertainty analysis, for parameterized planetary thermal history models, was to provide insight into their utility, not to advocate for the correctness or use of one model over another. Uncertainty analysis can not determine if the assumptions of one model are more or less physically valid than those of a competing model. It can provide guidelines for how to interpret and apply model results.

When intrinsic uncertainty is accounted for, model results are presented as probability distributions (Figure 2.5). This introduces a probabilistic element to model testing. A mean model path that may have been cast out as unsuccessful, because it did match observations, can remain viable if its’ uncertainty cloud overlaps data uncertainty. The degree of overlap allows a non-binary confidence measure to be determined. Model testing can still proceed in a probabilistically manner with confidence cutoffs being stated (e.g., “we consider initial condition and parameter combinations that can match data within two sigma confidence as viable”). If a model uncertainty cloud dwarfs data uncertainty, then it could become difficult to use the data itself to rule out particular model paths. In effect, a model could become irrefutable given the data. An irrefutable model is not necessarily based on invalid assumptions but it does lose a level of utility.

Though preliminary, we can provide some implications for using models to interpret the Earth’s thermal history in light of uncertainty. The larger the spread of model probability distributions, the lower the confidence one can place in using the model to make predictions and/or postdictions. The thermal history models with the highest uncertainty are associated with a positive feedback, which is already known to introduce a strong initial condition dependence to model outputs [Korenaga, 2016]. It also brings the possibility that any unmodelled effects (e.g., post formation impacts [O’Neill et al., 2017]) could have as large an effect on model outputs as do the assumed critical factors that define the model (the tighter the model uncertainty probability distribution, the lower the effect of unmodelled factors...
relative to the effect of the particular heat flux scaling that the model is based on). For illustration purposes we assumed models that operated in a plate tectonic mode. However, other convective regimes may be possible over Earth’s lifetime [e.g. Lenardic, 2018a]. The mean evolution path of a plate tectonic model may not lead it to cross a regime boundary that initiates a different mode of tectonics. If uncertainty is considered, then it is possible that the model allows for a tectonic transition. The greater the uncertainty cloud, the greater this potential. A final implication relates to how non-monotonic thermal-tectonic signals in the rock record are interpreted. Independent of intrinsic uncertainty analysis, certain classes of thermal history models predict monotonically decaying trends for mantle temperature, and by association plate velocities, over the bulk of an evolution path (e.g., classic thermal history models with a $\beta$ of 0.33). Considering intrinsic uncertainty can prevent the incorrect inference that evidence for mantle heating in the past and/or an increase of plate velocities is necessarily arguing against the validity of these models (Figure 2.8c). It is possible that the magnitude of inferred non-monotonic trends could favor one model over another but, in order to determine that, the magnitude of allowed fluctuations would need to be quantified. That is to say, an intrinsic/structural uncertainty analysis would need to be performed.

An implication for exoplanet modeling, in light of intrinsic model uncertainty, relates to planetary habitability. A prominent issue in planetary modeling is mapping conditions that allow a planet to maintain liquid water over geological time scales, a feature considered critical for life as we know it [e.g., Meadows and Barnes, 2018]. This has lead to the coupling of thermal history and climate models [e.g., Foley, 2015, Foley and Driscoll, 2016, Rushby et al., 2018]. These coupled models track volatile cycling between a planets interior and its surface envelopes. Coupling a thermal history model to a volatile cycling model can compound model uncertainty in a non-linear way (Figure 2.7). Climate models will be associated with their own uncertainties [e.g. Mahadevan and Deutch, 2010]. Propagating uncertainties in coupled solid planet and climate models has not been studied to date. If the coupling amplifies intrinsic uncertainty, then the predictions from a coupled model
may become highly uncertain, even for a unique combination of initial condition and parameter values. A situation can arise where a unique combination of model inputs (e.g. solar distance, planet mass, initial water content) leads to an uninhabitable planet based on models that do not account for intrinsic uncertainty. However, if intrinsic uncertainty is accounted for, then this conclusion may no longer hold. Accounting for uncertainty associated with model selection can compound this potential. Coupled climate and thermo-tectonic evolution modeling studies have dominantly used a single thermal history model, a single $\beta$ value. However, within uncertainty, different $\beta$ models can be consistent with Earth constraints (Figure 2.9). These models can produce very different predictions for planets older than the Earth. Collectively, this allows for the potential that the mapping of habitable conditions to date may be associated with an inflated level of confidence.

To determine the degree to which this is the case, future studies will be required that account for layered uncertainty analysis (coupling of structural/intrinsic uncertainty, input uncertainty, and model selection uncertainty). This motivates future work that extends beyond the scope of this paper.

The points raised herein are not a call for less modeling, but more. Specifically, a layer of modeling that is not geared at making predictions that can be compared to data or that can be used to guide future exploration to gather new observations but, instead, is designed to quantify model uncertainties before the models are put into application modes. With this comes the value of a multi-model approach. Different base level thermal history models can be used collectively to determine which conclusions remain robust in light of model selection uncertainty together with the intrinsic uncertainties of particular models. This adds a complication to using planetary models to guide our thinking about conditions that allow for planetary life beyond Earth but the problem is not intractable. The value of parameterized thermal history models, and the coupling of such models to simplified climate models, is that they remain simple enough to run millions of models given current computational power. This can provide uncertainty measures which, together with a multi-model approach, can be used to put confidence limits on planetary model predictions.
something that will be appreciated by observationally focused colleagues who strive to provide uncertainty measures on observational data. The search for planets that allow for life beyond Earth involves a synergy between modeling and observations [e.g., Kasting, 2012]. Both are associated with uncertainty. This does not stand in the way of moving the joint venture forward. Uncertainties can be evaluated and compared. If probability distributions are determined for observations and for models this opens the path to use, for example, Bayesian analysis to provide well defined confidence levels for planetary life potential under a range of potential scenarios [e.g., Walker et al., 2018]. This approach becomes most effective if exoplanet search strategies are designed with statistical analysis in mind [e.g., Lenardic and Seales, 2019].

2.6 Conclusions

Thermal history models are associated with uncertainty that is independent of imperfectly known initial conditions and/or input parameter values. This intrinsic uncertainty can be assessed. Isolating dominant model feedbacks can give qualitative insight as to whether uncertainty can be amplified or damped. Single perturbation analysis can bring a more quantitative assessment by determining the reactance time of a model (the time for perturbations to grow or decay by some amount). A perturbed physics approach can provide a further metric by determining an uncertainty shadow for a particular model evolution over time. It can also determine the structural stability of a model and provide model output probability distributions that account for uncertainty. Once intrinsic uncertainty is accounted for in thermal history models, model outputs/predictions and comparisons to observational data should be treated in a statistical/probabilistic way.
Chapter 3

Uncertainty Quantification in Planetary Thermal History Models: Implications for Hypotheses Discrimination and Habitability Modeling

3.1 Abstract

Multiple hypotheses/models have been put forward regarding Earth’s cooling history. Searching for life beyond Earth has brought these models into a new light as they connect to an energy source life can tap. Discriminating between different cooling models and adapting them to aid in the assessment of planetary habitability has been hampered by a lack of uncertainty quantification. Here we provide an uncertainty quantification that accounts for a range of interconnected model uncertainties. This involved calculating over a million individual model evolutions to determine uncertainty metrics. Accounting for uncertainties means that model results must be evaluated in a probabilistic sense, even though the underlying models are deterministic. The uncertainty analysis was used to quantify the degree to which different models can satisfy observational constraints on the Earth’s cooling. For the Earth’s cooling history, uncertainty leads to ambiguity - multiple models, based on different hypotheses, can match observations. This has implications for using such models to forecast conditions for exoplanets that share Earth characteristics but are older than the Earth, i.e., ambiguity has implications for modeling the long-term life potential of terrestrial planets. Even for the most Earth-like planet we know of, the Earth itself, model uncertainty and ambiguity leads to large forecast spreads. Given Earth has the best data constraints, we should expect larger spreads for models of terrestrial planets in general. The uncertainty analysis provided here can be expanded by coupling planetary cooling models to climate models and propagating uncertainty between them to
assess habitability from a probabilistic view.

3.2 Introduction

The surface conditions of the Earth have evolved over our planet’s history in response to two energy sources: solar energy and internal energy. Both energy sources have, themselves, evolved and continue to do so. Stellar models provide insights into the Sun’s energetic evolution [Feulner, 2012]. Thermal history models provide insights into the cooling of the Earth’s interior [Davies, 1980, Schubert et al., 1980]. Earth’s internal energy comes from the decay of radioactive isotopes within its rocky interior and from heat retained from planetary formation and early differentiation. This internal energy drives volcanic and tectonic activity, both of which influence the cycling of life-essential elements and volatile elements, such as greenhouse gasses, between the Earth’s interior and surface reservoirs (atmosphere, hydrosphere, biosphere). That connection to elemental cycling, along with the discovery of life that can tap into the Earth’s internal energy [Baross and Hoffman, 1985, Jannasch and Mottl, 1985] and an expanding search for life beyond Earth, has rejuvenated interest in the cooling history of the Earth and, by association, thermal history models. This renaissance has moved thermal history modeling from the realm of geosciences into the realm of astronomy and astrophysics [Kite et al., 2009, Schaefer and Sasselov, 2015, Komacek and Abbot, 2016, Foley, 2015, Foley and Driscoll, 2016, Tosi et al., 2017, Foley and Smye, 2018, Rushby et al., 2018, Barnes et al., 2020].

When a modeling methodology moves from one discipline to another there is the potential for synergies and for misconceptions. The Earth has the largest observational data set that can constrain planetary models. However, this does not mean that significant uncertainties do not remain. This has not been communicated as well as it could be across communities. Even within the geosciences’s community itself the role of uncertainty and ambiguity for thermal history models has not received the level of attention given to it in other modeling endeavors (e.g., water resources, climate [Loucks et al., 2017, Curry and Webster, 2011]). This provides the two-pronged motivation for this paper: 1) Given
data and model uncertainties, what is the confidence level we can give to different Earth cooling models and, by association, are multiple models viable? 2) What implications does uncertainty regarding the Earth’s thermal history carry for modeling the habitability of terrestrial planets?

The cooling history of a planet depends on its tectonic mode [Lenardic, 2018a]. The Earth’s present mode is plate tectonics [McKenzie and Parker, 1967, Morgan, 1968]. The simplest starting assumption is that plate tectonics has operated over the Earth’s geologic history (i.e., since the transition from a magma ocean phase to a phase of planetary evolution that preserves a rock record [Sleep, 2000]). This assumption has been made by the majority of Earth thermal history models to date, and we will follow suit herein. With knowledge of our conclusions, we can say that geologic proxy data (Figure 3.1) used to constrain the Earth’s cooling cannot rule out this possibility. Models that allow for tectonic transitions may also be able to match data constraints, but that will only increase the effects of model uncertainty. By assuming a single tectonic mode we will not only follow an Occam’s razor approach, but we will also be conservative in assessing model uncertainty.

The theory of plate tectonics is a kinematic one that defines the Earth’s surface as being divided into internally rigid, rocky plates that move relative to each other with deformation and volcanic activity concentrated along plate boundaries. For the Earth’s cooling, a key factor is that cold tectonic plates can sink back (subduct) into the Earth’s warmer, rocky interior (i.e., plates are a component of the upper thermal boundary layer of the solid Earth’s thermal convection system). Extending the kinematic theory of plate tectonics to a dynamic one involves quantifying the forces that drive and resist plate motions. Which are the primary, or dominant, forces and their magnitude within this force balance remains a debated issue. This debate is central to this paper, as it means that different models, based on different assumptions regarding the forces that resist plate motion, have been proposed (3.2).

The range of proposed plate tectonic cooling models for the Earth differ significantly
Figure 3.1: Geologic proxy data for mantle potential temperature ($T_p$) throughout Earth’s history. We use [Ganne and Feng, 2017] as a constraint for our model.

in terms of physical assumptions, and each, therefore represents a different hypotheses regarding the dynamics of plate tectonics. The cooling rate associated with the convective overturn of tectonic plates depends on resisting forces to plate motion. The earliest plate tectonic cooling models assumed that the dominant resistance to plate motions comes from the viscosity of the Earth’s mantle - the rocky interior that plates move over and subduct into [Tozer, 1972, Schubert et al., 1979, 1980]. Later models argued that the strength of plates needed to be considered as plate deformation and deformation at plate boundaries provided significant energy dissipation [Conrad and Hager, 1999a,b]. Those models assumed that plate strength would decrease under hotter conditions, i.e., in the Earth’s past, or remain constant. That assumption was challenged by another plate tectonic cooling model that assumed plate strength increased in the Earth’s past [Korenaga, 2003, 2006]. All of these models remain argued for to this day (discussed in more detail in the next section) with different authors arguing with variable degrees of ’argumentative force.’ The fact that debate remains signals that there is no singular, agreed upon, plate tectonic cooling model, which has implications for modeling planetary habitability beyond Earth.
Models that couple interior planet cooling to climate evolution, seeking to address long
term habitability of terrestrial planets in general, consider the potential of different tectonic
modes, with one example being a plate-tectonic cooling model [Driscoll and Bercovici,
2013, Foley and Driscoll, 2016]. A misconception that can follow is that there is a singular,
agreed upon, plate tectonic cooling model. As noted above, and detailed in what follows,
this is not correct.

How different are proposed plate tectonic cooling models in effect? That is, are the
differences in terms of model outputs small relative to data uncertainty? Over the full range
of the models that have been proposed to date, they are not. This is clearly demonstrated in
the fact that the sign of the dominant feedback for planetary cooling varies from negative
to positive over the full range of proposed models [Moore and Lenardic, 2015, Seales et al.,
2019]; the dominant feedback determines whether plate tectonics is less (positive feedback)
or more efficient (negative feedback) at cooling the mantle at hotter temperatures. The
implications for extrapolating Earth cooling models to "Earth-like" terrestrial planets is
significant [Tozer, 1972, Korenaga, 2016].

To date, no study has systematically compared model outputs for the range of proposed
plate tectonic thermal history models to observational data in light of model uncertainties,
though some have considered uncertainty in specific contexts [McNamara and Van Keken,
2000, Korenaga, 2011]. The bulk of this paper sets out to provide such a comparison. First,
the comparison is carried out for model evolutions over the Earth’s geologic age. That
exercise will isolate models that are consistent with Earth data constraints. From there, we
will project this range of “successful” models forward in time to model Earth-like planets
older than the Earth. This will provide insights into the level of certainty that exists for
making statements regarding the thermal state of terrestrial planets assumed to operate in
a plate tectonic cooling mode, an issue of interest to the planetary habitability community.
3.3 Methods

In this section we will define the thermal history models we used in this analysis, define the model uncertainties we evaluated, outline the geologic proxy data we used as model constraints, and define how these were combined to assign probabilities of model success.

In principle, thermal history models can be formulated to solve for the full three dimensional evolution of a planetary interior over time [e.g., Zhong et al., 2000]. In practice, such formulations (run over the Earth’s full geological history) remain computationally expensive, which limits the degree to which model output space can be explored. For this reason, thermal history models of the Earth have been formulated to track the average internal temperature of the Earth, and the majority of thermal history models presented for the Earth are of this variety.

Thermal history models that track averaged internal temperatures are also referred to as parameterized thermal history models. Different parameterizations reflect different assumptions regarding the operation of plate tectonics (discussed more fully below). That difference being noted, thermal history models share a common underpinning: The Earth’s average mantle temperature evolves over time based on the balance between heat produced within \(H\) and lost from \(Q\) the mantle according to

\[
CT_p = H - Q. \tag{3.1}
\]

where \(T_p\) is the average mantle temperature \(T_m\) removing the effect of adiabatic heating (which does not drive thermal convection). Heat is produced within the mantle by the radiogenic decay of \(^{238}U\), \(^{235}U\), \(^{232}Th\) and \(^{40}K\), and heat production over time is given by

\[
H(t) = H_0 \sum_{n=1}^{4} h_n \exp(\lambda_n t), h_n = \frac{c_n p_n}{\sum_n c_n p_n} \tag{3.2}
\]

where \(H_0\) is a reference heat production, \(h_n\) is the amount of heat produced by a given isotope, and \(t\) is time. We calculate relative isotopic concentrations by assuming present day proportions of \(U : Th : K = 1 : 4 : (1.27 \times 10^4\) and normalizing by total U [Turcotte,
The values used in equation 3.2 are listed in Table 3.1.

<table>
<thead>
<tr>
<th>Isotope</th>
<th>$\rho_n$ (W/kg)</th>
<th>$c_n$</th>
<th>$h_n$</th>
<th>$\lambda_n$ (1/Ga)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{238}\text{U}$</td>
<td>$9.37 \times 10^{-5}$</td>
<td>0.9927</td>
<td>0.372</td>
<td>0.155</td>
</tr>
<tr>
<td>$^{235}\text{U}$</td>
<td>$5.69 \times 10^{-4}$</td>
<td>0.0072</td>
<td>0.0164</td>
<td>0.985</td>
</tr>
<tr>
<td>$^{232}\text{Th}$</td>
<td>$2.69 \times 10^{-5}$</td>
<td>4.0</td>
<td>0.430</td>
<td>0.0495</td>
</tr>
<tr>
<td>$^{40}\text{K}$</td>
<td>$2.79 \times 10^{-5}$</td>
<td>1.6256</td>
<td>0.181</td>
<td>0.555</td>
</tr>
</tbody>
</table>

Heat from the Earth’s metallic core could be more directly included in equation 3.1 by building in a core evolution model. For simplicity, and to be consistent with the bulk of previous Earth cooling models, we will not do so herein. Adding a core evolution model would increase model uncertainties, and in not doing so we will follow the approach of being conservative in our uncertainty assessment. We also leave out tidal heating as it is not a major effect in the Earth context. Including it as a heat source term could be an interesting extension of this work for planetary bodies such as those of the TRAPPIST-1 system, but this is outside the scope of our analysis.

Heat is lost from the planetary interior by convective cooling. This cooling is parameterized according to the Nusselt-Rayleigh scaling law given by

$$Nu \sim Ra^\beta$$ (3.3)

[Turcotte, 1992]. The Nusselt number $Nu$ is a nondimensional heat flux calculated as the ratio of convective ($Q$) to conductive ($q$) heat flux across the convecting layer. Conductive heat flux is given by: $q = \frac{k \Delta T}{D}$, where $k$, $D$ and $\Delta T$ are the thermal conductivity, convecting layer thickness and temperature drop across the convecting layer, respectively. The Rayleigh number $Ra$ is a nondimensional number that describes the vigor of convection and is defined as

$$Ra = \frac{\rho g \alpha \Delta T D^3}{\kappa \eta}$$ (3.4)
where \( \rho \) is mantle density, \( g \) is the acceleration due to gravity, \( \alpha \) is the thermal expansivity, \( \Delta T \) is the temperature between the surface and interior, \( \kappa \) is mantle diffusivity and \( \eta \) is mantle viscosity. Here we assume surface temperature is zero and thus \( \Delta T \) reduces to \( T_p \). To change Nusselt-Rayleigh scaling to an equivalency requires a constant \( a \) be added to the right hand side of equation 3.3. The value of \( a \) is dependent on the geometry of the convecting system and the average aspect ratio of convection cells. One can use laboratory experiments, boundary layer theory, and/or numerical simulations to constrain the constant [Davies, 1980, Schubert et al., 1980]. One may also scale the heat flow to present day heat flow \( Q_0 \) and employ a scaling temperature \( T_0 \) as was down by Christensen [1985] and Korenaga [2003] to fix the scaling constant \( a \) and arrive at a heat flow scaling given by

\[
Q = Q_0 \left( \frac{T_p}{T_0} \right)^{1+\beta} \left( \frac{\eta(T_0)}{\eta(T_p)} \right)^\beta.
\]  

(3.5)

Equation 3.5 is dependent on viscosity \( (\eta(T_p)) \) which is defined as

\[
\eta(T_p) = \eta_0 \exp \left( \frac{A}{RT_p} \right).
\]  

(3.6)

where \( A, R \) and \( \eta_0 \) are the activation energy, universal gas constant and scaling constant [Karato and Wu, 1993], respectively. We only consider diffusion creep in our study. Adding dislocation creep, which holds at upper mantle conditions while diffusion creep holds in the lower mantle, would add parameter uncertainty to our models. The methods we present could be used to specifically address rheological uncertainties. However, this is not our principal focus and is better suited for future work. For comparison to previous studies, we set \( \eta_0 \) so that the upper mantle has a viscosity of \( 10^{19} \) \( \text{Pa} \cdot \text{s} \) at 1350 °C. Combining equations 3.3-3.6 and using the definition of \( Nu \) leads to the governing equation

\[
C T_p = H_0 \sum_{n=0}^{4} h_n \exp (-\lambda_n t) - Q_0 \left( \frac{T_p}{T_0} \right)^{1+\beta} \left( \frac{\eta(T_0)}{\eta(T_p)} \right)^\beta.
\]  

(3.7)

Choosing the value for \( \beta \) in equation 3.7 involves making assumptions/hypotheses
Figure 3.2: The sources of plate resisting forces, effective $\beta$ value associated with different hypotheses regarding the dynamics of plate tectonics, and a thermal depth profile relating mid mantle temperatures ($T_m$) to mantle potential temperatures ($T_p$).

regarding the dynamics of plate tectonics (3.2). The earliest thermal history models used a value of 0.33 [Schubert et al., 1980, Spohn and Schubert, 1982, Jackson and Pollack, 1984]. This assumes that the dominant resistance to convective motion comes from mantle viscosity [Tozer, 1972]. It also assumes very vigorous convection. For levels of convection pertinent to the Earth the scaling exponent is slightly lower, $0.30 \leq \beta \leq 0.32$ [Schubert and Anderson, 1985, Lenardic and Moresi, 2003], due to the upper boundary of mantle convection (plates in a plate tectonic mode) not being fully self-determined (a self-determined boundary layer is not influenced by boundary layer interactions which can affect convective scalings [Moore and Lenardic, 2015]). Later models, that more directly incorporated model analogues to tectonic plates, showed that values nearly matching this scaling would be recovered provided that very weak plate boundaries were also incorporated [Gurnis, 1989]. Later models that allowed weak plate boundaries to develop dynamically lead to a scaling exponent of 0.29 [Moresi and Solomatov, 1998]. If plate
boundaries are not assumed to be so weak that energy dissipation along them can be neglected and/or if plate strength offers significant resistance, then the scaling exponent will be lower with a range between $0 \leq \beta \leq 0.15$ having been proposed [Christensen, 1985, Giannandrea and Christensen, 1993, Conrad and Hager, 1999b,a]. A low viscosity channel below plates - the Earth’s asthenosphere [Richards and Lenardic, 2018] - allows different size plates to have different balances between plate driving and resisting forces [Höink et al., 2011]. This leads to a mixed mode scaling in which plate strength is the dominant resistance for small plates while mantle viscosity is dominant for larger plates. Considering the distribution of current tectonic plate sizes as a guide, this leads to a global heat flow scaling exponent of $0.15 \leq \beta \leq 0.20$ [Höink et al., 2013]. An argument for $\beta < 0$ has also been made [Korenaga, 2003]. The physical basis for this last class of models is that at hotter mantle temperatures enhanced melting would generate a thicker dehydrated layer below oceanic crust. This layer would be responsible for the bulk of plate strength. By this reasoning, hotter mantle temperatures in Earth’s past would allow for a thicker, stronger plates, which would slow plate velocities and decrease the rate at which the mantle cooled.

Given that different $\beta$ values represent different physical assumptions regarding the dynamics of plate tectonics, and by association Earth cooling, it follows that different values of $\beta$ represent different hypotheses. This means that we can think of the choice of $\beta$ as a model selection problem, which introduces model selection uncertainty into our analysis. To account for this, we will assume the different models historically put forth are unique; however, we will allow for $\beta$ values between them to represent gradational changes between the different hypotheses. Specifically, we will test a range of models with $\beta$ values between -0.15 and 0.3 at intervals of 0.025. In doing so, our analysis will generate relatively smooth model probability distributions in $\beta$ space, allowing us to map peaks in $\beta$ space to determine models with the highest probability of matching data constraints subject to a variety of uncertainties.

For each $\beta$ model, we will also evaluate combined initial condition and parametric
uncertainty. The values for each are listed in Table 3.2. Initial condition uncertainty for thermal history models comes from uncertainties about post-magma-ocean planetary temperatures. Parametric uncertainty for thermal history models is connected to the values used for radiogenic heating, the heat flow scaling constant, and the scaling temperature. The strength of temperature dependent viscosity is also a model parameter that can be subjected to a range of values. For simplicity, we will not consider that explicitly herein, as it is connected to variations in the mantle Rayleigh number, $Ra$, which will already be subjected to a range of variations due to the variations in the other parameters noted.

The final type of uncertainty that will go into our cumulative uncertainty quantification is the uncertainty associated with unmodeled factors. This is referred to as model inadequacy [Kennedy and O’Hagan, 2001] within the discipline of uncertainty quantification. It is also referred to as structural uncertainty as it is connected to the structural stability of a model [Guckenheimer and Holmes, 1983]. The outputs from a structurally stable model remain qualitatively similar if the model is perturbed - the perturbations represent low amplitude, unmodeled factors. Structural stability testing can be accomplished using a perturbed physics approach [Astrom and Murray, 2008]. Such an approach can also provide a measure of structural uncertainty for models that do maintain structural stability [Seales et al. 2019]. Models with low structural uncertainty can damp perturbations/fluctuations associated with physical factors not directly incorporated into them. Figure 3.3 shows an example output of such an analysis, henceforth referred to as an ensemble, from a model subjected to a perturbed physics analysis (see Seales et al. [2019] for a full description of this method). For each ensemble, we will use two standard deviations as our uncertainty metric. In performing this analysis, we found that increasing the standard deviation of the perturbation set itself (i.e., the maximum amplitude of perturbations) did not significantly effect the accumulation of uncertainty provided that the perturbations remained randomized in time and of an amplitude below a few percent - an assessment of the uncertainty associated with the particular uncertainty metric itself. Including this as well as all other forms of uncertainty, our analysis involved computing more than 1.25
millions of model evolutions.

Figure 3.3: Ensemble model runs for a single thermal history model (a) and an uncertainty measure for a single ensemble (b). An ensemble of 100 perturbed paths (gray lines) is plotted in (a), along with the ensemble mean (dark green line) and ensemble two standard deviation limits (lighter green lines). In (b) a present day time slice is taken through the ensemble evolution. This ensemble has an approximately 50% probability of satisfying present day constraints, which are temperature (1300°C-1400°C) and a Urey ratio between 0.2 and 0.5. Urey ratios are listed are labeled for the the mean, uncertainty limit and acceptable bound solutions.

The success or failure of an ensemble will be determined by comparing the ensemble mean – the ensemble mean and the unperturbed model solution are equivalent if the model is structurally stable – and two-standard deviation bounds to paleo and present day constraints. For paleo constraints we use the results of Ganne and Feng [2017] who calculated uncertainty bounds on mantle temperatures over time (3.3). They derived these bounds by using the MgO content of approximately 22,000 samples of mafic and ultramafic extrusive basalts, and calculating the potential temperature associated with these melts using PRIMELT [Herzberg and Asimow, 2015] with different assumed mantle redox conditions. According to Figure 10 of Ganne and Feng [2017], the maximum and minimum temperature estimates at 3.5 Ga are 1666 °C and 1500 °C. These values drop
to maximum and minimum of 1553 °C and 1321 °C, respectively, at 100 Ma. Ganne and Feng [2017] also calculated temperature range estimates at intermediate ages and provided extrapolations that bounded the estimates over geologic history (plotted as black lines in Figure 3.1). Ganne and Feng [2017] suggested the maximum and minimum bounds may represent the temperature of plumes and ambient mantle, respectively. That amounts to a model assumption. For the purpose of this study we will not impose any a priori model assumptions regarding the data. Rather, we will consider the data to represent the time dependent dispersion in paleo-temperature constraints (the data spread, under this interpretation, represents observational uncertainty). This is consistent with the work of Condie et al. [2016] and Herzberg and Asimow [2008], both of whose data sets are also shown in Figure 3.1. Both Condie et al. [2016] and Herzberg and Asimow [2008] considered their data to represent ambient mantle (which would correlate to \( T_p \) for thermal history models). The dispersion of the data in Ganne and Feng [2017] encompass both Condie et al. [2016] and Herzberg and Asimow [2008], therefore using those two data sources as constraints for our models, as opposed to the extrapolations of Ganne and Feng [2017], would produce similar results. For our present day temperature constraint we use a value of 1350 °C ±50 °C [Herzberg and Asimow, 2008].

A second present day constraint is the mantle Urey ratio, \( U_r \), which is the the ratio of \( H \) to \( Q \). Jaupart et al. [2015] estimate it to be between 0.3 and 0.5. Allowing for continents, the \( U_r \) upper bound can be extended [Grigné and Labrosse, 2001, Lenardic et al., 2011]. Lenardic et al. [2011] show that, for present day continental land fractions, heat flows are consistent for mantles with and without continental coverage. The argument is that continental insulation warms the mantle below it and this effect is transmitted to the entire mantle as it is assumed to be well mixed. Increased temperatures lead to increased plate speed and more efficient cooling of oceanic mantle, offsetting the insulating effect. However, in the continental case, adding the heat producing elements from the continents back into Earth’s mantle would increase heat production and therefore \( U_r \) would also increase by about 0.2. Therefore, the upper \( U_r \) bound, for models that do not directly
include continental effects, can be extended to approximately 0.7 as an adjustment to account for present day continental effects. We will consider model success with and without a continental adjustment.

Using the constraints above, we now define the ensemble probability for successful models. This involves identifying the upper and lower most bounds on the ensemble probability distribution that fall within constraints and calculating the probability that an ensemble member falls between these two points. For example, in Figure 3.3b the mean of the ensemble is \( \sim 1380 \, ^{\circ}\text{C} \). The upper temperature bound occurs at \( 1400 \, ^{\circ}\text{C} \), where the present day \( U_r \) is 0.45, within present day constraints. The lower temperature bound is not set to \( 1300 \, ^{\circ}\text{C} \) because at this temperature \( U_r \) is greater than 0.5. An \( U_r \) value of 0.5 occurs at \( 1365 \, ^{\circ}\text{C} \). Therefore, for this ensemble, any output temperature between 1365 and 1400 \( ^{\circ}\text{C} \) (the hachured region in 3.3b) satisfies present day constraints with a probability of 0.5. This hachured region, then, is the fraction of models within this ensemble that can match present day constraints.

### Table 3.2: Model Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Values</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T_i )</td>
<td>1000, 1250, 1500, 1750, 2000</td>
<td>(^{\circ}\text{C})</td>
<td>Initial Temperature</td>
</tr>
<tr>
<td>( T_0 )</td>
<td>1300, 1350, 1400</td>
<td>(^{\circ}\text{C})</td>
<td>Scaling Temperature</td>
</tr>
<tr>
<td>( Q_0 )</td>
<td>3.0e13, 3.5e13, 4.0e13</td>
<td>TW</td>
<td>Scaling Heat Flow</td>
</tr>
<tr>
<td>( H_0 )</td>
<td>2.19e13, 2.55e13, 2.92e13, 4.38e13, 5.12e13, 5.84e13, 6.57e13, 7.66e13, 8.76e13, 1.02e14, 1.09e14, 1.17e14, 1.28e14, 1.46e14</td>
<td>TW</td>
<td>Initial Radiogenics</td>
</tr>
<tr>
<td>( \eta_0 )</td>
<td>2.21e9</td>
<td>Pa\cdot\text{s}</td>
<td>Viscosity constant</td>
</tr>
<tr>
<td>A</td>
<td>300</td>
<td>kJ mol(^{-1})</td>
<td>Activation Energy</td>
</tr>
<tr>
<td>R</td>
<td>8.314</td>
<td>J / (mol\cdot K)</td>
<td>Universal Gas Constant</td>
</tr>
</tbody>
</table>

### 3.4 Results

Figure 3.4 shows mean ensemble cooling paths for models with different \( \beta \) values, input parameter values, and initial conditions. We leave off the ensemble uncertainty bounds
for ease of viewing, but they were calculated for all the plotted ensembles. The mean ensemble paths that satisfy the present day $T_p$ constraints are shown as red lines. Mean ensemble paths that fall outside the constraint, but are associated with models that can match the constraint within structural uncertainty bounds (i.e., within ensemble bounds), are shown as light red lines. Solutions that do not match the constraint, even allowing for ensemble uncertainty bounds, are shown as grey lines. A model with $\beta < 0$ is very initial condition and input value dependent. This leads to a wide model solution space. Models with $\beta \geq 0$ had weaker initial condition and input value dependencies, resulting in a more concentrated solution space.

The number of cases that satisfy the present day $T_p$ constraint for variable $\beta$ are shown in Figure 3.5. The number of cases where the mean matches the present day constraint (darker green) is small at the most negative $\beta$ endmember. The number of mean ensemble paths remained below 10% until $\beta$ became positive and the number of successful cases began to grow. At a $\beta$ value near 0.2 the number of successful cases plateaued around 30%. Accounting for structural uncertainty increases the number of successful cases for all $\beta$ values (lighter green). The rise in successful cases occurred while $\beta$ was still negative, around $\beta = -0.1$, and plateaued at a $\beta$ slightly greater than 0.1.

The number of cases matching present day $U_r$ are shown in Figure 3.6. The color scheme is the same as Figure 3.5 with mean solutions in darker green and those that include structural uncertainty in lighter green. Results are shown for cases that match present day $U_r$ without accounting for the effect of present day continental distribution (lighter green) and for cases in which the effect of continents, on the present day Urey ratio [Grigné and Labrosse, 2001, Lenardic et al., 2011], is accounted for (lightest green). The distribution of successful cases peaked around 60% for $\beta = 0.05$. Accounting for structural uncertainty had little effect for the $U_r$ constraint. Accounting for continents increased the number of successful cases. At its peak, near $\beta = 0.1$, the number of successful matches was greater than 90%. Including a continental effect disproportionately benefited more positive $\beta$ values.
Figure 3.4: Mean ensemble paths from a range of models. Mean ensemble paths are classified into three groups – those that satisfy an observational constraint (red), those that satisfy it within structural uncertainty (light red), and those that do not satisfy the constraint (gray).
Figure 3.5: Probability distribution for models that satisfy the present day temperature constraint.

Figure 3.6: Probability distribution for models that satisfy the present day Urey ratio constraint.

Figure 3.7 shows the number of successful cases when assigning equal weight to the present day $T_p$ and $Ur$ constraints. The distribution is non-normal. Mean ensemble paths
resulted in less than 10% of successful cases across the board. The peak for the mean solutions is at a $\beta$ value slightly less than 0.1. Below this value successful models fall to nearly zero before increasing slightly when values of $\beta < 0$ were considered. Increasing the upper $Ur$ bound to 0.7, to account for the potential effect of continents, shifted the peak $\beta$ value to be greater than 0.1 and increased the number of successful cases to nearly 60% at the peak. Considering structural uncertainty preferentially benefited the lower half of the tested $\beta$ space. A very low percentage of models could match both constraints for $\beta$ values greater than 0.2 unless the effects of continents were considered (and it should be kept in mind that doing so adds its own layer of uncertainty as the continental correction comes from models [Grigné and Labrosse, 2001, Lenardic et al., 2011]).

Figure 3.7: Probability distribution for models that satisfy the present day temperature and Urey ratio constraint.

The distributions that resulted from using only paleo temperature constraints are shown in Figure 3.8. The trends are similar to those in Figure 3.5. One difference is the uniform decrease in the fraction of ensembles able to match the paleo constraints. This intuitively makes sense in that to be successful an ensemble must stay within a temperature window over an extended evolution time rather than match a value at a single time. A
subtle, but noteworthy difference between Figure 3.5 and Figure 3.8 is that the fraction of successful ensembles matching the paleo constraints increased to a greater degree as $\beta$ was increased. Positive $\beta$ models tend to lessen initial condition dependence. Nearer to the model start time there is less time to eliminate the influence of the initial condition. As a result, some models that converge to present day temperatures were too hot or too cold at 2.5 Gyr and thus considered unsuccessful. Even with this change in slope, the distribution peaked around 0.2.

Figure 3.8: Probability distribution for models that satisfy the paleo temperature constraints.

Figure 3.9 shows models that can match paleo and present day constraints. Figure 3.9(a) shows the fraction of models for each $\beta$ that have some portion of the ensemble that satisfies all three constraints. Distributions are bi-modal, having one peak in the negative $\beta$ domain and one in the positive domain. Accounting for structural uncertainty increased the fraction of successful solutions across the board and produced nearly identical peaks in both the positive and negative domains. Allowing for continental effects shifted the largest peak close to a $\beta$ value of 0.2, but a sub-zero peak remained. A representation of the total probability is shown in Figure 3.9(b). For each $\beta$, the total probability is the sum
In (b) we tabulate the total fraction of models that match all constraints by calculating the probability of success of each ensemble as shown in Figure 3.3 of each ensemble probability divided by the total number of initial conditions and model input combinations assessed. For a constraint on present day \( U_r \) that does not account for continents, a peak probability of approximately 10% occurred at a \( \beta \) value of 0.1. This distribution has a single peak with a heavy left tail, which is caused by the hard upper \( U_r \) limit of 0.5 that cast out a large portion of the more positive \( \beta \) ensembles. Relaxing this constraint resulted in more normal distribution peaked around 0.15. This is close to the value argued for by Conrad and Hager [1999a]. We have given all data constraints equal weight. If one of the constraints is found to be more reliable than the others, then the distribution peak will shift towards the \( \beta \) values that coincide with matching that constraint.

Using only the mean ensemble paths that matched paleo and present day constraints, we projected mantle potential temperature out to 10 Gyr (Figure 3.10). Figure 3.10(a) projects only those mean ensemble paths that matched Earth constraints (darker green models in Figure 3.9(a)). The differing feedbacks within the models become apparent as time evolves.
with negative $\beta$ models (positive feedback) reaching far cooler mantle temperatures. These models lead to cooling runaways and once temperatures drop too low the models are cut off as they have lost structural stability, that is small perturbations/fluctuations significantly affect model evolution, pulling the perturbed solution far from the unperturbed solution [Seales et al., 2019]. Models with $\beta > 0.1$ cooled more slowly, maintaining temperatures above 1000 °C throughout. Figure 3.10(b) shows the projected models that match paleo and present day constraints with structural uncertainty now accounted for. Projections were limited to those models that matched $Ur$ values between 0.3 and 0.5. Including structural uncertainties allowed for run away cooling behavior to occur nearly one billion years nearer to present day for models with the most negative $V$ values. If we take into account the total probabilities, which peak between $\beta$ values of 0.1 and 0.2, and only use those cases, then projected temperature vary between 1000 and 1200 °C at 10 Gyr of model evolution. However, as each of the mean ensemble paths plotted match Earth constraints, they all remain possible. Stated another way, there is no reason why the evolution path of a particular planet, the Earth, needs to follow a most probable path within a model solutions space.

### 3.5 Discussion

Our analysis considered multiple forms of uncertainty to assess the probability that any given thermal history model fits Earth constraints. Any model with a probability greater than zero is viable. One model being less probable than another, in model solution space, does not eliminate the possibility that the lower probability model captures the essential physics of plate tectonics, as related to planetary cooling. Having said that, we can also weigh probabilities to assess which models can match Earth constraints over the widest range of uncertainties. Figure 3.9(b) indicates models with $\beta$ between 0.1 and 0.2 fall into this category. High $\beta$ models can match present day temperature over a wide input range but struggle to match the lower $Ur$ constraint. Lower $\beta$ models can match the $Ur$ constraint but struggle with present day temperature constraints if $\beta$ drops too low as they then run
Figure 3.10: Mantle water volume (a) and thermal (b) paths that matched all constraints. In (b) we show paleo thermal history estimates that suggest a multi-phase thermal history along with successful thermal paths for qualitative comparison with the output of our analysis.

hot [McNamara and Van Keken, 2000]. That a “sweet spot” could exist between the two end-members is not, in hindsight, qualitatively surprising.

Uncertainties in the data constraints we used influenced the calculated probability of successful models. We assumed equal weighting for each constraint. Of the two present day constraints, the present day mantle temperature is a harder constraint. This is because there is less uncertainty in estimating its value than there is in estimating the Urey ratio Jaupart et al. [2015]. The difficulty of considering different weightings is that, although the distribution of uncertainties associated with temperature data has been calculated [Condie et al., 2016, Ganne and Feng, 2017], the same is not true for the Urey ratio. At this stage, we did not consider it warranted to apply different weightings but this could be done in the future.

Our analysis explored a slice of potential model space. A more extensive exploration would change quantitative results but key qualitative results are likely to be robust. The qualitative differences between positive and negative β models comes from the fact that the former is dominated by a negative system feedback and the latter by a positive feedback
[Moore and Lenardic, 2015]. More sophisticated models can be constructed (e.g., fully 3-D models) but the dominant feedback will still dictate end-member behavior and uncertainty structure. Uncertainties will be smaller in models dominated by negative feedbacks, as will be model solution space. The latter means that model cases may be less likely to match any given data constraints, but if they can match constraints, then a narrow solution space will lead to a larger percentage of model ensembles being successful. Models with high positive feedbacks will have greater uncertainty and an associated larger solution space. A large solution space increases the potential that at least some cases can match a given data constraint and, at the same time, it favors a smaller percentage of potential model ensembles being successful.

The connection between uncertainty and successful models relates to another conclusion we argue is robust: Models based on different hypotheses, regarding the dynamics of plate tectonics, are consistent with constraints on the Earth’s thermal history, i.e., competing hypotheses remain viable. Phrased another way, model and data uncertainties lead to ambiguity - more than one model is viable. Considering more sophisticated (complex) models will not, we argue, change this conclusion, provided full model uncertainties are assessed. Increasing complexity can increase model uncertainty [Saltelli, 2019]. More complex models come with more parameters and assumptions which increases both parametric and model selection uncertainty [Saltelli, 2019], as well as the potential to overfit data. That can increase the number of potential model solutions and the computational time needed to find them. It can also greatly increase the time and work load needed to quantify model uncertainty. More complex models may be able to better match data constraints but this should not be confused with the models being more certain. The ability of a model to match constraints is not the same thing as a model’s uncertainty. Model uncertainty can, however, affect the ability of a model to match constraints. More uncertain models are associated with a larger potential model solution space. A larger solution space increases the potential that some combinations of model inputs, initial conditions, and ensemble paths will match constraints. It is possible that new and/or more certain data
constraints could bridle this to a degree, though historical data from the Earth will always have uncertainty. As such, we argue that multiple hypotheses will likely remain viable into the near future, particularly if there is a trend toward developing more complex models.

Model ambiguity in Earth science is not new [e.g., Richards and Lenardic, 2018] and it has been considered for endmember cases in thermal history modeling [Korenaga, 2008]. Accounting for model uncertainties extends the range of model ambiguity such that multiple hypotheses, regarding the dynamics of planetary cooling, can be consistent with data constraints. Hypothesis discrimination can continue, but it must proceed in a statistical manner. This, we argue, is another robust conclusion. We can ask which models come with higher probabilities of success in light of uncertainties. This is the utility of Figure 3.9. The degree to which one is willing to push this further depends on a question that cannot be answered at present: of all the possible evolution paths, consistent with physical and chemical principles, did a single planet, the Earth, follow what is the most likely path in that potentiality space? The conservative stance is to say 'We don’t know,' which means we consider all models with greater than zero probability as viable.

The question above relates to the extension of thermal history studies from Earth to planetary application, habitability in particular [Kite et al., 2009, Schaefer and Sasselov, 2015, Komacek and Abbot, 2016, Foley, 2015, Foley and Driscoll, 2016, Tosi et al., 2017, Foley and Smye, 2018, Barnes et al., 2020]. Thermal history models applied to the Earth are postdictive: they set out to match historical data. In the context of habitability studies, thermal history models are used in a predictive mode to determine whether liquid water may be present on the surface of terrestrial planets with variable planetary and orbital properties. These predictions are made by calculating the flux of volatiles from the interior of the planet to the surface using mantle temperatures along with melting modules and testing whether surface water can persist over time scales that allow life to develop. Using models in a predictive mode increases the potentiality space of model outputs. With the thought of limiting the vastness of this space, many studies have focused on planets similar to Earth in size and composition as a starting point [e.g., Foley, 2015, Foley and Driscoll,
Implicit to this is the thought that uncertainties will be lowest for modeling this subset of planets. Our analysis suggests that even if we consider the most Earth-like planet possible, with the most observational data (the Earth), significant uncertainty remains (3.9).

The above leads to a few suggestions on moving forward. First, even if we focus on a plate tectonic mode of planetary cooling, we should consider all viable models (Figure 3.9). To date, habitability models have considered a single plate tectonic model [e.g., Driscoll and Bercovici, 2013, Foley, 2015, Foley and Driscoll, 2016, Rushby et al., 2018]. This bypasses model selection uncertainty. Second, all models should be subjected to uncertainty quantification. Typically only a range of initial conditions and input values are tested. An ensemble approach is generally not employed, which leaves out structural uncertainty. One uncertainty measure is not a substitute for another and all need to be evaluated before model implications can be assessed and/or before a model can be validated. A corollary is that model implications need to be viewed in a probabilistic manner by presenting results as probability distributions. This becomes particularly important for models used to make forecasts.

All of the projections in Figure 3.10(b) should, we argue, be considered as potentialities. In that view, they are all counterfactuals [Taleb, 2011] with very different implications if used as forecasts. One such forecast is to estimate when melting ceases, likely leading to the shutdown of plate tectonics; however, caution must be used as a transition out of plate tectonics to another convective regime, such as stagnant lid, could occur prior to when melting shuts off. With that in mind, we consider a ballpark value of 1000 °C to indicate when melting ceases. An example prediction our analysis makes is that a family of ensembles imply that plate tectonics could end in about 1.5 billion years as the mantle becomes too cold. This family of ensembles is consistent with a study that did forecast the end of plate tectonics in 1.45 billion years [Cheng, 2018]. Such a forecast has implications for life beyond Earth. The fact that some of our projections are in line with the study of Cheng [2018] speaks to model reproducibility, as that study used a
negative β model, which is also the one we found leads to cold runaways. However, it is 
the negative β models that are associated with the largest uncertainty and are prone to 
structural instability Seales et al. [2019]. Not being clear about uncertainty, especially for a 
provocative conclusion, only invites misinformation (e.g., presenting a highly uncertain 
model forecast as a singular "result"). We would suggest that if full uncertainty analysis 
was as strong a component of planetary modeling studies as, for example, a methods 
section, then the odds of unintentionally making conclusions that can send misinformation 
would be reduced. Uncertainty quantification also has the potential to prevent red-herring 
debates of the type that have surfaced in the past [Valencia et al., 2007a, O’Neill and 
Lenardic, 2007].

In the exoplanet modeling field, thermal history models are being coupled to other 
models to explore how interior planet evolution co-evolves with other systems – stellar, 
orbital, volatile cycling, climate, weathering and life [Barnes et al., 2020]. Each system sub-
model is subject to the types of uncertainty we have presented for thermal history models, 
making the full model potentiality space large. This can make a grid search approach, to 
map out the coupled model solution space in light of uncertainties, intractable. However, 
the full model potentiality space is often not of primary interest. A more primary driver 
behind the coupled models is mapping the subspace that allows water to exist at the surface 
of a planet over geologic time (this connects the models to the search for life beyond Earth 
– life as we know it relies on water). Having a search target, within model potentiality 
space, can reduce the computational work load, but a grid search, akin to that of this paper, 
would still be impractical given the large dimensionality of the problem. More efficient 
computational methods can bring the modeling back to a tractable level (e.g., machine 
learning based methods [P. Fleming and VanderPlas, 2018]). This will introduce further 
uncertainty that will need to be evaluated – the uncertainty associated with the particular 
search method. All of this will increase the workload and the move toward a probabilistic 
framework. Such a framework, in turn, would move the field from a binary assessment of 
habitability toward an assessment that provides probability distribution functions for the
potential habitability of exoplanets (functions that can be progressively adjusted in light of new observations).

3.6 Conclusions

We applied an uncertainty analysis to solid Earth cooling models. The analysis accounted for the combined effects of: 1) Model selection uncertainty; 2) Model structural uncertainty; 3) Uncertainty in initial conditions; 4) Uncertainty in model input values. Accounting for model and observational uncertainties allows for model validation (testing the degree to which model outputs can match data constraints). Validation, once full uncertainty measures are evaluated, requires a probabilistic approach and results are presented as probability distributions. Given we only have one planet evolutionary path, the Earth, we have argued that any models that maintain finite probabilities of accounting for observational data, over model potentiality space and in light of uncertainties, remain viable. For the thermal history models we examined this leads to ambiguity. That is, multiple hypotheses, differing significantly in fundamental assumptions and implications, remain viable for the Earth’s thermal history. When thermal history models move from a postdictive mode (accounting for existing Earth data) into a predictive mode designed to constrain conditions that allow for clement surface environments on terrestrial planets, uncertainty quantification can become all the more critical as model results may then be guiding target selection for planetary observations and/or aiding in data interpretation.
Chapter 4

Deep Water Cycling and the Multi-Stage Cooling of The Earth

4.1 Abstract

Paleo-temperature data indicates that the Earth’s mantle did not cool at a constant rate over geologic time. The data are consistent with slow cooling from 3.8 to 2.5 billion years ago with a transition to more rapid cooling extending to the present. This has been suggested to indicate a change in global tectonics from a single plate to a plate tectonic mode. However, a tectonic change may not be necessary. Multi-stage cooling can result from deep water cycling coupled to thermal mantle convection. Melting and volcanism removes water from the mantle (degassing). Dehydration tends to stiffens the mantle, which slows convective vigor and plate velocities causing mantle heating. An increase in temperature tends to lower mantle viscosity which acts to increase plate velocities provided that mantle viscosity offers resistance to plate motion. If these two tendencies are in balance, then mantle cooling can be weak. If the balance is broken, by a switch to net mantle rehydration, then the mantle can cool more rapidly. We use coupled water cycling and mantle convection models to test the viability of this hypothesis. Within model and data uncertainty, the hypothesis that deep water cycling can lead to a multi-stage Earth cooling is consistent with present day and paleo data constraints on mantle cooling. It is also consistent with constraints that indicate a change from net mantle dehydration to rehydration over the Earth’s geologic evolution. Probability distributions, for successful models, indicate that plate and plate margin strength play a relatively minor role for resisting plate motions relative to the resistance from interior mantle viscosity.
4.2 Introduction

The rock record helps us unravel the Earth’s geologic history and also its thermal history, i.e., interior cooling over geologic time. Volcanic rocks can be used to estimate mantle temperatures at the time they formed. Figure 4.1 shows results from several studies that provide constraints on the cooling history of the Earth’s mantle [Ganne and Feng, 2017, Condie et al., 2016, Herzberg et al., 2010]. Uncertainty in the paleo data allows for a range of cooling paths. Present day constraints on mantle temperature and the ratio of radiogenic heat generation to mantle heat flow can narrow the range, but multiple model paths remain viable. This being acknowledged, the data trends are suggestive of multi-stage cooling. More specifically, data constraints are consistent with the hypothesis that the mantle experienced a change in cooling slope between 2 and 3 billions of years ago [C16,H10]. That possibility is bolstered by independent studies that indicate changes in the Earth system occurred within the same time window [e.g., Parai and Mukhopadhyay, 2018, Lee et al., 2016, Lyons et al., 2014].

A conceptual hypothesis for a change in mantle cooling rate at 2.5 billion years ago (Ga) invokes a change from a single plate mode of mantle convection (i.e., stagnant lid convection) to a plate tectonic mode [Condie et al., 2016]. To date, thermal history models have not been used to determine the degree to which the hypothesis can match data constraints. Doing so faces challenges from a geodynamical perspective. At present we do not know how long transitions from one convective regime to another would take, over the full range of parameter conditions pertinent to a planets evolution path, and/or the efficiency of mantle cooling through the transition [e.g., Weller and Kiefer, 2020]. If transitions are as long as Weller and Kiefer [2020] argued for (potentially a billion year time scale), then globally averaged heat balance models would be inapplicable and 3-D numerical simulations would be needed to map thermal histories. Running such models over geologic time scales is computationally intensive and requires large amounts of wall time, particularly if we also require model uncertainty quantification. This is not insurmountable, but it is not the only option.
Invoking tectonic transitions may be an intuitive way to account for changes in the Earth’s cooling rate. It is not, however, the only one. There are alternative hypotheses that do not invoke tectonic transitions. That class of hypotheses does not violate critical assumptions of globally averaged thermal history models; i.e., parameterized thermal history models [Schubert et al., 1980, Davies, 1980]. Such models allow for efficient hypothesis testing as mapping of parameter space, and associated uncertainty quantification, is not subject to computational and wall time restrictions that come with 3-D numerical simulations. One specific alternative invokes increasing plate strength in the Earth’s past [Korenaga, 2003]. That hypothesis assumes that the resistance to plate motion comes from plate strength. This is a break from classic thermal history models which are based on the

Figure 4.1: Estimates of Paleo mantle potential temperature.
idea that mantle viscosity dominates resistance to plate motion [Tozer, 1972]. Thermal history models based on increasing plate strength in the past have been used to show that the hypothesis can lead to a multi-stage thermal history. The change in mantle temperature slope over time they predict is more extreme than a change in cooling slope at ∼2.5 Ga. The models predict a change in the sign of the slope with the mantle heating before the transition and cooling after it. Less extreme versions of this hypothesis, that still assume plate strength dominates resistance to plate motion but not that it increases in the past, can also lead to changes in cooling rate over time [Christensen, 1984, Conrad and Hager, 1999a].

A systematic study that ran over one million thermal history models, with variable assumptions as to plate resisting forces, showed that models invoking plate strength as dominating resistance to plate motion showed peaks in the probability distribution of model cases that could account for observational constraints [Seales and Lenardic, 2020]. That is, such models are successful (the principal objective measure of model success being its ability to account for the observations it sets out to model). However, the physical validity of the plate resistance parameterizations used within such models has not been confirmed [Gerardi et al., 2019]. The physical validity of a parameterization that assumes internal viscosity offers the dominant resistance to convective and associated plate motion has, on the other hand, been confirmed via experiments [e.g., Giannandrea and Christensen, 1993] and numerical simulations [e.g., Schubert and Anderson, 1985, Lenardic and Moresi, 2003, Gurnis, 1989]. This does not over-ride the ability of plate strength models to account for observations. It does, however, suggest the question of whether a class of hypotheses that does not hinge on plate strength resisting plate motion can also lead to multi-stage cooling.

Classic thermal history models assume that mantle viscosity resists plate motion and that mantle viscosity is a function of temperature. This leads to a strong negative feedback [Tozer, 1972]. As a result, the thermal paths for such models rapidly settle on a near constant cooling path over geologic time [Schubert et al., 1979]. However, mantle viscosity
also depends on hydration [e.g., Karato and Wu, 1993]. Mantle hydration effects allow for water cycling feedbacks to interact with thermal feedbacks and variations in the strength of the feedbacks allows for different mantle cooling rates [Crowley et al., 2011, Sandu et al., 2011]. Resistance to plate motion still comes from mantle viscosity, but the viscosity is no longer determined solely by a thermal feedback.

In this paper we will test the ability of a deep water cycling hypothesis to account for thermal history data constraints with a focus on how changes in coupled water cycling and thermal feedbacks can lead to multi-stage cooling. More specifically, we will explore the hypothesis that a change from net mantle dehydration to net mantle rehydration can cause an increase in mantle cooling rates consistent with paleo data constraints. This specific add on is based on the feedback analysis of Crowley et al. [2011]. Those authors argued that a net dehydrating mantle could drive a component of mantle heating, which could compete with thermally driven cooling. They showed the opposite for a net rehydrating mantle. The hypothesis is also based on a geochemical argument that showed the Earth has transitioned from net dehydrating mantle to net rehydrating mantle over geologic time [e.g., Parai and Mukhopadhyay, 2018]. In the next section we define our model and discuss the model’s feedback structure. We then present model results and discuss implications for the Earth’s coupled thermal and deep water cycling history.

4.3 Methods

In this section we define the coupled model we use, discuss the observational data constraints applied to model outputs, and provide an overview of the feedback structure associated with this model.

4.3.1 Coupled Deep Water Cycling and Thermal History Model

Figure 4.2 shows a cartoon of how our model works conceptually. Plate generation and subduction cools the interior mantle and also cycles water between mantle and surface reservoirs. Mantle viscosity, which effects the vigor of mantle convection and associated
mantle cooling, depends on both temperature and mantle hydration. This leads to coupled thermal and water cycling feedbacks on mantle viscosity and, by association, mantle cooling. Variations in the strength of each feedback over geologic time allows for the potential of differing cooling efficiencies. We lay out the coupled model starting with the thermal component, then moving to the water cycling component, and then defining the mantle viscosity function that provides a coupling between the two.

Figure 4.2: Schematic of the coupled thermal and deep water cycle model.

**Thermal Component**

We used a parameterized thermal history model [Schubert et al., 1979, 1980] to track how Earth’s mantle temperature \( T_m \) changed with time. This model is a global energy balance expressed as

\[
p_CV\dot{T}_m = -3Aq_m + VQ(t)
\]  

(4.1)
where $\rho$ is mantle density, $C$ is mantle heat capacity, and $q_m$ is the mantle surface heat flux. Mantle volume, $V$ and surface area $A$ are calculated as $R_m^3 - R_c^3$ and $R_m^2 - R_c^2$, respectively. $R_m$ is the radial distance from Earth’s center to the surface, and $R_c$ is the radial distance to the Core-Mantle-Boundary (CMB). We define the Urey ratio ($Ur$) here

$$Ur = \frac{VQ}{Aq_m},$$

(4.2)

since we use it later as a model constraint.

The decay of radiogenic elements produces heat within the the mantle ($Q(t)$) according to

$$Q(t) = Q_0 e^{-\lambda t},$$

(4.3)

where $Q_0$ and $\lambda$ are constants, and $t$ is time in millions of years.

The Rayleigh number ($Ra$) is the ratio of forces driving convection to those resisting it. It is defined as

$$Ra = \frac{g\alpha \Delta T Z^3}{\eta \kappa},$$

(4.4)

where $g$, $\alpha$, $Z$, $\eta$ and $\kappa$ are gravity, thermal expansivity, depth of the convecting layer, kinematic viscosity and thermal diffusivity, respectively. The value $\Delta T$ is the temperature difference driving convection defined as $T_m - T_s$, where $T_s$ is surface temperatures. In parameterized thermal history modeling heat flux from the mantle is typically solved for using the Nusselt-Rayleigh scaling [Schubert et al., 2001], where the Nusselt number ($Nu$) is a nondimensional heat flux. The scaling takes the form

$$Nu = \frac{q_m Z}{k \Delta T} = \left(\frac{Ra}{Ra_{cr}}\right)^{\beta}.$$

(4.5)

Equation 4.5 is used to solve for $q_m$. In Equation 4.5, $k$ is thermal diffusivity, $Ra_{cr}$ is the critical Rayleigh number, which determines the onset of convection, and $\beta$ is a scaling exponent.

The choice of $\beta$ in Equation 4.5 has a rich history. The earliest thermal history models
used a value of 0.33 [Schubert et al., 1980, Spohn and Schubert, 1982, Jackson and Pollack, 1984]. This assumes that mantle viscosity provides the dominant resistance to plate motion [Tozer, 1972]. It also assumes very vigorous convection. For levels of convection pertinent to the Earth the scaling exponent is slightly lower, \(0.30 < \beta < 0.32\) [Schubert and Anderson, 1985, Lenardic and Moresi, 2003, Moore and Lenardic, 2015]. Models that more directly incorporated analogues to tectonic plates, showed that values nearly matching this scaling would be recovered provided that weak plate boundaries were also incorporated [Gurnis, 1989]. Later models that allowed weak plate boundaries to develop dynamically lead to a scaling exponent of 0.29 [Moresi and Solomatov, 1998]. If plate boundaries are not so weak that energy dissipation along them can be neglected and/or if plate strength offers significant resistance, then the scaling exponent has been argued to be lower with a range between \(0 < \beta < 0.15\) having been proposed [Christensen, 1985, Giannandrea and Christensen, 1993, Conrad and Hager, 1999b,a]. A low viscosity channel below plates - the Earth’s asthenosphere - allows different size plates to have different balances between driving and resisting forces [Crowley et al., 2011, Höink et al., 2011]. This leads to a mixed mode scaling. For the current distribution of plate sizes, the mixed mode leads to a global heat flow scaling exponent of 0.20 [Höink et al., 2013]. Korenaga [2003] made an argument for \(\beta < 0\). The physical basis for \(\beta < 0\) is that at hotter mantle temperatures enhanced melting would generate a thicker dehydrated layer below oceanic crust. This layer would be responsible for the bulk of plate strength. By this reasoning, hotter mantle temperatures in Earth’s past would allow for thicker, stronger plates, which would slow plate velocities and decrease the rate at which the mantle cooled.

As noted in the introduction, our goal is to explore models that do not rely on plate strength and/or plate margin strength providing the dominant resistance to plate motions. As such, we will not consider models with \(\beta < 0.20\). We will not, however, assume an \textit{a priori} preferred \(\beta\) value. Rather, we will test a range of models with \(0.2 \leq \beta \leq 0.33\) at intervals of 0.01.

In defining a velocity scale \((u_c)\), we first rearranged Fourier’s law to determine litho-
spheric thickness \( (D_b) \)

\[
D_b = k \frac{(T_b - T_s)}{q_m} .
\]

(4.6)

In Equation 4.6 \( k \) is thermal conductivity. The temperature \( T_b \) is the temperature at the base of the lithosphere. Using boundary layer theory [Schubert et al., 2001] we calculated the boundary layer breakaway time associated with subducting lithosphere according to

\[
t_s = \frac{1}{5.38 \kappa_m} D_b^2 .
\]

(4.7)

Finally, we defined \( u_c \) as

\[
u_c = \frac{(R_m - R_c)}{t_s} .
\]

(4.8)

This is of the form \( u_c \sim Ra^{2\beta} \). More precisely:

\[
u_c = \frac{a_1 \kappa}{2 (R_m - R_c)} \left( \frac{Ra}{Ra_{crit}} \right)^{2\beta} .
\]

(4.9)

Here \( a_1 \) is a scaling parameter. It takes the value of 5.38 for the case where \( \beta = 1/3 \) [Schubert et al., 1979, 1980]. For simplicity, we assumed this holds for all tested \( \beta \) values.

To achieve consistent velocities at present day mantle temperatures, we accounted for the effect of \( \beta \) on \( u_c \). We used the present day values for convective vigor \( (Ra_{now}) \) and velocity \( (u_{now}) \) along with the \( \beta = 1/3 \) scaling used as a reference. This gave us

\[
u_{now} = a_1 \frac{\kappa}{2 (R_m - R_c)} \left( \frac{Ra_{now}}{Ra_{crit}} \right)^{\frac{2\beta}{3}} .
\]

(4.10)

\[
u_{now} = a_2 \frac{\kappa}{2 (R_m - R_c)} \left( \frac{Ra_{now}}{Ra_{crit}} \right)^{2\beta} .
\]

(4.11)

\[
a_2 = a_1 Ra_{crit}^{\frac{2\beta - \frac{2}{3}}{2\beta}} Ra_{now}^{\frac{-2\beta}{2\beta}} .
\]

(4.12)

where we calculated the value of \( a_2 \) for each \( \beta \).
Deep Water Cycle Component

The deep water cycling component tracked the flow of water between the surface and interior reservoirs. We used the model of Sandu et al. [2011]. Water leaves the mantle as an incompatible element via batch melting. We assumed that melting at mid-ocean ridges dominated water loss from the mantle. Subducting slabs deliver water back into the mantle. In the following we detail how we tracked these flows.

Melting  We tracked mantle melting by defining a geotherm, a solidus and a liquidus. We assumed the geotherm was in conductive equilibrium in the lithosphere and followed the adiabat below this. We Based on these two assumptions, We calculated the geotherm as

$$T(z)|_{z \leq D_b} = T_s + \frac{q_m}{k} z. \quad (4.13)$$

$$T(z)|_{z > D_b} = T_p + \frac{q \alpha T_m}{C_p} z. \quad (4.14)$$

In equations 4.13 and 4.14, the potential temperature ($T_p$) is mantle temperature minus the adiabatic component and is calculated at the base of the lithosphere. We would like to emphasize that a depth dependent thermal profile in the interior mantle is used to calculate melt volumes only; in our models, the geotherm does not contribute to the convective dynamics (i.e., adiabatic heating does not drive convection).

The solidus defines the temperature vs. depth profile below which all mantle material will remain in its solid phase. If the mantle is warmer than the solidus, it will start to melt. Increasing mantle temperature further increases the volume of melt produced. If temperature increases enough enough, the entire parcel of mantle melts. This temperature defines the liquidus. We used two second-order polynomial curves to define the solidus and liquidus [Hirschmann, 2000]. For hydrous melting these functions are

$$T_{sol-hydr} = T_{sol-dry} - \Delta T_{H_2} \quad (4.15)$$
\[ T_{\text{liq-hydr}} = T_{\text{liq-dry}} - \Delta T_{\text{H}_2\text{O}} \]  \hspace{1cm} (4.16)

where \( T_{\text{sol-dry}} \) and \( T_{\text{liq-dry}} \) are the dry solidus and liquidus, respectively. \( T_{\text{sol-hydr}} \) is the hydrated solidus and \( T_{\text{liq-hydr}} \) is the hydrated liquidus. In equations 4.15 and 4.16, the second term represents the temperature shift of each curve caused by hydrous melting. This adjustment temperature scales with water concentration in the melt according to

\[ \Delta T_{\text{H}_2\text{O}} = K X^\gamma_{\text{melt}} \]  \hspace{1cm} (4.17)

where \( K \) and \( \gamma \) are constants, which were calibrated by [Katz et al., 2003]. The parameter \( X_{\text{melt}} \) is the ratio of water in the melt fraction expressed in kg of water per kg of melt. It is calculated as

\[ X_{\text{melt}} = \frac{C_{\text{mo}}}{D_{\text{H}_2\text{O}} + F_{\text{melt}} (1 - D_{\text{H}_2\text{O}})} . \]  \hspace{1cm} (4.18)

In Equation 4.18, \( C_{\text{mo}} \) is the bulk water composition in the solid mantle (expressed as a weight fraction), and \( D_{\text{H}_2\text{O}} \) is the bulk distribution coefficient which takes the value of 0.01 – highlighting it behaves as an incompatible trace element. The term \( F_{\text{melt}} \) is the degree of melting expressed as melt fraction. It is parameterized by a power-law as

\[ F_{\text{melt}} = \frac{T - (T_{\text{sol-dry}} - \Delta T_{\text{H}_2\text{O}} (X_{\text{melt}}))^\beta}{T_{\text{liq-dry}} - T_{\text{sol-dry}}} . \]  \hspace{1cm} (4.19)

This definition of \( F_{\text{melt}} \) is valid from the surface to a depth of 300 km as constrained by observation and melting experiments. Therefore, we prohibited any melt production below this depth.

The melt zone thickness \( D_{\text{melt}} \) is dependent upon the relative positioning of the geotherm and the solidus. The base of the melt zone is the deepest temperature at which these curves intersected. At this depth, a parcel of upwelling mantle starts melting. The top of the melt zone is where geotherm and solidus intersect closer to the surface. At this depth, the parcel of mantle has cooled to the point where melt is no longer produced. The vertical distance between these two depths defines \( D_{\text{melt}} \). We integrated \( F_{\text{melt}} \) and \( C_{\text{mo}} \) over
\(D_{\text{melt}}\) to provide average values for our water budget calculations.

**Degassing** Melting at mid-ocean ridges (MOR) transfers water from the mantle reservoir to the surface reservoir. The degassing rate \(r_{\text{MOR}}\) depended on the volume of mantle moving through the melt zone, the amount of melt produced within the melt zone, and how much of the water within the melt makes it to the surface. We modeled this process as

\[
r_{\text{MOR}} = \rho_m F_{\text{melt}} X_{\text{melt}} D_{\text{melt}} S \chi_d
\]

(4.20)

where \(F_{\text{melt}}\) is the integrated melt fraction in the melt zone and \(\chi_d\) is the degassing efficiency factor. Both \(D_{\text{melt}}\) and \(X_{\text{melt}}\) are calculated as specified above. The areal spreading rate \(S\), which is derived from a boundary layer model [Schubert et al., 2001], is defined as

\[
S = 2L_{\text{ridge}} u_c.
\]

(4.21)

We have assumed symmetrical spreading along a constant ridge length \((L_{\text{ridge}})\) and use the definition of \(u_c\) given in equation (4.8).

**Regassing** Subducting slabs deliver water bound in the serpentinized and thin sedimentary layers back into the mantle [Rüpke et al., 2004]. We assumed that most water held in the sedimentary layer degassed from the slab and found its way back to the surface. Therefore, we only accounted for water delivered by the serpentinized layer. Water was delivered back into the mantle at a rate of

\[
r_{\text{SUB}} = f_h \rho D_{\text{hydr}} S \chi_r,
\]

(4.22)

where \(f_h\), \(D_{\text{hydr}}\), and \(\chi_r\) are the mass fraction of water in the serpentinized layer, the thickness of the serpentinized layer and the regassing efficiency factor, respectively. The hydrous phase of serpentinite decomposes at a temperature around 700 °C [Ulmer and Trommsdorff, 1995]. We calculated the depth of this isotherm in the subducting slab...
(\(D_{\text{hydr}}\)). We assumed the maximum value \(D_{\text{hydr}}\) could take was 20 km. This is a rough approximation of the depth to which fractures may penetrate and deliver water into the lithosphere during slab bending at convergent margins.

We calculated the flow rate of mantle water \(r_{\text{Mwa}}\) according to

\[
r_{\text{Mwa}} = r_{\text{SUB}} - r_{\text{MOR}}.
\]

(4.23)

Positive values of \(r_{\text{Mwa}}\) indicate a net influx of water into the mantle.

**Temperature- and Water-Dependent Mantle Viscosity**

The temperature dependence of mantle viscosity is defined as:

\[
\eta = \eta_0 \exp \left( \frac{A}{RT_m} \right)
\]

(4.24)

where \(\eta_0\), \(A\), \(R\) are a reference viscosity, activation energy for dislocation creep [Weertman and Weertman, 1975] and the universal gas constant, respectively. The amount of water in the mantle also played a role in determining mantle viscosity. Experiments have shown that mantle viscosity and mantle water volumes are related by a power-law [Carter and Ave’lallémant, 1970, Chopra and Paterson, 1984, Mackwell et al., 1985, Karato and Wu, 1993]. The power law was further refined to include dependence on water fugacity in olivine [Hirth and Kohlstedt, 1996, Mei and Kohlstedt, 2000]. Assuming an empirical relation for water fugacity based on mantle water concentrations [Li et al., 2008], we calculate the effective viscosity as

\[
\eta_{\text{eff}} = \frac{\tau}{\dot{\varepsilon}} = \eta_0 A_{\text{cre}}^{-1} \left( \exp \left( c_0 + c_1 \ln C_{\text{OH}} + c_2 \ln^2 C_{\text{OH}} + c_3 \ln^3 C_{\text{OH}} \right) \right)^{-r} \exp \left( \frac{A}{RT} \right)
\]

(4.25)

where \(\tau\) is stress and \(\dot{\varepsilon}\) is strain rate. Li et al. [2008] determined the constants \(c_0\), \(c_1\), \(c_2\) and \(c_3\). The water concentration \((C_{\text{OH}})\) is expressed in \(H/10^6\) Si. The values \(\eta_0\) and \(A_{\text{cre}}\) are a calibration and material constant, respectively. Table 4.1 lists the values of the fixed
parameters we used in our study. Table 4.2 lists the parameter space we tested.

Table 4.1: Deep Water Cycle Model Parameters

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convective Model</td>
<td>$T_s$</td>
<td>Surface temperature</td>
<td>300</td>
<td>K</td>
</tr>
<tr>
<td></td>
<td>$H(0)$</td>
<td>Initial radiogenic heat</td>
<td>4.51</td>
<td>$J/(m^3 \text{yr})$</td>
</tr>
<tr>
<td></td>
<td>Rm</td>
<td>Mantle radius</td>
<td>6371</td>
<td>km</td>
</tr>
<tr>
<td></td>
<td>Rc</td>
<td>Core radius</td>
<td>3471</td>
<td>km</td>
</tr>
<tr>
<td></td>
<td>$\rho_m$</td>
<td>Mantle density</td>
<td>3000</td>
<td>$kg/m^3$</td>
</tr>
<tr>
<td></td>
<td>$k_m$</td>
<td>Thermal conductivity</td>
<td>4.2</td>
<td>$W/(mK)$</td>
</tr>
<tr>
<td></td>
<td>$c_p$</td>
<td>Specific heat</td>
<td>1400</td>
<td>$J/(kgK)$</td>
</tr>
<tr>
<td></td>
<td>$\alpha$</td>
<td>Thermal expansivity</td>
<td>3.00 $\times$ 10^{-5}</td>
<td>K^{-1}</td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>Convective exponent</td>
<td>0.33</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$\lambda$</td>
<td>Decay constant</td>
<td>3.4 $\times$ 10^{-10}</td>
<td>yr^{-1}</td>
</tr>
<tr>
<td></td>
<td>$Ra_{cr}$</td>
<td>Critical Rayleigh number</td>
<td>1100</td>
<td>-</td>
</tr>
<tr>
<td>Water Cycling</td>
<td>$\eta_0$</td>
<td>Viscosity constant</td>
<td>1.7 $\times$ 10^{17}</td>
<td>Pa·s</td>
</tr>
<tr>
<td></td>
<td>$A_{cre}$</td>
<td>Material constant</td>
<td>90</td>
<td>$MPa^{-r/s}$</td>
</tr>
<tr>
<td></td>
<td>r</td>
<td>Fugacity exponent</td>
<td>1.2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$Q_a$</td>
<td>Creep activation energy</td>
<td>4.8 $\times$ 10^{5}</td>
<td>J/mol</td>
</tr>
<tr>
<td></td>
<td>$\chi_d$</td>
<td>Degassing efficiency factor</td>
<td>0.03</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$\chi_r$</td>
<td>Regassing efficiency factor</td>
<td>0.015</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>OM</td>
<td>Mass of 1 Earth ocean</td>
<td>1.39 $\times$ 10^{21}</td>
<td>kg</td>
</tr>
<tr>
<td></td>
<td>OM(0)</td>
<td>Ocean masses initially in mantle</td>
<td>2</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4.2: Tested Parameter Space

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>Convective exponent</td>
<td>0.2-0.33</td>
</tr>
<tr>
<td>$H_i$</td>
<td>Initial radiogenic heat</td>
<td>4.51, 3.157</td>
</tr>
<tr>
<td>$T_{m_i}$</td>
<td>Initial mantle temperature</td>
<td>3300, 2300, 1300</td>
</tr>
<tr>
<td>$\chi_d$</td>
<td>Degassing efficiency factor</td>
<td>0.002, 0.02, 0.04, 0.4</td>
</tr>
<tr>
<td>$\chi_r$</td>
<td>Regassing efficiency factor</td>
<td>0.001, 0.003, 0.01, 0.1</td>
</tr>
<tr>
<td>OM$_i$</td>
<td>Ocean masses initially in mantle</td>
<td>0.01, 0.25, 0.5, 0.75, 1, 2, 4, 6</td>
</tr>
</tbody>
</table>


4.3.2 Analyzing Structural Stability

A model is structurally stable if its outputs do not qualitatively change in the presence of low amplitude unmodeled effects [Guckenheimer and Holmes, 1983, George and Oxley, 1985]. Applying structurally unstable models to account for observational data sets is problematic as the robustness of conclusions will be compromised. For this reason, the first step we took in testing our model was to determine whether it was structurally stable. We assessed structural stability using a “perturbed physics” approach [Astrom and Murray, 2008]. Our approach followed the specifics detailed in Seales et al. [2019]. In short, we randomly perturbed the model over time. Perturbations were randomly drawn from a normally distributed set that had a fixed mean and variance. We repeated this process 100 times to form an ensemble of perturbed paths. Each perturbed path started with an identical initial condition and had the same parameter values. However, as the perturbations were random, the perturbed paths differed from each other. To determine whether the model was structurally stable, we compared the mean of the ensemble to the unperturbed path. If the two are within some tolerance, the model is structurally stable.

For a structurally stable model, we can relate the ensemble spread to the model’s structural uncertainty [Strong and Oakley, 2014, Wieder et al., 2015]. The structural uncertainty of the model is the time dependent form of the ensemble. It provides a probability distribution of how far the perturbed paths stray from the mean (in effect, it is a model confidence interval that accounts for structural uncertainty). For the purposes of our study, we define the structural uncertainty bounds as the two-sigma window from the full ensemble about the mean path (further details can be found in Seales et al. [2019] and Seales and Lenardic [2020]).

4.3.3 Observational Data Constraints

Successful models are defined as those that can satisfy observational constraints. We use both present day and paleo proxy data to constrain successful model paths. For present day constraints, we use mantle temperature and $Ur$ (Equation 4.2). The present day mantle
temperature falls between 1300 and 1400 °C [Herzberg et al., 2010]. Jaupart et al. [2015] estimated the present day $U_r$ is between 0.2 and 0.5. Accounting for the thermal effect of continents allows for an upward $U_r$ correction of 0.2 [Lenardic et al., 2011, Grigné and Labrosse, 2001].

Figure 4.1 shows paleo temperature proxy data constraints. The upper and lower bounds of Ganne and Feng [2017] encompass the data sets of Condie et al. [2016] and Herzberg et al. [2010]. Ganne and Feng [2017] suggested that their maximum and minimum bounds may represent the temperature of plumes and ambient mantle, respectively. This is not consistent with Condie et al. [2016] Herzberg et al. [2010], who both considered their data to represent ambient mantle (which would correlate to $T_p$ for thermal history models). We will do the same herein. Under this view the combined data spread, from different groups, represents observational uncertainty in paleo temperature constraints.

The full range of observational uncertainty allows multiple models to be viable [Seales and Lenardic, 2020]. As noted in the introduction, our aim is to test the idea that the Earth experienced a multi-stage cooling. As such, we will follow the conceptual interpretation of the data offered by Condie et al. [2016], who argued that the data was indicative of a change in cooling slope at 2.5-2.0 Ga. Our specific hypothesis for this change in slope is a change from net mantle dehydration to rehydration. For this reason our successful models will not only need to match thermal data, within uncertainty, but will also need to allow for a change in net hydration between 3.0 to 1.75 Ga and an associated change in cooling slope (or potentially a change from heating to cooling) within that time window.

### 4.3.4 First Order Model Feedbacks

Before moving to model results, it is worth conceptually overviewing critical model feedbacks. In our models, we assumed that mantle viscosity was the dominant resistor to plate motions. As shown in Equation 4.25, both mantle temperature ($T$) and mantle water concentration ($\chi_m$) influence mantle viscosity ($\eta$). Figure 4.3 shows a feedback loop diagram for coupled hydration and thermal feedbacks [Crowley et al., 2011]. The left
hand side of Figure 4.3 shows the thermal feedback structure. If mantle temperature were to increase, mantle viscosity would decrease. This would increase plate velocities. Faster velocities increases heat flow and cool the mantle. An increase in mantle temperature results in mantle cooling, a negative feedback. Therefore, the thermal part of our model wants to buffer itself from changes as has been know for some time [Tozer, 1972].

Figure 4.3: Simplified feedback loops associated with the thermal (negative feedback) and deep water cycle (positive feedback) modules of the coupled model.

The right hand side of Figure 4.3 shows the water cycling feedback loop. Mantle viscosity effects both the thermal and the water cycling loop which leads to a coupling between the two. If there is a net flow of water out of the mantle, \( \chi_m \) decreases. Removing water from the mantle increases mantle viscosity. This causes plates to move more slowly, decreasing mantle heat flow. This results in a hotter mantle, which has two effects on water transport. First, a hotter mantle is associated with a thinner lithosphere. Thinning the lithosphere also thins the hydrated layer held within it. A thinner hydrated layer can deliver less water back into the mantle. A second effect of a hotter mantle is that it decreases the solidus, which generates more melt. Both decreased return of water to
the mantle and increased melting by a depressed solidus cause a net decrease in mantle water concentration. The water cycle feedback, then, introduces the potential of a positive feedback. If water is lost, feedbacks can lead to a tendency to lose more. This positive feedback can dominate the overall system feedback if it is not offset or balanced by the thermal feedback. Which type of behavior will prevail depends on the strength balance between the feedbacks which can change over time. For this reason we will need to be able to quantify the strength of different feedbacks over model evolution times.

4.3.5 Assessing Feedback Strengths

Testing the hypothesis that a change in the deep water cycle can account for a multi-stage thermal history requires that the strength of feedbacks over model evolution time can be quantified. We will do so using the method developed by Crowley et al. [2011]. The method determines which feedback, thermal or water cycling (Figure 4.2), dominated mantle viscosity at a particular time. The method defines the change of mantle viscosity ($\dot{\eta}$) as a function of these feedback as

$$\dot{\eta} = \frac{\partial \eta}{\partial T} \frac{\partial T}{\partial t} + \frac{\partial \eta}{\partial \chi_m} \frac{\partial \chi_m}{\partial t} = \eta_T \ddot{T} + \eta_\chi \ddot{\chi}_m. \quad (4.26)$$

The first term on the right hand side ($\eta_T \ddot{T}$) represents the thermal feedback and the second term ($\eta_\chi \ddot{\chi}_m$) the water cycling feedback. In our forward model, we solve for $\ddot{T}$ from equation 4.1 and $\ddot{\chi}_m$ using equation 4.23. We can calculate $\eta_T$ and $\eta_\chi$ by taking partial derivatives of equation 4.25, which are

$$\eta_T = -\frac{A \eta}{R T_m^2}. \quad (4.27)$$

$$\eta_\chi = -\frac{r \eta}{\chi_m} (c_1 + 2 * c_2 \ln(C_{OH} + 3 * c_3 \ln^2 C_{OH})). \quad (4.28)$$

Crowley et al. [2011] defined a nondimensional ratio of these two values

$$S_{WT} = \left| \frac{\eta_\chi \ddot{\chi}_m}{\eta_T \ddot{T}} \right|. \quad (4.29)$$
When $S_{WT} > 1$ the water cycling feedback starts to dominate. For lower values, the thermal feedback becomes progressively more dominant in the overall feedback structure. Using the outputs of our model, we determined time intervals with different feedback structures. Doing so allowed us to more fully quantify our principal hypothesis.

4.4 Results

In this section we assess model structural stability, isolate model paths that satisfy geologic proxy data (from more than $10^5$ model paths with variable initial conditions, model inputs, and $\beta$ values, see Table 4.2), and analyze the degree to which coupling deep water cycling to the Earth’s thermal evolution can account for a multi-stage cooling history.

4.4.1 Structural Stability and Structural Uncertainty

Figure 4.4(a) shows the structural uncertainty of a coupled model compared to that of a thermal history model with no water cycling [Seales et al., 2019]. Each model is for $\beta = \frac{1}{3}$. Coupling the deep water cycle to the thermal history model caused a four fold increase in the structural uncertainty. The uncertainty maxed out at a two-sigma value of $\sim 100 \degree C$ at present day.

Figure 4.4(b) plots a structural uncertainty window for a model path (shaded gray). The window is determined from 100 different perturbed cases of the coupled model used for Figure 4.4(b). This provides, in effect, a structural confidence interval for a model path. The solid line shows the ensemble mean from all the perturbed cases. The ensemble mean matched the evolution path of the unperturbed model which indicated that the model maintains structural stability. We also show one perturbed path from the ensemble (dashed line). It retained the same first order trend of the mean path. The structural uncertainty for present day temperature was very near the uncertainty in present day data constraints. For paleo temperatures, structural uncertainty was less than the uncertainty associated with paleo proxy data constraints.

In Figure 4.4(c) and 4d we show the ensemble window for both mantle and surface
Figure 4.4: Structural uncertainty analysis. a) Comparison of simple and coupled thermal history model structural uncertainties. b) Sample thermal history ensemble and one ensemble path. c) Mantle water volume ensemble with structural uncertainty. d) Surface water volume ensemble with structural uncertainty.
water volumes, respectively. Each water volume ensemble is for the thermal path shown in Figure 4.4(c). The structural stability of the model lead to the mean of ensemble paths tracking the unperturbed model path. For water volumes, individual perturbed paths more closely tracked means values than was the case for mantle temperature. The structural uncertainty for water volumes maxed at \( \sim 11\% \).

What is critical, in terms of model application, is that our uncertainty analysis shows that the structural uncertainty of our model is comparable to uncertainty in present day data and is smaller than the uncertainty of proxy paleo data. If this was not the case, and the structural uncertainty of a model was considerably greater than that of observational constraints, then the ability of constraints to knock out model (i.e., rule out hypotheses) would be weakened. In effect, models with high levels of structural uncertainty relative to data become harder to rule out using the data itself. This is not the case for our models over the \( \beta \) range we will test. For lower \( \beta \) value models, particularly negative \( \beta \) models, this is no longer the case as structural uncertainty can be larger than data uncertainty [Seales et al., 2019]. This is another reason why we focused on higher \( \beta \) models.

The perturbed paths that make up the ensemble provide a proxy means for modeling shorter time-scale fluctuations about a mean thermal history trend [Lenardic et al., 2016]. Our interest herein is on whether distinct cooling stages are possible over geologic time. If so, these stage would show different means. Our structural uncertainty analysis shows that, for a structurally stable model, that will hold in light of low amplitude, unmodelled effects. As such, we will focus the results section on mean trends. In terms of assessing the ability of models to match data constraints, using mean trends will rule out a larger class of models as accounting for structural uncertainty would increase the probability that any particular case could match data. That increase would scale as the structural uncertainty of the model. We will take the more restrictive approach in testing our main hypothesis noting that structural uncertainty allows for an extended range of successful model paths.
4.4.2 Models That Match Observational Data Constraints

Observational constraints were used to eliminate unsuccessful model paths. By model path we mean the thermal path we computed using a unique set of initial condition and parameter values for models with different values of $\beta$. We applied different constraints individually to all model paths. This provided a distribution of model paths that satisfied each constraint (Figure 4.5 and 4.6). Figure 4.5 shows the distribution of model paths that satisfied the present day $T_p$ constraints (Figure 4.5(a)) and present day $Ur$ constraints (Figure 4.5(b)). Figure 4.6 shows the distribution of model paths that satisfied the paleo proxy data constraints. Figure 4.6(a) is the most conservative paleo constraint applied as models only needed to fall within the full range of observational uncertainty (Figure 4.1). The constraint of Figure 4.6(b) provides a less conservative paleo constraint as it also rules out models that do not allow for a changes in mantle hydration from net degassing to regassing between 3.0 and 1.75 Ga.

Each constraint eliminated different paths, which progressively decreased the number of successful paths. Figure 4.7 shows the distribution of model paths that satisfied all
Figure 4.6: Distributions of model paths that matched present day constraints: a) thermal and b) $Ur$

constraints (Figure 4.7). The paths are distributed between a minimum $\beta$ value of 0.25 and a maximum of 0.3, peaking around $\beta = 0.29$. Figures 4.5 and 4.6 hinted at this outcome. In both the paleo and present day proxy data, one constraint favored paths with a higher $\beta$ value whereas the other favored paths with a lower $\beta$ value. This indicated the likelihood of a sweet spot. When we inspected the parameters of the successful paths, we found a commonality between 19 of 21: the value of $\chi_d = 0.04$ and $\chi_r = 0.1$.

Figure 4.8 shows the present day mantle and surface water volumes associated with the thermal paths that matched all thermal constraints. Paths that assumed initial mantle water volumes of 4 and 6 OMs peaked at just below 3 OMs of present day mantle water volumes (Figure 4.8(a)). Paths that started with an initial mantle water volume of 2 OM did not end in a similar range, as there was not enough water available. Present day surface water volumes split into groups based on the amount of assumed initial mantle water (Figure 4.8(b)). This is not a surprise. Model feedbacks can regulate internal mantle water volumes, but the model has no such feedbacks that would regulate surface water volumes. As such, final surface water volumes are more strongly dependent on assumed initial water.
volumes. In our analysis, we did not account for any initial surface water or late stage water addition. If we did, models that assumed an initial condition of 2 OMs water volumes could potentially match present day surface water volume. We also did not account for surface water loss to space over model evolution times. If we did, models that assumed an initial condition of greater than 4 OM volumes could also potentially match present day water volumes. With those two caveats, the assumed 4 OM initial condition does most closely satisfy the present day surface water constraint.

We conclude this section by plotting mantle water volume and mantle temperature over time for all successful models (Figure 4.9). Figure 4.9(a) shows that the mantle water volumes collapsed to a relatively narrow range over the age of the Earth for most models.
Figure 4.8: Present day mantle (a) and surface (b) water volumes, in present day ocean masses, for model paths that matched all constraints. Histogram is colored coded by the amount of water that was assumed as an initial condition.

The exceptions being cases that began with 2 OM of water. This supports the idea that internal mantle feedbacks regulate the amount of water in the mantle. Figure 4.9(b) shows the range of successful thermal paths satisfying all constraints. There are two distinguishable groups: those that start warm and those that start cooler. Both groups lead to multi-stage cooling paths but of different natures. The models that started cooler experienced a mantle heating stage followed by cooling. The models that started warmer experienced a reduced cooling stage between 4.0 to 2.5 Ga followed by accelerated mantle cooling. It is interesting to note that the warm start trend is in line with the conceptual interpretation of the paleo temperature data offered by Condie et al. [2016] while the cooler start is consistent with the interpretation of Herzberg et al. [2010].

4.4.3 Feedback Analysis

In this section we provide further support that a multi-stage thermal history can result from changes in the deep water cycle. For clarity, we focus on the path highlighted in
Figure 4.9: Mantle water volume (a) and thermal (b) paths that matched all constraints. In (b) we show paleo thermal history estimates that suggest a multi-phase thermal history along with successful thermal paths for qualitative comparison with the output of our analysis.

Figure 4.10 shows the relative and absolute influences of the deep water cycle and thermal effects on mantle viscosity. The dark red line shows how the deep water cycle changed mantle viscosity with time ($\eta_X \dot{X}_m$), and the light red line shows how thermal effects changed mantle viscosity with time ($\eta_T \dot{T}$). The black line ($S_{WT}$) is the ratio of $\eta_X \dot{X}_m$ to $\eta_T \dot{T}$. The first 0.8 billion years of the thermal history was characterized by $S_{WT} < 1$.

Figure 4.9(b).

Over half of that stage $S_{WT} < 0.1$. This indicates that thermal effects dominated the overall system feedback. The second stage of the thermal history began at roughly 4 Ga when $S_{WT}$ grew larger than one. This persisted until $\sim 2.8$ Ga. During this stage, $S_{WT}$ peaked at 3.6 Ga, indicating a fundamental change to the overall model feedback structure. This coincided with a minimum in ($\eta_T \dot{T}$). The third stage, with $S_{WT}$ again dropping below unity, lasted from 2.8 Ga to present. Early in this phase, at $\sim 2.7$ Ga, the deep water cycle experienced a dramatic shift: the net flow of water began to enter rather than exit the mantle. The change to a rehydrating mantle means the water cycle tended to lower mantle viscosity.
which, in turn, tended to enhance mantle cooling. The timing of the shift in water cycling is expressed in Figure 4.10 by \( \eta X \dot{\chi}_m \) dropping below zero. During this last stage, \( S_{WT} \) approached one as both \( \eta X \dot{\chi}_m \) and \( \eta T \dot{T} \) grew quickly but in opposing directions. This ended abruptly when the maximum dehydrated layer thickness was reached at 1.6 Ga, and \( S_{WT} \) slowly drifted towards lower values.

![Figure 4.10: Quantitative analysis of feedbacks. When \( S_{WT} > 1 \), water cycling influence mantle viscosity (\( \eta \dot{\chi}_m \)) more than thermal effects (\( \eta T \dot{T} \)).](image)

The feedback analysis of Figure 4.10, and associated model trends of Figure 4.9, provide insights into the competing factors that allow for the three thermal history stages discussed.

During the first stage, the strong negative feedback associated with the thermal loop caused rapid adjustment from an initially hot or cool start (Figure 4.9(b). This worked to bring heat flow and radiogenic heat production toward a balance. In this sense, the model
behavior was not drastically different than classic thermal history models with $\beta$ values near 0.3 [Tozer, 1972, Schubert et al., 1980, Spohn and Schubert, 1982, Jackson and Pollack, 1984]. This is consistent with $S_{WT}$ being below 0.1 over much of this stage (Figure 4.10).

As heat flow and radiogenic heat production approached a balance, water cycling effects on mantle viscosity could start to compete with the negative thermal feedback. This marked the start of the second thermal history stage. The mantle was degassing during this stage (Figure 4.9(b). This worked to stiffen the mantle. This, on its own, tended to drive less efficient mantle convection and associated heating. However, radiogenics in the mantle were decaying and producing less heat which allowed for a balance between hydration and thermal feedbacks. As a result, a stage of relatively mild mantle cooling could be maintained. As the mantle continued to dehydrate, the solidus shifted to warmer temperatures. This reduced the melt fraction and slowed the rate at which water was lost from the mantle. This limited the peak of the water cycling effect during the second stage. As a result, the flat line temperature trend could not be maintained and the mantle started to progressively cool. The geotherm moved towards the solidus, further reducing melt volumes. Cooling also thickened the lithosphere and the thickness of the hydrated layer within it. This started to deliver more water back into the mantle. This further damped the rate at which water was lost. With this process set in motion, the second stage gave way to the final thermal history stage.

During the third stage, water cycling transitioned to net rehydration of the mantle. This caused a change from water cycling tending to stiffen the mantle to water cycling tending to weaken the mantle (Figure 4.10). Thermal effects on viscosity were stronger than water effects but the two were not strongly out of balance (Figure 4.10). As such, both contributed to the overall trend in the final stage, particularly over the first $\sim$1 billion years of the third stage. As the mantle cooled, the geotherm shifted further towards the solidus and the melt zone began to shrink. This reduced the melt fraction, which allowed less water to leave the mantle. The lithosphere continued to thicken as did the hydrated layer embedded within it. This increased the rate at which water was delivered into the
mantle. The combined effects of a shrinking melt zone and increased delivery of water into the mantle lead to a mild rise in the water cycling effect over the start of the third stage (Figure 4.10). That rise ended once the hydrated lithosphere layer reached its maximum allowable thickness. If this limit did not exist, then $S_{WT}$ could have continued its mild rise but it would have remained below unity as the rise in thermal effects exceeded it. With the limit in place, net rehydration of the mantle, and associated decline of surface water, slowed. None the less, both thermal and water cycling effects continued to influence thermal history with the thermal effect being stronger by a factor of $\sim2$ at a model time representing present day (Figure 4.10).

### 4.5 Discussion

A three stage thermal history, as analyzed in the previous section, is consistent with geologic proxy data and can match present day constraints. Figure 4.11 shows the uncertainties of our successful model paths projected onto the paleo temperature proxy data. We calculated the mean and two-sigma uncertainty bounds for the 21 thermal paths that satisfied all constraints. The time domain spans from the present day to 4 Ga, as this is the extent of the proxy data. However, in all of our models, the second thermal history stage began prior to 4 Ga. Within model uncertainties, the second stage is predicted to extended to 3.5-2 Ga. The mean of the uncertainty window predicts that a transition to stage three occurs at $\sim2.7$ Ga. That transition is associated with a change from net mantle degassing to net regassing. Collectively, our results argue that paleo temperature data is not necessarily indicative of a change in tectonic regime [Condie et al., 2016]. Changes in the deep water cycling can account for the data and for present day data constraints. This provides a viable alternative hypothesis.

Using data constraints to eliminate potential model paths also provided us with a $\beta$ value range that allowed for successful model paths. Figure 4.7 indicates that the range is $0.25 \leq \beta \leq 0.3$. We can compare that with results from thermal history models that do not include deep water cycling. Seales and Lenardic [2020] systematically explored a
Figure 4.11: Summary of all results and uncertainties. We show the mean and two-sigma uncertainty of model paths that matched all constraints. The red dots indicate the mean and uncertainty of the change from water cycling (phase two) to thermal dominance (phase three) of the thermal history. This change in phase coincided with the change in slope of both paleo mantle temperature estimates, suggesting our hypothesis is viable and that there is no need to invoke a change in convective regime.

range of such models and subjected them to the same thermal constraints used herein. The probability density function for successful models spanned a wider range than that of Figure 4.7. It was also double peaked with a mild peak at $\beta = -0.10$ and a stronger peak at $\beta = 0.15$. Both those peaks are associated with models that assume that plate and/or plate margin strength provides the primarily Resistance to plate motions [Conrad and Hager, 1999a, Korenaga, 2003]. Our results, on the other hand, suggest that mantle viscosity primarily resists plate motions.

The principal reason for the different modeling conclusions noted above is that the study of this paper coupled deep water cycling to thermal history. This allowed us to add the constraint that deep water cycling transitions from net mantle dehydration to net rehydration over geologic time. The time window of the transition, within model uncertainty, that we predict is in line with observational constraints [Parai and Mukhopad-
Not surprisingly, adding a new constraint lead to a lower overall number of successful model paths than was determined by the study of Seales and Lenardic [2020]. There is another difference that should be noted from our approach and an alternative modeling methodology. Some thermal history modeling approaches build present day thermal constraints directly into the models and calculate model paths in "reverse time" starting from the present and extending to the past [Christensen, 1985, Korenaga, 2003]. We did not follow this approach. This allowed us to use present day constraints to eliminate model paths rather than build the constraints into the model. This had a critical effect on our viable $\beta$ range. Models with lower values of $\beta$ struggled to match present day temperatures (Figure 4.5(a)). This is consistent with McNamara and Van Keken [2000] who found that lower $\beta$ value models tended to run too hot.

It is worth specifically comparing our model to another thermal history model that also invoked hydration effects but in a very different way. Korenaga [2003] took the view that plates rather than mantle viscosity primarily resist plate motions. The associated model they explored assumed that an elevated mantle temperature produces larger melt volumes which, in turn, creates a thicker dehydrated lithosphere in the Earth’s past. This leads to thicker and stronger plates further back in geologic time. The thicker and stronger plates impeded plate motions and associated mantle cooling. Korenaga [2003] determined that an effective $\beta$ value of -0.15 could capture these effects within a thermal history model. This is different from the models herein which assumed that mantle viscosity provides the principal resistance to plate motion. As noted in the introduction, this was motivated, in part, by studies that have argued against the physical viability of negative $\beta$ models [Gerardi et al., 2019].

Future extension of our models could include effects we did not directly model. Chotalia et al. [2020] showed that including a finite delay in the mixing time for water in the mantle could affect hydration feedbacks on the solidus at the mid-ocean ridge. They found that when they included this delay, it restricted the period of mantle regassing. It also opened up the possibility for shorter timescale oscillations between mantle regassing and
degassing states. Karlsen et al. [2019] also found that a perturbed water cycle, this time by supercontinental breakup, could perturb sea-level by more than 100 m. In both cases, the added model complexities lead to fluctuations about mean trends. Thermal history models traditionally have tracked mean values over time and it has been noted that fluctuations about the mean are to be expected and that should be considered when evaluating thermal history model results [Lenardic et al., 2016, Silver and Behn, 2008]. Although we did not consider the specific effects noted above in our analysis, we did consider the role of unmodeled effects, and associated fluctuations about mean trends, when we evaluated structural uncertainty [Seales et al., 2019]. Models, by definition of the word model, will exclude some physical factors and the value of a structural uncertainty analysis is that it can test the robustness of first-order model trends in the face of this. The fact that the models presented herein are structurally stable goes hand in hand with the robustness of mean trends to potential fluctuations (Figure 4.4(b)).

Our results, as related to changes in deep water cycling, have implications extending beyond mantle cooling. The change from stage two to stage three, for our successful thermal history models, is associated with a switch to net regassing of the mantle. The sequence of events that led to this involved the solidus shifting to warmer temperatures below mid-ocean ridges as melting dehydrated the mantle. Simultaneously, heat flow from the mantle decreased, which thickened lithospheric plates. This thickened the hydrated layer that carried water back into the mantle. This implies a switch from relatively dry to wet subduction. Increased water volume delivered into the mantle by subducting slabs could preferentially produce felsic rather than mafic crust. This would increase the area of Earth’s surface covered by felsic crust. As a result, its oxidative efficiency would decrease, leading to a rise in atmospheric $O_2$ [Lee et al., 2016]. The exposure of larger felsic areas would allow for higher weathering rates, which could enhance an influx of carbonates that would further bolster this rise in $O_2$ [Eguchi et al., 2020]. Our hypothesis predicts, then, that a rise in atmospheric $O_2$ should coincide with the timing of the deep water cycle changing convective efficiency and the associated change in the cooling rate of the mantle.
Our models predicts that this change occurs at a mean model uncertainty time of $\sim 2.7$ Ga. This coincides with the inferred timing of the Great Oxidation Event [e.g., Lyons et al., 2014].

4.6 Conclusions

We have used coupled models of deep water cycling and thermal history to explore the hypothesis that changes in water cycling, over geologic time, could lead to multi-stage mantle cooling. We tested the viability of this hypothesis by applying observational constraints on a wide range of calculated model paths. The hypothesis was shown to be compatible with data constraints and it did not require changes in the tectonic style of the Earth over geologic time. It also implied that mantle viscosity provides the dominant resistance to plate motions with plate and plate margin strength playing a lesser role.
Chapter 5

Different is More: The Value of Finding an Inhabited Planet that is Far From Earth2.0

5.1 Abstract

The search for an inhabited planet, other than our own, is a driver of planetary exploration in our solar system and beyond. Using information from our own planet to inform search strategies allows for a targeted search. It is, however, worth considering some span in the strategy and in a priori expectation. An inhabited, Earth-like planet is one that would be similar to Earth in ways that extend beyond having biota. To facilitate analysis, we employ a metric akin to the Earth-similarity index of Schulze-Makuch et al. [2011]. The metric extends from zero, for an inhabited planet that is like Earth in all other regards (i.e., zero differences), toward end-member values for planets that differ from Earth but maintain life potential. The analysis shows how finding inhabited planets that do not share all other Earth characteristics could improve our ability to assess galactic life potential without a large increase in time-commitment costs. Search strategies that acknowledge the possibility of such planets can also minimize the potential of exploration losses (e.g., searching for long durations to reach conclusions that are search strategy biased). Discovering such planets could additionally provide a test of the Gaia hypothesis - a test that has proved difficult using only the Earth as a laboratory. Lastly, we discuss how an Earth2.0 narrative that has been presented to the public as a search strategy comes with nostalgia-laden philosophical baggage that does not best serve exploration.
5.2 Introduction

The idea of a second Earth has a long history [Couprie, 2011]. Recently, NASA has entered into second Earth thinking: 1) “This discovery gives us a hint that finding a Second Earth is not a matter of if but when.” - Thomas Zurbuchen, Assoc. Admin., Science Mission Directorate at Nasa, 2017; 2) “This exciting result brings us one step closer to finding an Earth 2.0.” - John Grunsfeld, Assoc. Admin., Science Mission Directorate at Nasa, 2015. The quotes come from press conferences that announced discoveries regarding planets orbiting stars other than our own (exoplanets). A motivator behind these statements is the search for life beyond Earth and placing observational constraints on the probability of life in our galaxy.

Debates about life beyond Earth have a long history. An even-handed referencing would take pages. Interested readers can easily track down a multitude of books, articles, blogs, and the like. As a starting point, two books that encapsulate competing views are Ward and Ward and Brownlee [2000] and Kasting [2012]. End-member views regarding galactic life potential are “rare-Earth” and “plenitude” (i.e. life requires specific environmental conditions, of the kind that exist on Earth, versus the idea that life can thrive in a range of conditions that might exist across planets and/or moons within our galaxy). To be clear (and avoid strawmen): 1) Rare does not mean singular (a rare-Earth view does not exclude life beyond Earth - it posits that in order for a planet to have life, particularly complex life, it must share certain essential features with the Earth, beyond having life, that make the Earth the planet it is and, it argues, that the combination of such features is rare for planets in our galaxy); 2) Plentitude does not mean all things are possible (a plentitude view acknowledges that there will be galactic bodies with conditions that do not allow for life); 3) Nature need not care about end-members (as an example: conditions that make the Earth the Earth, beyond having life, may not be rare in the galaxy such that there are many planets that are essentially analogous to Earth; life may only exist on such planets which would not be in line with a plentitude end-member view and, although it would not support the idea that Earth conditions are rare, it would confirm one aspect of the
rare-Earth argument, i.e., Earth conditions are needed for life).

We are now at the stage of exoplanet research where missions are being designed with the goal of detecting remote signatures of planetary life [e.g., Schwieterman et al., 2018]. This is a new phase of exploration into galactic life potential. It generates a level of excitement that motivated the press release statements quoted in the first paragraph. It brings with it the potential that observational data can be brought to bear on long-standing debates about galactic life. It also brings with it an equally old question related to any exploration: how will we choose to look? This is a decision problem. It is a decision problem under uncertainty. Any exploration comes with risk. Uncertainty adds to that risk. Acknowledging uncertainty and risk does not damp excitement. However, the fact that a particular search strategy is gaining traction, one that is centered around Earth analogs, without being weighed against alternates in light of uncertainty and risk, does motivate the step back and rethink that is the core of this paper (note added in revision: over the time that this paper was in review we became aware that we are not the first to suggest a rethink regarding search strategy [Bean et al., 2017, Kite et al., 2018, Lingam and Loeb, 2018].

A search strategy can be centered on planets that share a high number of Earth attributes, with the thought that this maximizes our chances of finding signs of life. Our central thesis is that a search strategy that allows for deviations from Earth analogs can bring more information and less risk at moderately higher cost. Fleshing this statement out forms the bulk of this paper.

There is an added reason why we argue for a break from an Earth centered view. Wanting to know if life exists beyond Earth, and what that implies for life potential across our galaxy, is not driven by scientific motivations alone. The humanistic implications of the search, and the humanistic/cultural factors that feed into it, should be acknowledged. More specifically, the desire to find a planet like our own is driven by factors that extend well beyond pure scientific curiosity [Messeri]. Our intent is not to argue that humanistic/cultural aspects be removed from the discussion. Rather, we want to layout how finding
inhabited planets that do not look like home to us could carry broader scientific and hu-
manistic implications than finding an Earth2.0. The humanistic/philosophical implications
are discussed in the final discussion section.

5.3 Non-Earthness, Planetary Life Potential, and Search Strategies

Exploration goal(s) determine search strategies. The goal(s) of exoplanet exploration, as
related to galactic life, can be expressed as a single question or intertwined questions.
There is an existence question: ‘does life exist beyond Earth?’ If this is the only motivating
question, then the search endeavor could be of the ‘find one and done’ type. We suspect
that most people who have dedicated time and thought to exoplanet explorations do not
subscribe to this. Thus, there must be a broader goal. We would argue that addressing the
issue of galactic life potential (rare versus plentiful) is that broader goal. That introduces
the intertwined questions of ‘what conditions allow for and/or maximize life potential (i.e.,
the probability that life exists on a planet)?’ and ‘what percentage of planets in our galaxy
are inhabited (i.e., what percentage have a biosphere)?’ For what follows we will, unless
otherwise noted, assume all of these questions are motivators in the collective search
for life beyond Earth and, as such, all should be considered when determining search
strategies.

We now pose an added question: ‘would finding inhabited planets that share many
other Earth attributes or finding inhabited planets that differ significantly from Earth better
help us to address the question of galactic life potential and should this consideration feed
into search strategies?’ We will argue that it should be considered. This leads to a practical
issue that needs to be addressed from the start: an overly broad search is not practical.
Given that, we would like some constraints on how different we think a planet could be
from Earth and still maintain life potential. This motivates the first sub-section below.
5.3.1 How Different can Different Be and Maintain Life Potential

The term “Earth-like” is often used in a qualitative way. A quantitative metric can be defined as per the Earth-similarity index of Schulze-Makuch et al. [2011]. Our interest is more specific than asking how close a planet is to Earth in terms of the full range of variables that come into play. We are interested in the question of how different a planet can be from Earth and still allow for life. We can, for example, use the Earth-similarity index and define a sub-range over which a planet can maintain life potential (this can also be applied to moons). Since our interest is on how different a galactic body can be from the Earth and allow for life, it will be useful to center our metric on zero (i.e., if a planet with life potential is like Earth in all other ways then the value of the metric is zero). However one defines such a metric a useful next step will be to ask what are the extreme values it can take while maintaining a non-negligible probability for life.

We start with an agreed upon criteria: life requires energy. The energy sources for a biosphere are energy from the star a planet or moon orbits and internal energy from the decay of radioactive isotopes within its interior, heat retained from formation, and/or tidal heating.

Figure 5.1a is a schematic of how energy sources affect life on Earth. Solar and internal energy can be used as direct energy sources to power photosynthesis or chemosynthesis. The energy sources also drive cycles that influence environmental conditions. If there is a limited range of environmental conditions under which a biosphere can exist, this implies that energy sources can affect planetary life potential by maintaining livable conditions. For life as we know it, the existence of liquid water is crucial. The classic idea of a "Habitable Zone" ties into delineating conditions required for a planet to maintain liquid water over time scales that allow for life development and evolution [Kasting et al., 1993].

For the Earth, the buffering of environmental conditions is generally considered to rely on hydrological and geophysical/geochemical cycles. Internal energy drives volcanic and tectonic activity that transfers volatiles (CO$_2$, H$_2$O) between a planet’s surface envelopes (atmosphere, hydrosphere, biosphere) and its rocky interior (crust, lithosphere, mantle).
Volcanism cycles greenhouse gases into the atmosphere. Tectonics creates weatherable topography and weathering reactions draw greenhouse gases out of the atmosphere. Weathering depends on hydrology. This means that surface and deep planet cycles are linked in so far as discussions of buffering Earth’s climate are concerned [Kump et al., 2000]. Life also links in as it has the ability to effect the cycles that maintain environmental conditions suitable for its own existence [Lovelock and Margulis, 1974]. Over geologic timescales, atmospheric CO$_2$ content and associated greenhouse climate forcing is influenced by the balance between volcanic degassing and weathering [Berner et al., 1983]. Weathering depends on processes governed partly by surface temperature, which allows for the potential that a planet can buffer/stabilize climate and surface conditions in a manner that allows liquid water to exist over extended time frames [Walker et al., 1981].

The silicate-weathering negative feedback, outlined above, is the currently preferred hypothesis for how the Earth’s climate has been regulated so as not to enter a prolonged hard snowball state or a runaway greenhouse state. This has cast a significant influence on ideas about how to search for inhabited planets beyond our solar system [e.g., Kasting,
It is worth stressing that, as formulated for Earth climate stabilization, the feedback relies on both solar and internal planetary energy. Removing one of the two energy sources thus affects direct energy sources for life and also a mechanism for maintaining environmental conditions conducive to life. This stresses how a planet or moon that lacks one of the two energy sources would be distinctly non Earth-like.

The idea that a planetary body can have life even if solar energy is negligible (Figure 5.1b) is driving exploration of icy moons within our own solar system [Schulze-Makuch and Irwin, 2001, Wenz, 2017] and has been suggested for planets that do not orbit stars [Abbot and Switzer, 2011]. For planets that do not orbit stars the difficulty of remote life detection is extreme. For planets/moons that do orbit a star there is the potential that life on such bodies would not interact with an atmosphere so as to create biosignatures that can be observed remotely [Schwieterman et al., 2018]. However, it is not clear that life would not leave detectable signatures that are not connected to atmospheric chemistry [Lingam and Loeb, 2019]. Detectability cannot be ignored but at this stage we are concerned about limits. What is key to our arguments is that an inhabited planet that lacks solar energy affecting life directly or indirectly remains a viable possibility.

An inhabited planet that lacks internal energy sources (Figure 5.1c) would imply that geophysical/geochemical cycles are not required to maintain conditions conducive to life. Habitable conditions can be maintained on waterworlds (planets with water masses 10-1000 times that of Earth) as a result of stochastic variations in formation conditions with no need for geochemical cycling [Kite and Ford, 2018]. The potential that habitable conditions can be maintained without geo-cycling can extend beyond waterworlds if life itself maintains conditions that allow for its continued existence. This is the core of Gaia theory [Lovelock and Margulis, 1974, Lovelock, 1979, 1988, Watson and Lovelock, 1983]. Over the course of debating the theory, different levels of Gaia have been suggested [Kirchner, 1989, 2003]. Soft Gaia considers life on Earth to have influence on the geophysical/geochemical cycles that modulate Earth’s surface environment while the strong form of Gaia considers life to be critical to modulating surface conditions at livable levels [Barlow, 1992, Schneider,
Under a strong Gaia view, life could exist on a planet or moon that has tapped all of its internal energy. Akin to the discussion of the previous paragraph, the key for what follows is that an inhabited planet that lacks internal planetary energy affecting life directly or indirectly remains a viable possibility.

The scenarios of Figures 5.1b,c are energetic extremes. Between them sits the potential of livable planets that differ from the Earth in other ways. Some examples: Planets without oceans could allow for habitable conditions [Abe et al., 2011]; Planets that are ocean worlds could allow for life [Kaltenegger and Sasselov, 2011]; Planets with internal energy principally driven by tidal heating (a minor factor for Earth) could allow for life [Barnes et al., 2009]; Planets without plate tectonics (the geologic mode of Earth) could allow for conditions conducive to life [Lenardic et al., 2016, Foley and Smye, 2018].

Figure 5.2: Future observations, limited by a window of potential habitability ($\theta$) and by the search strategy employed, are grouped ($\theta_i$) relative to Earth ($\theta_{Earth}$). Each group has a different habitability potential and there is a set of conditions that maximize habitability potential ($\mu$). The width of the search window ($\theta_{Search}$) affects how accurately we can predict these conditions.
5.3.2 Galactic Potentialities, Cost, Gain, and Risk: Statistical Thought Experiments

In this sub-section, we explore the degree to which different search strategies can alter our ability to address questions of galactic life. The analysis will be in the form of double ‘what if’ thought experiments. The distribution of life in the galaxy is unknown but different potentialities can be considered (e.g., ‘what if life potential is not peaked around Earth’) to evaluate how different search strategies perform under variable possibility space (e.g., ‘what if a targeted search strategy is used under this potential life probability distribution’). The mathematical details of the experiments can be found in the appendix. Below we lay out the conceptual approach.

We assume that conditions describing a galactic body can be expressed as an index. The difference between the index values for any galactic body and the Earth defines an Earth-difference metric, $\theta$. A more precise definition for $\theta$, in terms of delineating all the physical/chemical factors that could feed into it, is not required for the analysis that follows. Figure 5.2 illustrates how $\theta$ is used in our thought experiments. The analysis is referenced to our own planet as $\theta_{Earth}$. We assume there is a range of conditions ($\theta_{min}$ to $\theta_{max}$) over which planetary life is possible. Galactic bodies are grouped into bins, denoted as $\theta_i$, each having a fixed width. Each bin holds an equal number of objects, an assumption that can be relaxed. Of those objects, some will be inhabited and some will not. Based on this premise, a probability index, characterizing the life potential of each set of conditions, is assigned to each bin. Effectively, we are assigning a habitability index, akin to that of Barnes et al. [2015], to each bin. The bin with the highest probability index is defined to be the set of conditions that maximizes life potential. The position of this bin, on the Earth-difference scale, is denoted by $\mu$.

The procedure above allows a large number of distributions to be generated. As a starting point, we choose a normal distribution to model the frequency of inhabited planets based on their specified conditions. We then vary $\mu$ and create different probability distributions. Here, $\mu$ is varied between 0 and 0.5 at an increment of 0.1. Allowing the life
potential peak to vary is resonant with Heller and Armstrong [2014] who have argued that Earth conditions may not be those that maximize life potential. To be clear, we are not arguing that this is or is not the case. Our stance is that at present we do not know. We consider it a possibility in the same sense that it is possible that Earth conditions do indeed maximize life potential. Effectively we are considering multiple working hypotheses and exploring the implications of each for different search strategies.

Variable synthetic distributions, constructed as per above, represent different galactic potentials. The number of observations we have to date do not allow for discrimination between different potentials. For our experiments we will assume that future observations can be used toward this end and we will ask how different search strategies can achieve it. A search strategy is defined in terms of a search window centered about $\theta_{\text{Earth}}$ and extending to a value of $\theta_{\text{Search}}$ in both directions. The greater the extent, the more we are willing to search for inhabited, non Earth-like planets. Increasing values of $\theta_{\text{Search}}$ are evaluated for their ability to recover the mean of the inhabited distribution, referred to as $\mu_{\text{test}}$, and to provide accurate estimates of galactic life.

Within this framework, we constructed different galactic life distribution potentials and considered different search strategies for each. For any distribution, we drew from it at random, subject to a specific search window, observing what conditions defined each object and whether they were inhabited or not. The number of observations needed to find a fixed number of inhabited objects was tracked. This number can be varied and we will also consider search strategy performance if the goal is to only find a single inhabited planet. If the goal is to provide some constraints on galactic life distribution then more than a single find would be required. The minimum number of observations needed to potentially constrain a distribution is not a fixed, agreed upon value. As a starting point, we settled on 30 inhabited objects - this number is only a ‘rule of thumb’ for the minimum number of observations needed to potentially constrain a distribution [Hogg and Tanis, 1997]. Once this number of observations was obtained, the conditions that maximize life were estimated for varying search window widths (differences in the estimates could then
be compared to actual distributions to gauge uncertainty as a function of search strategy). This process was performed a number of times, tracking the number of observations and life potential maximizing conditions that were arrived at each time. We then assessed the accuracy and cost of each search strategy. Accuracy was defined as the percent error between the conditions predicted by each search strategy and the true solution. Cost is related to the number of observations. The more observations needed, the greater the cost. Some function could be devised to approximate how this cost translates to total time and resources. Here we assume that the number of observations provides a useful starting point in considering relative costs (we appreciate that the scaling between number of observations and monetary cost will likely be a non-linear).

Figure 5.3: a) A cost-benefit relationship based on search window width (color of dot) and how different the conditions can be from Earth and allow for life (larger dots are further from Earth-like). b) Accuracy is assessing how prevalent life is within the galaxy ($P_{galaxy}$) versus search window width for different potential galactic distributions ($\mu_{test}$).

For the initial suites of experiments, the probability of life amongst galactic bodies at the distribution peak was set to 20%. The probability then dropped toward zero as conditions moved away from the peak toward the most extreme cases that maintained life.
potential (Figure 5.2). This equated to assuming that roughly 5% of galactic objects that reside within the window of planetary conditions that allow for life would be inhabited. That number can and will be varied extending to 25% and less than 1%.

The first scenario considered was one under which the conditions that maximize life potential are centered on Earth conditions. In this case, a focused search strategy (i.e. a narrow search window) is ideal for finding a minimum number of inhabited planets in a cost efficient way (Figure 5.3a). It can under predict how prevalent life is in the galaxy by about a third (Figure 5.3b). Doubling the search window can bring the estimate of galactic life down to the lowest possible error at approximately twice the cost of a narrow search.

If the conditions maximizing life potential are different from Earth then the accuracy and efficiency of different search windows changes (Figure 5.3). The changes for some search windows are large while for others they are relatively mild (i.e., different strategies have different levels of robustness). When the optimal conditions are similar to Earth ($\mu_{test} = 0.1$), the error of the narrowest search strategy in estimating galactic life is not significantly increased (Figure 5.3b) and cost remains low (Figure 5.3a). The gains associated with moving to an intermediate search window are similar to the Earth centered case (Figure 5.3b).

If conditions that maximize life potential deviate further from Earth ($\mu_{test} = 0.3$) then changes in the cost-benefit relationships, for variable search strategies, can become significant. For the narrowest window, accuracy continued to decrease while cost, in terms of number of observations, increased relatively rapidly. The cost of a narrow search exceeded that of an intermediate search as $\mu_{test}$ exceeded 0.3. For those cases, a narrow search window leads to significant bias. As a result, a lot of observations (and time) are required and the prediction they lead to regarding galactic life potential is in significant error. This is a dangerous combination. For $\mu_{test} = 0.3$, going from a narrow to an intermediate search can lower the error in assessing galactic life potential from over 70% to under 10% with close to zero increase in cost. For $\mu_{test} > 0.3$, going to an intermediate search can lower both error and cost.
Thus far, we have focused on the contrast between narrow and intermediate search strategies. Extending the search to observe everything that we think is capable of harboring life increases accuracy over all potentialities. This comes with increased cost. Over most of the potentiality space we have explored the increased cost in going from an intermediate to a wide search does not come with a significant increase in accuracy. For the most extreme cases tested, there is a significant increase in accuracy. For those cases, the accuracy increase for a wide relative to an intermediate search is a factor of 4 with a factor of 1.25 increase in cost. In short, the principal gain in going from an intermediate to a wide search is that the risk, in terms of misrepresenting galactic life, for a “worst-case” situation is minimized.

The absolute number of observations needed to find a fixed number of inhabited planets, for any search strategy, depends on the assumed percentage of inhabited galactic bodies. Figure 5.4 shows results from experimental suites that vary that value under the goal of finding 30 (Figure 5.4a) or a single inhabited planet (Figure 5.4b). For the experiments of Figure 5.4a, the accuracy in assessing galactic life follows the same trends as in Figure 5.3b.

Figure 5.4: Results from experiments, as per Figure 5.3, that vary the total percentage of inhabited galactic bodies.
The relative difference between search strategies remains robust for different assumptions as to the total number of inhabited planets. The exception is that as the total number becomes very low the potential advantage of a wide search becomes weaker (Figure 5.4a, lower right corner). If the total number of inhabited planets is very low, then the number of planets that must be assessed to find more than a single inhabited planet becomes large under all search strategies. If that is the case, then we are pushing toward the limit at which statistical analysis is not justified.

To this point we have assumed that finding a single inhabited planet is not the sole goal for explorations of galactic life. Nonetheless, it is useful to consider time commitment toward the first milestone within the overall search endeavor. The results of Figure 5.4b can be viewed in this way. They also have potential utility if the total number of inhabited planets in the galaxy is so low that efforts to determine a ‘life potential distribution’ will be effectively doomed from the start. If finding a single inhabited planet is the sole goal and/or if it turns out that a very large number of planets need to be assessed before we find any signs of life on one of them, then Figure 5.4b is pertinent while issues of accuracy as per Figure 5.3b are less pertinent. There are some difference between a single planet and multi planet goal, particularly if the total number of inhabited galactic bodies is low. However, the main trend we are highlighting remains robust: a narrow search can be more efficient under some galactic life potentials but it comes with greater risk when evaluated over a broad range of potentials that extend beyond the assumption that Earth conditions maximize life potential.

5.3.3 Changing Minds

Different ideas about life in our galaxy (e.g., rare versus plentiful) can be viewed as competing hypotheses or, in a probabilistic framework, different *a priori* assumptions that have not been fully tested (in this case because observations are currently limited). Each end-member idea is a viable hypothesis. Each is testable. That is, each is a scientific prior. Any scientific prior will adjust to new observations and, given a large number of
observations, all such priors should converge toward the hypothesis that is most consistent with observations. In principal, it does not matter what prior assumptions a search strategy is based on - the observations will decide in the end. In practice, we need to consider that the observations we will make over the next wave of missions will be discrete in number. Using the Earth-likeness metric of the previous sub-section we can ask which hypothesis is more responsive to discrete new observations.

The discussion above relates to search strategies. One can imagine that a scientist who holds to a rare-Earth, or Earth-centered, view would opt for a narrow search window on the grounds that this would be the most efficient strategy. On the other hand, a scientist who holds to the view that a wide range of conditions can allow for planetary life might argue against a narrow search on the grounds that it will introduce a bias. We can pose the question of how many discoveries of inhabited planets would be required to change either scientist’s mind if their hypothesis happens to be incorrect.

![Figure 5.5: The number of discoveries, of inhabited planets, it takes to converge to the attributes that maximize inhabitance for different priors and for different life probability potentials.](image)

The question above may seem highly subjective. However, asking the question of how a priori ideas or beliefs adjust to new observations lends itself to a Bayesian analysis (Appendix), which can provide quantitative insight. For an Earth-centered prior, initial confidence begins tightly peaked around $\theta_{Earth}$. A plentitude prior assigns equal weight
across some $\theta$-range. As in our previous analysis, we can draw observations randomly and track the conditions of the body and whether or not it is inhabited. If the body is indeed inhabited, both end-member priors are updated to reflect this. This is repeated and the number of observations tracked until each scientist’s confidence is in alignment with the true set of conditions that maximize life potential within an equal tolerance. This is quantified using a maximum a posterior (MAP) metric [Robert, 2007]. That metric is defined as the parameter with the maximum posterior probability (further details can be found in the appendix).

Figure 5.55 shows how long it takes each end-member hypothesis to converge for different potential galactic life distributions. If galactic conditions are Earth centered, it would not take long to alter either a priori view. If the conditions are different than Earth, it takes more observations for both priors to converge to the true solution. This increase occurs because our first data point is always Earth. That initial data constraint is a restatement of our current understanding of life in the galaxy: we have only observed one inhabited body, Earth. If conditions maximizing life potential are far different from Earth, then our planet can introduce an initial bias. A plentitude view adjusts relatively quickly, regardless of our planets galactic status. An Earth-centered prior has a slower adjustment in the face of new observations, i.e., it comes with a greater resistance. If galactic life potential is not Earth centered, then going in with the assumption that it is, as opposed to the assumption that a wider range is possible, will lead to $\sim 6$ times more observations being required to remove the shadow of the a priori assumption.

5.4 Discussion

In the previous section we used an Earth-likeness (equivalently a nonEarthness) metric to evaluate the cost, accuracy, and risk associated with different search strategies and the more subjective cost associated with competing \textit{a priori} hypothesis regarding planetary life potential. We employed a generic metric to provide a quantitative measure of the degree of nonEarthness that maintains life potential. In the statistical thought experiments,
the results were dependent on the value of this metric to variable degrees - in some scenarios that cannot be observationally ruled out the degree of dependence was large. An upshot is the suggestion that there is value in developing, as a community, a more tightly defined metric. This could provide some nonEarth1.0 balance to Earth2.0 thinking and discussions. The balance could have scientific and humanistic implications. In the following sub-sections, we wade into these discussion topics, not as a final say, but as a starting point.

5.4.1 Different is More

What is the value of considering the potential of inhabited, nonEarth-like planets in the search for life? What is the value of finding such a planet? The first question was at the core of the previous section. A key implication was related to risk minimization, which we expand on below. We then move to the second question. We focus in on Gaia theory and on how exoplanet exploration could provide a useful Gaia test bed that extends beyond Earth.

Search strategies should be able to achieve goals in cost efficient ways that minimize risk in the face of worst-case scenarios. What is the worst-case scenario for investing resources to assess galactic life potential? Is it finding no signs of life beyond Earth? That is a worst case if one believes that it is a negative conclusion. If the goal is to reach a conclusion we can be as confident of as we can be, then no conclusion is negative if it is the one that follows from the preponderance of un-biased evidence. We would argue that the worst case would be spending a large amount of time and resources to reach the wrong conclusion. Figure 5.3 indicates that an overly narrow search can open us to that risk.

The above can be viewed as an overly “formalized” assessment. Exploration depends on public interest, which doesn’t fit as easily into formalized analysis. From that angle, a worst-case scenario could be viewed as one that tries public patience (although we would argue that an equally valid worst-case scenario would be misinforming the public). This could motivate the idea that we need to find signs of an inhabited planet beyond Earth as
soon as we can (the opening statements quoted in the introduction reflect that, rightly or wrongly, this has already been effectively promised to the public). This, in turn, could be used to argue for an Earth-centered search based on the idea that it would minimize time commitment. The difficulty is that we would then be making a time investment decision based on assuming we know the answer to a critical question before we put the investment strategy into practice. This again opens us to risk as a narrow search could come with a higher time cost if we are wrong in terms of our *a priori* assumption (Figure 5.4b).

In regards to assumptions, all of our results required assumptions regarding galactic life distributions. This, we would argue, does not weaken our conclusions but is an overarching point of the analysis to begin with. The tighter a search focuses around an Earth2.0 mode, the more a soft form of rare-Earth is being assumed to be valid (i.e. life potential peaks at Earth). One could argue that models of planetary habitability and observations from our solar system, to date, lend support to this assumption. Fair enough but it remains an assumption. It may prove correct but if it is used to justify a search strategy then the strategy is designed, effectively, to test only one hypothesis. One could take an alternate view. Rather than arguing for one hypothesis over another (e.g., life potential peaking near or away from Earth [e.g., Heller and Armstrong, 2014], we can assume that a range of hypothesis remain viable and seek to minimize risk in the face of the various potentialities. This is the view we have argued for. Considering the potential of inhabited non-Earth like planets can minimize risk at an increased cost that, in our opinion, is not excessive (Figures 5.3 and 5.4). It also allows for the potential of discriminating between different hypotheses regarding galactic life (Figure 5.5).

We now turn to the question of “what is the value of finding an inhabited planet that is non Earth like in other regards?” The classic concept of a “Habitable Zone” assumes that delineating planetary conditions that allow for the potential of life can proceed without considering life’s role [Kasting et al., 1993]. In effect, it is assumed that habitability can be determined without explicitly considering inhabitance. This assumption has been challenged and the degree to which removing it from exoplanet discussions could influence
our thinking about life in our galaxy is significant [Goldblatt, 2016]. The difficulty has been and remains that observations from this planet can not unravel the degree to which life influences cycles that regulate environmental conditions (the Earth has life and remains geologically active - which of these is more critical to the Earth being habitable is difficult to unravel as life has entwined itself in a range of geophysical/geochemical cycles [Goldblatt, 2016]).

From an exoplanet perspective, we can push Gaia theory to a limit. If a strong form of Gaia can operate then planetary regulation could occur without abiotic cycles (Figure 5.1c). That is, life could do the heavy lifting for maintaining habitable conditions. The implication is that although planetary internal energy, that drives volcanism and tectonics, plays a role for the Earth’s inhabitance it is not critical for planetary bodies in general. The internal energy of a planet will depend on its age and composition. The potential of determining composition and age for exoplanets is actively being discussed within the community. Both could be within reach for next generation observations. If a planet is found that has low potential of being geologically active and shows biosignatures this would provide a step toward confirming Gaia (internal energy may still have been crucial for the origin of life [Baross and Hoffman, 1985] and for maintaining habitable conditions early in the planets lifetime but the extension of habitable conditions past the geologic lifetime of the planet would be due to life itself). At present, search strategies are focused on planets that are likely to be geologically active with the thought that this is critical for life [Ward and Brownlee, 2000, Kasting, 2012]. That remains an assumption. It is an assumption that has the potential to be refuted if a single strong Gaia1.0 is found (a greater number of Earth2.0’s would be required to confirm the assumption at a statistical level).

Finding strong Gaia1.0 would dramatically change our views about planetary habitability (finding Earth2.0 would be a confirmation of a prevalent idea). The degree of rethinking can be hinted at by posing a question: Is habitability a characteristic of a planet, like temperature, or is it something that flows through it, like heat? Stated another way, is it a state or a process variable? The classic “habitable zone” concept assumes that it
can be treated as a state variable. From that, follows the assumption that its limits can be determined so as to effectively make a phase diagram that delineates regions that allow for life. On a strong Gaia the origin and evolution of life are dominant for habitability. With that comes contingency and the potential of multiple temporal paths leading to variable end states of inhabitance [Walker et al., 2018]. This is a process variable view [Bridgman, 2013] under which multiple equilibrium states are possible and path-dependence cannot be ignored [Dyke and Weaver, 2013, Weaver, 2015, Lenardic et al., 2016]. This brings in layers of potentiality associated with evolution and historical contingency (to be clear, a contingent process is not the same as a random process [e.g., Bohm, 2004]). The approach to planetary life research would, as a result, need to move toward one that is more statistical/probabilistic than it is at present [Walker et al., 2018].

The issue of evolution leads to a final point. Rare-Earth ideas acknowledge that planets different from the Earth could have simple (microbial) life but argue that higher life (plants and animals) requires conditions like that of Earth [Ward and Brownlee, 2000]. Life can respond to environmental changes. Many of the changes, that are considered critical to the development of higher life on Earth, are ascribed to internal energy sources driving changes in surface conditions such that if a planet lacked the geological changes that occurred on Earth it could have microbial life but it would not have developed higher life [Ward and Brownlee, 2000, Stern, 2016]. The idea that evolution requires environmental changes is not agreed upon for the evolution of life on Earth [McKee, 2000] and, as such, extending it to planetary bodies in general is an a priori assumption. Exoplanet search strategies have incorporated the potential that atmospheric bio-signatures might be of the kind that prevailed on early Earth, before the rise of complex life, or of the kind associated with higher life [Schwieterman et al., 2018]. Finding signs of higher life on a geologically inactive planet could provide a new layer of evidence that evolution can proceed in an autocatalytic mode with no need for externally driven environmental changes [Kauffman, 1992, McKee, 2000, Gatti, 2011].

Finding signs of life on another planet, be it like our own or significantly different,
would be a major discovery. The implications of that discovery could be further reaching for planets that are different. In that sense, different is more - it could bring more information content about life potential in our galaxy.

5.4.2 Narratives

The search for life beyond Earth is no small undertaking. It can benefit from efforts to engage the public. The engagement is often framed as a narrative. Narratives can go well beyond public relations. A narrative can frame a problem in a way that favors one decision/conclusion over alternatives that are just as rational as the frame-favored decision [Tversky and Kahneman, 1981]. For the issue at hand there is a feedback as public opinion can influence exploration strategies. Whether it is intended or not, building a narrative around a problem will influence the way humans think about the problem and, in keeping with the cultural/humanistic connections of this sub-section, ‘what we think changes how we act’ [c.f., Gang of Four, Solid Gold, EMI/Warner Bros., 1981]. In short, this encapsulates the value of re-framing problems and considering multiple frames when addressing problems involving uncertainty [Tversky and Kahneman, 1988].

An Earth2.0 narrative that has been used to frame exoplanet exploration reinforces the idea that Earth conditions are the ideal ones for habitability (a variant of a rare-Earth narrative [Ward and Brownlee, 2000]. We would argue that we do not have the observations needed to discriminate between different assumptions regarding galactic life potential at this stage of our exploration and it is not in the best interest of the search to send messages, explicit or implicit, that we do. An Earth2.0 narrative, as it is being presented to the public, walks the line of promising specific returns and the dangers of that for science, in the public realm, should be kept in mind [e.g., Westfall, 2016].

Beyond sending messages that do not accurately reflect the state of our knowledge, an Earth2.0 narrative comes with philosophical and cultural baggage that may not best serve its intended purpose. An Earth2.0 narrative is nostalgia heavy. There have been many words put forward in the service of an Earth2.0 narrative but images, arguably, can give a
better sense of the messages this narrative carries. Artwork depicting travel posters to exoplanets with white picket fences [https://exoplanets.nasa.gov/alien-worlds/exoplanet-travel-bureau/] invoke a sense of the familiar - a sense of home (note: only a sense of home for those who grew up culturally and economically in places that had white picket fences to begin with). The desire for a place just like home is fed by nostalgia and a desire for an ideal [Messeri] - an ideal that may never have existed. The potentially false narrative of an ideal past has been taken up in artistic works (e.g., some movie examples: Pleasantville, Midnight in Paris, Trainspotting 2) and in historical studies [e.g., Webb, 2017]. We mention these works to point out that, to a portion of humanity, a nostalgia filled narrative may not have the effects hoped for.

We are not implying that the cultural value of nostalgia in general, or specifically for developing public narratives, can be broken down to 1’s and 0’s. It is not as simple as it’s always good or always bad. It depends on context [e.g., Bonnett, 2010]. Within the context of space exploration, a narrative built on finding a second Earth is intertwined with the idea of finding a second home [Messeri]. Home is something we know, something comfortable, and ideas of home can lead to thoughts of "better times". This can send the message that our goals are to recapture something. It invokes a sense of looking toward the past, which runs counter to the idea of exploration. When pushed to limits, such nostalgic messages are associated with populist movements. From our perspective, this also runs counter to the idea of space exploration, which is a global endeavor (i.e., an internationalist as opposed to a national endeavor).

An alternate narrative, as compared to searching for an Earth2.0, is one of galactic diversity in terms of livable and living planets. In this alternative narrative, the future becomes more prominent with all the lack of certainty, immediate comfort and familiarity that a future holds. Planets with life beyond Earth may be something foreign to us. They may be uncomfortable for us to live on at first. To know them we need to find them, as opposed to starting our exploration on the premise we already know them (we are not the first to remark on the dangers of assuming we know the answer before hand when it comes
to planetary exploration [Moore et al., 2017, Tasker et al., 2017]). The alternate framework we are proposing will also not resonate across all of humanity as a public engagement narrative but which of the two, diversity of living planets versus finding a second Earth, better represents the sense of exploration that got us, as human beings, to begin exploring space in the first place?

5.5 Appendix: Methods

5.5.1 Statistical Cost-Benefit-Risk Analysis

We used a statistical analysis to address our motivating question: How does search window width affect the cost-benefit relationship based on the conditions that maximize life potential? In the analysis, we tested three search window widths - narrow, intermediate and wide, which are defined by the $\theta_{\text{Search}}$ values 0.2, 0.6 and 1.0, respectively. For each search window we tested six different life distributions, each having increasingly nonEarth-like life potential maximizing means ($\mu_{\text{test}}$), ranging from 0.0 to 0.5. For each combination of $\theta_{\text{Search}}$ and $\mu_{\text{test}}$, we conducted 100 trials that randomly sampled a synthetic data set, characterized by the conditional probability $P(I|\Theta)$, the probability that an object is inhabited given a specific set of conditions. We assumed that each subset of conditions, broken up into discrete intervals ($\theta_i$), 0.1 units in width, were equally likely to represent any observable object. Within each $\theta_i$ interval, some of the objects were inhabited and some were not. The percentage of inhabited bodies in each $\theta_i$ interval was approximated as the probability for that bin given a normal distribution defined by $\mu_{\text{test}}$ and a standard deviation ($\sigma$) of 0.2, which we held constant.

For each combination of $\theta_{\text{Search}}$ and $\mu_{\text{test}}$, a set of trials were performed to estimate the sample statistics. In each trial, we used a uniform random number generator to determine the bin from which an observation would be drawn (subject to limits imposed by the search window). Then, we did a second random draw, from that bin, and tracked whether the observation was of an inhabited or uninhabited object (the greater the habitability probability from that bin, the greater the chances that an inhabited object would be
observed). During the duration of the trial, we tracked the total number of observations, the conditions that defined each observed object, how many objects were accumulated in each interval \( \theta_i \) and whether or not each object was inhabited. Once 30 inhabited objects were observed, we calculated the mean \( \bar{x} \) and standard deviation (s) of the sample. Following the 100 different, randomized trials, we calculated the means \( \mu_{\bar{x}}, \mu_s, \mu_{n_i} \) and \( \mu_{n_0_i} \) of the sampled statistics. In this paper, we refer to \( \mu_{\bar{x}} \) as the observed mean, \( \mu_s \) as observed standard deviation, \( \mu_{n_i} \) as the number of observed inhabited objects in the interval \( \theta_i \) and \( \mu_{n_0_i} \) as the total number of observations in the interval \( \theta_i \).

Using the results of the statistical analysis, we calculated the percent error between the true (\( P_{\text{galaxy}} \)) and predicted number of inhabited galactic objects (\( P^*_{\text{galaxy}} \)) for each combination of \( \theta_{\text{Search}} \) and \( \mu_{\text{test}} \). The statistical analysis results tell us the average ratio between the number of inhabited objects and total observed objects is defined as

\[
\frac{\mu_{n_i}}{\mu_{n_0_i}}
\]

for each \( \theta_i \). Equation 5.1 can be scaled to the entire population by the relation

\[
\frac{\mu_{n_i}}{\mu_{n_0_i}} = \frac{N_i}{N_{\theta_i}}
\]

where \( N_i \) is the total number of inhabited objects in the interval and \( N_{\theta_i} \) is the total number of objects in the interval. Using our mean of the distribution of the means (\( \mu_{\bar{x}} \) and \( \mu_s \)), we can approximate the probability distribution of life conditions as

\[
P(I|\Theta) = N(\mu_{\bar{x}}, \mu_s).
\]

Furthermore, for each interval, we know that

\[
P(I|\theta_i) = \frac{N_i}{P_{\text{galaxy}}},
\]

where \( P_{\text{galaxy}} \) is the number of objects in the galaxy that are inhabited. Combining equations
5.1, 5.2 and 5.4, we arrive at our estimate of the galactic population

\[ P_{\text{galaxy}}^* = \frac{N_0 n_i}{\mu n_0} P(I|\theta_i). \]  

(5.5)

Following the same procedure, except using \( \mu_{\text{test}} \) and \( \sigma \) rather than \( \mu_{\bar{x}} \) and \( \mu_s \) in equation 5.3, the true galactic population is

\[ P_{\text{galaxy}} = \frac{N_0 n_i}{\mu n_0} P(I|\theta_i). \]  

(5.6)

where \( n_i \) is the true number of bodies present in the interval \( i \) if a representative sample were taken of size \( \mu n_0 \). That is

\[ n_i = \mu n_0 P(I|\theta_i). \]  

(5.7)

Substituting equation 5.7 into 5.6 we get

\[ P_{\text{galaxy}} = N_0. \]  

(5.8)

This relationship is a result two main assumptions: (1) that every object is equally likely to be characterized by any set of conditions \( \theta_i \), and (2) the proportion of objects that are inhabited within a given interval \( i \) is represented as the probability of that interval for the normal distribution characterized by \( \mu_{\text{test}} \) and \( \sigma \). To obtain the percent error \( (\epsilon) \), we use the relation

\[ \epsilon = \left| \frac{P_{\text{galaxy}} - P_{\text{galaxy}}^*}{P_{\text{galaxy}}} \right|. \]  

(5.9)

which, after combining with equations 5.6 and 5.9, simplifies to

\[ \epsilon = \left| 1 - \frac{\mu n_i}{\mu n_0} P(I|\theta_i) \right|. \]  

(5.10)
5.5.2 Changing Minds

Bayesian inference provides a framework to evaluate how new data can alter a priori assumptions that define different end-member search strategies. Assumptions regarding our current state of knowledge are used to assign probabilities to a particular set of beliefs known as the prior, \( P(\theta) \), where \( \theta \) is a discrete vector representing parameters. As new data \( (D) \) are discovered, the chances of them occurring, given that a particular parameter is correct, is then computed to produce the likelihood function, \( P(D|\theta) \). The likelihood is telling us the probability of getting the data given the set of conditions. The prior and likelihood function are then combined to produce our updated knowledge, known as the posterior \( P(\theta|D) \), which is interpreted as the probability that parameter \( \theta \) is correct, given the observed data set \( D \). These three components are related by

\[
P(\theta|D) \propto P(D|\theta)P(\theta)
\] (5.11)

In our analysis, \( \theta \) represents how Earth-like a body is and \( D \) represents the number of such galactic bodies that are inhabited. Earth-likeness is a sliding parameter scale normalized to Earth at a value of zero and extending to negative and positive infinity. We consider a finite range between some assumed minimum and maximum value, \( \theta_{\text{min}} \) and \( \theta_{\text{max}} \) (\( \theta_{\text{min}} \) and \( \theta_{\text{max}} \) are assigned normalized values of -1.0 and 1.0, respectively, for the most extreme version of a plentitude hypothesis). We consider differing values of Earth-likeness at intervals of 0.1, totaling 21 different types of Earth-likeness bins.

We use different priors to represent competing hypotheses for the range of earth-likeness an inhabited planet can have. One prior assigns a high probability around zero in \( \theta \)-space and much lower values everywhere else. It is represented as a normal distribution with a mean value of zero and variance of 0.1. The second prior assumes that all earth-likenesses, within our finite range of consideration, are equally probable to be inhabited.

Discoveries of future hypothetically inhabited planets are used to produce the likelihood function. It is assumed that the Earth-likeness of each new discovery can be categorized.
Using this data, the standard Gaussian functional form is used to derive the likelihood function,

\[ P(D|\theta) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(\theta - D_i)^2}{2\sigma_i^2} \right] P(\theta) \]  \hspace{1cm} (5.12)

where we assume each new data point is independent of any other, potentially a limiting assumption as the sampling may be highly selective. This allows for the multiplication of the prior and likelihood function to produce the posterior. The posterior is then normalized such that the sum of probabilities from all bins equal one.

A plentitude hypothesis uses an uninformative prior, which allows the data to directly influence the posterior distribution. The posterior is computed by calculating the likelihood function and normalizing the sum of probabilities to one. On the other hand, an Earth-centered prior is normally distributed and is multiplied with the likelihood function producing a posterior that is also normally distributed. The mean and variance of this posterior is computed as

\[ a = \frac{1}{\sigma_{\text{prior}}^2}, \]  \hspace{1cm} (5.13)

\[ b = \frac{n}{\sigma^2}, \]  \hspace{1cm} (5.14)

\[ \mu_{\text{post}} = \frac{a\mu_{\text{prior}} + b\bar{x}}{a + b}, \]  \hspace{1cm} (5.15)

\[ \sigma_{\text{post}}^2 = \frac{1}{a + b}, \]  \hspace{1cm} (5.16)

where \( \sigma, \bar{x} \) and \( n \) are the known uncertainty of the likelihood, the mean value of the data and number of data points, respectively. Once the probabilities have been computed, they must be normalized to sum together to one.

To gauge the convergence of a given prior we use the maximum a posteriori (MAP), defined as the parameter, which has a maximum posterior probability [Robert, 2007]. The MAP is used to calculate the convergence of end-member priors to a final solution. The number of discoveries it takes for the MAP to equal the assumed mean of \( \theta \), given random draws from an assumed data set, defines the “convergence metric” plotted in Figure 5.5.
To test the convergence of each prior, we assume a normally distributed subset of 1,000 inhabited planets (other values can be tested - our initial interest is a relative comparison between two end-member priors and, as such, absolute numbers for convergence are not particularly meaningful as they are overly dependent on the assumed number of inhabited bodies). The mean of different potential galactic distributions ($\mu_{\text{test}}$) is varied between 0 and 1 in $\theta$-space with a standard deviation of 0.3. Planets are drawn at random, following Earth as the initial data point, and the number of draws it takes to converge to a solution is tracked. The process is repeated for each $\mu_{\text{test}}$ and each prior 150 times and the average number of draws, from the subset of 150, is reported as the convergence metric for each prior under each variable galactic life potential (Figure 5.5).
References


C. Herzberg and P. D. Asimow. PRIMELT3 MEGA.XLSM software for primary magma calculation: Peridotite primary magma MgO contents from the liquidus to the solidus.


J. Korenaga. Thermal evolution with a hydrating mantle and the initiation of plate
tectonics in the early Earth. *Journal of Geophysical Research: Solid Earth*, 116(12):20,

e1601168–e1601168, aug 2016. ISSN 23752548. doi: 10.1126/sciadv.1601168. URL
http://advances.sciencemag.org/cgi/doi/10.1126/sciadv.1601168.

L. R. Kump, S. L. Brantley, and M. A. Arthur. Chemical Weathering, Atmospheric CO₂,
ISSN 0084-6597. doi: 10.1146/annurev.earth.28.1.611. URL


A. Lenardic. The diversity of tectonic modes and thoughts about transitions between
them. In *Philosophical Transactions of the Royal Society A: Mathematical, Physical and
Engineering Sciences*, volume 376, page 20170416. The Royal Society Publishing, nov

A. Lenardic. Volcanic-Tectonic Modes and Planetary Life Potential. In *Handbook of
Exoplanets*, pages 2897–2916. Springer International Publishing, Cham, 2018b. doi:
10.1007/978-3-319-55333-7_65. URL


D. P. Fleming and J. VanderPlas. approxposterior: Approximate Posterior Distributions in


