FROM THE MOON TO PLUTO: THE USE OF IMPACT AND CONVECTION MODELING AS A WINDOW INTO PLANETARY INTERIORS

by

Alexander J. Trowbridge

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Earth, Atmospheric, and Planetary Sciences

West Lafayette, Indiana

August 2020

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. David A. Minton, Chair
School of Earth, Atmospheric, and Planetary Sciences

Dr. Andrew M. Freed
School of Earth, Atmospheric, and Planetary Sciences

Dr. H. Jay Melosh
School of Earth, Atmospheric, and Planetary Sciences

Dr. Brandon C. Johnson
School of Earth, Atmospheric, and Planetary Sciences

Approved by:
Dr. Daniel J. Cziczo

Dedicated to the memory of my grandfather, John Peter Zaremba Jr.

ACKNOWLEDGMENTS

I gratefully acknowledge the developers of iSALE-2D (www.isalecode.de), including Gareth Collins, Kai Wünnemann, Dirk Elbeshausen, Tom Davison, Boris Ivanov, and Jay Melosh. Some plots in this work were created with the pySALEPlot tool written by Tom Davison. In addition, I acknowledge the efforts of David Blair and Kevin Graves, who developed several codes used to support this research. This dissertation was supported by the NASA Earth and Space Science Fellowship (80NSSC18K1265) and the Fredrick N. Andrews Fellowship.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Planetary science is often limited to only surface observations of planets requiring the development of modeling techniques to infer information about the planet's interior. This work outlines three separate scientific problems that arose from planetary surface observations, the methodology utilized to explain the formation of these observation, and what we learned about the planet's interior by solving these problems.

Chapter 1 discusses why lunar mascon basins (impact basins associated with a central free-air gravity positive) form for only a limited range of basin diameters. Modeling the full formation of South-Pole Aitken (SPA) basin using a sequential two-code (hydrocode and Finite Element Model) shows that due to SPA's great size (long wavelength) and the high geothermal gradient of the Moon at impact, the basin's relaxation process was controlled by isostatic adjustment with minimal influence from lithospheric rigidity or membrane stresses. Additionally, the modeling shows that the Moon was hot and weak at impact.

Chapter 2 addresses why there is a lack of olivine abundance on Mars around large impact basins, and the formation of the megabreccia that is associated with an orthopyroxene signature in the circum-Isidis Planitia region. Hydrocode modeling of the excavation of the Isidis forming impact shows the impact was more than capable of excavating mantle material and reproducing the observed megabreccia. This coupled with the lack of olivine signature indicates that the Martian upper mantle is orthopyroxene-rich.

Chapter 3 covers the investigation into why the nitrogen ice sheet on Pluto, Sputnik Planitia, is the youngest observed terrain and why the surface is divided into irregular polygons about 20–30 kilometers in diameter. The utilization of a new parameterized convection model enables the computation of the Rayleigh number of the nitrogen ice and shows that the nitrogen ice is vigorously convecting, making Rayleigh–Bénard convection the most likely explanation for these polygons (Trowbridge et al., 2016). Additionally, the diameter of Sputnik Planitia's polygons and the dimensions of its 'floating mountains' of water ice suggest that its nitrogen ice is about five to ten kilometers thick (Trowbridge et al., 2016). The estimated convection velocity of 1.5 centimeters a year indicates a surface age of only around a million years (Trowbridge et al., 2016).

The accumulation of this work is three chapters that use three separate techniques to further understand three separate planets.

INTRODUCTION: A HISTORICAL PERSPECTIVE

The Mystery of the Lunar Craters: Galileo to GRAIL

In 1609-1610, Galileo Galilei used the newly developed telescope to identify mountains and craters on the lunar surface, sparking the beginning of understanding planetary surfaces (Galileo, 1609). Explaining the origin of the craters observed by Galileo might be the first attempt by the scientific community to try and understand how a feature on another planet formed. In 1665, Robert Hooke conducted experiments observing a pot of boiling Alabaster and noticed craters forming on its surface as a result of bursting bubbles (Hooke, 1665, p. 243). Additionally, Hooke dropped musket balls into a mixture of Tobacco-pipe clay and water and noted that after removing the musket ball the resulting circular depressions were "exactly like these of the Moon" (Hooke, 1665, p. 243). As a result of these experiments, Hooke proposed two origin hypotheses for the lunar craters: (1) an internal volcanic source like the boiling Alabaster; (2) and an impact origin as a result of falling heavenly bodies onto the Moon like his musket balls (Hooke, 1665, p. 243). However, having no explanation for where these heavenly bodies would come from disregarded the impact theory (Hooke, 1665, p. 243).

The first creditable support for the impact origin came from Grove Gilbert in 1893 (Gilbert, 1893). Gilbert compared the morphology of lunar craters to volcanic craters on Earth (i.e. depth to diameter ratios, the depth of the floor of the crater compared to the outer plain, the height of the "central hill" of the lunar crater, etc.) and concluded that only impacts could have caused them (Gilbert, 1893, pg. 246-253). Gilbert recognized a crater size-dependent morphology, illustrating that as craters increase in size, central peaks are observed within the crater's interior and at further crater diameters the craters transition into basins with flat-floored centers (Gilbert, 1893, pg. 245).

Gilbert conducted low velocity impact experiments reproducing the observed lunar craters with vertical impacts (Gilbert, 1893, pg. 261). He also noted that as you decrease the impact angle the craters get more elliptical in shape (Gilbert, 1893, pg. 261) and determined that the most probable impact angle for a meteoric body would be 45 degrees, which results in an oval scar (Gilbert, 1893, pg. 261-263). As almost all observed lunar craters were circular, these impact experiments remained unconvincing to the general scientific community (Melosh, 1989).

In 1924, Algernon Gifford compared meteorite impacts to buried explosions showing that the same kinetic energy is released between an object moving nearly two miles per second and an explosion of an equal mass of TNT (Gifford, 1924; Gifford, 1931). Gifford inferred that high velocity meteor impacts could generate circular craters regardless of impact angle and that the meteorites were much smaller than the craters they formed (Gifford, 1924; Gifford, 1931).

Despite this, the volcanic origin hypothesis for lunar craters persisted but began to wane as the general scientific community gained more insight from terrestrial impact craters. In 1906, Daniel Barringer observed on top of the southern rim of the Coon Butte Crater of Arizona (now known as Meteor Crater) an inverted deposit of fragments from the yellow and red sandstone that sit at the lowest strata (Barringer, 1906). Additionally, he observed pieces of metallic iron scattered on the surface outside the crater, and correctly deduced that the crater was formed from an impact of a large nickel-iron meteorite; ironically, Grove Gilbert studied this same crater earlier but concluded it was of volcanic origin (Barringer, 1906).

Barringer's work set a standard to which future geologists would identify new craters with potential impact origin (Baldwin, 1978). In 1927, Ivan Reinvaldt conducted a detailed geologic survey of the Kaali-järv Crater in Estonia and determined the crater was impact in origin (Reinvaldt, 1928). In 1928, Daniel Barringer Jr. identified similar meteoritic nickel-iron as Meteor

Crater along the crater rims of craters in Odessa, Texas; craters discovered in 1921 (Barringer, 1928). In 1931, Malcolm Maclaren impressed by the raised rims of the Bosumtwi (Ashanti) crater in Ghana proposed an impact origin of a stony meteorite traveling at around 50 miles per second (Maclaren, 1931); although, there was no meteoritic material found (Spencer, 1933). Published in 1932, Arthur Alderman proposed an impact origin for the craters at Henbury, Australia after discovering a great number of metallic meteorite fragments scattered over the area of several craters (Alderman, 1932). This same year, Harry St. John Philby discovered two distinct craters (Wabar craters) in the Arabian Desert and deduced their impact origin from their raised rims and metallic iron fragments scattered outside these craters (Philby, 1933). In 1933, Leonard Spencer wrote a summary of discovered craters, detailing the depth to diameter ratios, compared them to mine craters formed during WW1, and proposed the Camp del Cielo craters in Argentina should be added to the list of known meteorite craters (Spencer, 1933).

Soon larger terrestrial craters would be discovered further supporting the impact origin of lunar craters. In 1936, Boon and Albritton identified localized shattering, and a lack of volcanic materials around the previously called "cryptovolcanoes," now known as central peak craters (Boon and Albritton, 1936). These structures were much larger than the previously identified craters. Additionally, Boon and Albritton calculated that a body of only 250 feet in diameter and with an impact velocity of 19 miles per second was needed to eject all the mass to form Meteor Crater (Boon and Albritton, 1938). This calculation combined with the discovery of the larger impact craters reinforced the concept that a small object traveling at great speed could generate a much larger sized crater. In 1946, using the insight gained by the work of Boon and Albritton, Robert Dietz revaluated the lunar craters' origin and argued that due to the size, shape, and distribution of the lunar craters they must have been impact in origin (Dietz, 1946). In 1949, Ralph

Baldwin showed the similarity between the crater depth and diameter on a log-log scale for lunar craters, terrestrial meteoritic craters, and explosion pits (Baldwin, 1949). Additionally, Baldwin showed that the depth and diameter plots of the terrestrial volcanic and lunar craters had very little correlation (Baldwin, 1949). In a summary of early impact history, Baldwin wrote it was this figure that became the most convincing argument for the impact origin for lunar craters (Balswin, 1978).

As explosive technology advanced, so did our understanding of impacts. In 1955, the 1.2-kiloton Teapot ESS nuclear weapon was detonated (Melosh, 1989). E. M. Shoemaker compared the crater generated by this explosion to his geologic survey of Meteor Crater and developed the foundations of modern impact cratering (Shoemaker, 1960; Shoemaker 1963). Testing his theories on the lunar craters, Shoemaker would publish a compelling argument for the impact origin of the lunar craters (Shoemaker, 1962).

The impact origin of Galileo's craters had essentially been settled, but as our observational technology advanced the lunar craters revealed further questions. From 1966 through 1967 the United States sent five unmanned lunar orbiter missions to map the gravity and topography of the Moon's surface, intending to help select an appropriate landing site for the Apollo missions (Muller and Sjogren, 1968). Much to the surprise of the scientific community, these lunar orbiter probes accelerated towards some of the bigger impact basins. This is because these basins contained mass concentrations, or mascons, as indicated by positive free-air gravity anomalies, denoting excess mass being supported superisostatically (Muller and Sjogren, 1968; Neumann et al., 1996). During Apollo 16, orbital perturbations due to the mascons cut the life of its subsatellite to a mere 35 days (Apollo 16 Mission Report, 1972).

The formation of mascon basins remained a mystery for decades, motivating several origin theories, including that of a buried nickel-iron mass, presumably from the impactor (Muller and

Sjogren, 1968). Another mascon origin theory suggests that mascons are a result of flexural support of mare loads; however, it was shown that the nearside mare basins would still be mascons even if the gravitational contribution of the mare deposits were removed (Neumann et al., 1996). Neumann et al. (1996) suggested that the mascons are formed from support of a superisostatic configuration caused by mantle rebound during transient crater collapse. This explanation ignores the fact that basins retain a large thermal anomaly after impact, making support by a cool, strong lithosphere unlikely (Melosh et al., 2013). It has since been shown that impact basins are invariably not in a superisostatic configuration following transient crater collapse (Freed et al., 2014). The final proposed origin for the lunar mascons was flexural uplift of the center of a basin into a superisostatic position during post-transient crater collapse cooling and isostatic adjustment (Andrews-Hanna, 2013). In this scenario, a strong lithospheric bridge develops between the basin center and a subisostatic crustal collar, enabling the basin center to rise above isostatic equilibrium as the crustal collar rises do to isostatic forces (Andrews-Hanna, 2013). However, this theory was difficult to numerically test, as little was known about the mechanical state of basins immediately after transient crater collapse (Melosh et al., 2013).

Solving the mystery of these basins was one of the primary goals of the Gravity Recovery and Interior Laboratory (GRAIL) mission, which measured the lunar gravity and collected the best gravity data of any celestial body to date (Zuber et al., 2013). With this new data, topography data from the Lunar Orbiter Laser Altimeter (LOLA), and new advancements in computer modelling of impacts, the mystery of the formation of mascons was finally solved. Numerical modeling studies (Melosh et al, 2013; Freed et al., 2014) of the complete evolution of two lunar mascons, the ~200 km diameter Freundlich-Sharonov and Humorum basins, demonstrated how a superisostatic basin can develop after a large impact. These studies showed that following transient

crater collapse the basins emerged in a sub-isostatic state due to low topography at the basin center surrounded by a thickened crustal collar. Low pressure associated with subisostatic topography and the buoyancy of the depressed crust resulted in mantle flow toward the basin center from outside the basin, resulting in uplift that brings basin topography closer to isostatic equilibrium. Since such pressure gradients will dissipate as isostatic equilibrium is approached, there is generally no mechanism from which isostatic forces can raise topography beyond isostatic equilibrium and into a superisostatic state.

However, in the millions of years following impact, cooling of the basin leads to the development of a lithosphere inside the basin that mechanically couples the inner basin to the crustal collar region; this occurs prior to completion of isostatic adjustment for these medium-sized basins. As the crustal collar continued to rise, this lithospheric bridge elevated the inner basin above isostatic equilibrium into a superisostatic state, or a mascon. This is possible because the crustal collar region represents a significantly larger volume of initially subisostatic material, thus enabling the rising crustal collar to easily lift the inner basin above isostatic equilibrium. Lithospheric stiffness, however, prevents the depressed crustal collar itself from reaching isostatic equilibrium. Therefore, it remains in the subisostatic state we observe today. In the outer basin, lithospheric support of higher topography associated with impact ejecta surrounds the entire basin with a superisostatic annulus. Thus, these lunar basins are associated with a bullseye free-air gravity anomaly, whether or not the basin is filled with mare.

While the general formation mechanism for mascons has been solved, there still remains the unexplained observation that mascon basins are generally found only within a limited range of diameters: basins smaller than ~200 km are not typically observed to be mascons, nor is the largest basin, South Pole-Aitken (SPA) (~2300 km in diameter). Smaller basins are generally thought to

not form into mascons because they never develop sufficient isostatic forces to overcome lithospheric rigidity. Chapter 1 of this work is in an investigation into why very large lunar basins do not form into mascons. Modeling the full evolution of South-Pole Aitken using a sequential two-code approach (hydrocode and Finite Element Model) shows that due to SPA's great size (long wavelength) and the high geothermal gradient of the Moon at impact, the basin's relaxation process was controlled by isostatic adjustment with minimal influence from lithospheric rigidity. Additionally, the modeling shows that the Moon had to be hot and weak at impact.

Composition of the Martian Upper Mantle

In 1871, the Italian astronomer Giovanni Schiaparelli pointed his telescope to Mars to map its surface. He drew long, straight lines across Mars, which he named canali (meaning channels), leading to an interpretation that Mars had river networks and an ocean, a thick atmosphere, and vegetation (Zahnle, 2001). This theory stayed popular within the scientific community until Mariner 4 flew by Mars in 1964, observing no water on the surface, heavily cratered terrains, a thin atmosphere, and showed that these canali were a result of an optical illusion associated with poor resolution (Zahnle, 2001).

Although disproving the notion of vegetation on Mars, Mariner 4 led to a new misconception that Mars was a Moon-like planet with a surface riddled with impact craters (Zahnle, 2001). In 1969, the flyby of Mariners 6 and 7 furthered this notion by observing a heavily cratered terrain (McCauley et al., 1972). Finally, in 1971, Mariner 9 photographed 80 percent of Mars' surface revealing a geologically diverse landscape with massive volcanoes (i.e. Tharsis Regio), canyon systems (i.e. Valles Marineris), and ancient dried out river beds (McCauley et al., 1972).

More information of Martian geology and its interior came with further missions. In 1976, Viking 1 and Viking 2 arrived at Mars and began measuring (among other things) the composition of the Martian atmosphere, showing a CO_2 denominated planet (Nier and McElroy, 1977). In 1983, trace gases in meteorites previously found in Antarctica, called the SNC meteorites, were compared to the composition of the Martian atmosphere collected by the Viking spacecrafts suggesting that these meteorites were from Mars (Bogard and Johnson, 1983). These SNC meteorites proved important for petrological modeling of the Martian mantle. Assuming the SNC meteorites are partial melts derived from the mantle, many models have been developed to represent the primitive mantle of Mars, all of which show a Martian mantle that is more Fe-rich than Earth (Smrekar et al., 2019). Notably, there is the more olivine-rich upper mantle model of Dreibus and Wänke (1985), which is currently the most widely accepted compositional model (Smrekar et al., 2019), and the pyroxene-rich upper mantle model (Sanloup et al., 1999).

In 1988, the Imaging Spectrometer for Mars instrument on Phobos-II enabled mineralogical mapping of the surface of Mars (Sagdeev and Zakharov, 1989). Later analysis of this data identified two pyroxene-basalts near presumed volcanoes, which had a composition consistent with the mafic minerology of the SNC meteorites (Mustard et al., 1997). The conclusion was that the Martian mantle was relatively depleted in Al early in Martian formation, which could have come about through the formation of a basaltic crust (Mustard et al., 1997). In 1997, the Mars Pathfinder mission landed a rover (Sojourner) on the surface to measure the composition of the Martian rocks (Golombek et al., 1997). Using the Alpha Particle X-ray Spectrometer, Sojourner identified rocks consistent with both basaltic and andesitic parent materials, pointing out that the high silica content of some of the rocks appears to require crustal differentiation of mantle-derived parent materials (Golombek et al., 1997).

Within the same year, the Mars Global Surveyor went into orbit around Mars taking higher resolution images, and making topography and gravity measurements (Zuber et al., 2000). This newly collected data revealed large mascon basins, similar to the lunar mascons, such as Isidis Planitia and Argyre (Zuber et al., 2000). Additionally, the gravity data collected constrained previous crustal thickness estimates, i.e. a global mean of 50 km with the thinnest crust (~3 km) occurring within Isidis Planitia (Zuber et al., 2000).

In the early 2000s, Mars Express, using the Visible and Infrared Mineralogical Mapping Spectrometer (OMEGA) instrument, mapped the abundance of mafic minerals on the surface of Mars (Mustard et al., 2005). OMEGA observed olivine and HCP-rich regions spanning an age range of geologic units; however, the LCP-rich (orthopyroxene) regions were only found in oldest aged units (Mustard et al., 2005). This was interpreted as LCP-rich melts derived from a mantle depleted in Al and Ca (Mustard et al., 2005). Additionally, the OMEGA dataset showed olivine in the northern rims of large basins, such as Hellas, Isidis, and Argyre, and was interpreted to be impact excavated mantle material (Ody et al., 2013). Other groups also interpreted the olivine deposit in the circum-Isidis region as impact melt based upon the OMEGA dataset (Mustard et al., 2007).

In 2006, the Mars Reconnaissance Orbiter (MRO) entered orbit around Mars collecting the highest resolution images and spectral data to date of the surface. The High-Resolution Imaging Science Experiment instrument on board imaged megabreccia blocks in the circum-Isidis region (Mustard et al., 2009). Analysis of these megabreccia blocks place their origin to the Isidis impact event (Weiss et al., 2018). Additionally, the Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) instrument on the MRO showed that the breccia blocks consisted of unaltered LCP-

rich (orthopyroxene) mafic rocks together with rocks showing signatures of Fe/Mg-phyllosilicates (Mustard et al., 2009).

Recent analysis of the olivine unit across Mars reinterprets the deposition history as a global volcanic pyroclastic deposit; therefore, suggesting the olivine in this region is not ejecta from the Isidis forming impact (Kremer et al., 2019). Crater scaling for a basin the size of Isidis predicts a maximum excavation of 50-100 km, which is deeper than the mean crustal thickness of Mars (Neumann et al., 2004). Moreover, the thinnest crust on Mars is always within the center of Isidis Planitia regardless of parameters choosen to produce the crustal thickness map (Smrekar et al., 2019). This leads to the question addressed in Chapter 2, if the Martian upper mantle is olivine-rich (predicted by the widely accepted Dreibus and Wänke (1985) model), and the Isidis forming impact excavated mantle material, where is the olivine in the circum-Isidis region from the impact event and why does the observed megabreccia correspond with an LCP-rich (orthopyroxene) spectral signature? Hydrocode modeling of the excavation of the Isidis forming impact shows the impact was more than capable of excavating mantle material and reproducing the observed megabreccia. This coupled with the lack of olivine signature indicates that the Martian upper mantle is orthopyroxene-rich, consistent with the pyroxene-rich upper mantle model (Sanloup et al., 1999).

Pluto's Geologic Activity

In 1834, Thomas Hussey documented a conversation with another astronomer, Alexis Bouvard, that the perturbation of an unknown planet may cause the irregular motions of Uranus (Grosser, 1964). In 1846, these same motions of Uranus led Urbain Le Verrier to calculate the location of a planet he believed was responsible (Grosser, 1964). The same night Verrier sent these

coordinates to an astronomer colleague, Johann Gottfried Galle, and the planet Neptune was discovered (Kemp, 1990). Over the next century, many scientists (e.g., D. P. Todd, G. Forbes, M. C. Flammarian, and Perceival Lowel) showed that Neptune alone did not explain the motions of Uranus and predicted a trans-Neptunian planet (Chhabra et al., 1982). In 1930, Clyde Tombaugh discovered this trans-Neptunian planet, which would later be known as Pluto (Kemp, 1990).

Earlier work on Pluto was limited due to uncertainty in its mass, diameter, and bulk density; however, observations in the mid-1970s would improve these uncertainties (Cruikshank et al., 1997). In 1976, estimates of Pluto's diameter and bulk density were improved by photometric detection of CH_4 (Cruikshank et al., 1976), which indicated that Pluto was more ice-like than rock and sparked speculation on the presence of a Pluto atmosphere (Cruikshank et al., 1997). Additionally, the discovery of Pluto's satellite Charon in 1978 led to determining the mass of Pluto (Christy and Harrington, 1978). Spectroscopic observations confirmed frozen CH_4 on the surface of Pluto, which combined with these new estimates for the mass of Pluto further constrained the bulk density and radius of Pluto (Cruikshank and Silvaggio, 1980).

More information was gained about Pluto's surface after the detection of Pluto's atmosphere via the observed effects on a stellar occultation in 1988 (Millis et al., 1988; Elliot et al., 1989). Soon after, spectral observations led to the discovery of N_2 ice (Owen et al., 1993), CO ice (Owen et al., 1993), and H_2O ice (Owen et al., 1996). The fact N_2 ice has an absorption coefficient $\sim 10^5$ less than CH_4, combined with the trace amounts of CO and H_2O observed, led to the conclusion that N_2 ice was the dominate ice on the planet (Owen et al., 1993; Cruikshank et al., 1997).

Further constraints on the distribution of ice on Pluto came from albedo maps. The first maps produced used a collection of circular spots on Pluto and a constant albedo on Charon to model the lightcurve of the Pluto-Charon system; this revealed that Pluto's surface was heterogenous in

albedo, having bright polar caps (Marcialis 1988; Buie and Tholen 1989). By modeling the lightcurve of the Pluto-Charon system using data collected during eclipse and transit events, Buie et al. (1992) further constrained the albedo of Pluto; however, the biggest improvement would come about by a series of works that observed Pluto using the Hubble Space Telescope (Stern et al. 1997; Buie et al., 2010). At a much higher imaging resolution, the Hubble Space Telescope showed that not only were the poles of Pluto bright, but the brightest region of the planet was actually at 0º latitude, 180º longitude (Buie et al., 2010). The accumulation of knowledge of the composition of Pluto up until this point would correctly predict that this brightest region was likely a large reservoir of ice predominately N_2 in composition; however, no one predicted the extent to which this region would be geologically active (Stern et al., 2015).

In 2015, the New Horizons spacecraft flew by Pluto getting the highest resolution images of its surface ever recorded (Stern et al., 2015). Prior to the New Horizons flyby, Pluto was believed to be most similar to Triton, though perhaps less geologically active given the lack of tidal heating that Neptune provides (Moore et al., 2015). Remarkably, that brightest region of the planet observed by the Hubble Space Telescope turned out to be a large nitrogen ice reservoir devoid of any observable craters (Stern et al., 2015), called Sputnik Planitia, which is interpreted to be an impact basin. Preliminary crater counting placed the age of this surface at less than 10 million years old, implying an incredibly young and geologically active surface (Stern et al., 2015). Nearly the entire surface of this region was observed to be divided into irregular polygons about 20–30 kilometers in diameter, whose centers rise tens of meters above their sides (Stern et al., 2015). Both thermal contraction and convection were initial proposed to explain this terrain (Stern et al., 2015; Trowbridge et al., 2016).

The formation of these Pluto polygons is the subject of Chapter 3. In this chapter, the utilization of a new parameterized convection model enables the computation of the Rayleigh number of the nitrogen ice and shows that the nitrogen ice is vigorously convecting, making Rayleigh–Bénard convection the most likely explanation for these polygons (Trowbridge et al., 2016). Additionally, the diameter of Sputnik Planitia's polygons and the dimensions of its 'floating mountains' of water ice suggest that its nitrogen ice is about five to ten kilometers thick (Trowbridge et al., 2016). The estimated convection velocity of 1.5 centimeters a year indicates a surface age of only around a million years (Trowbridge et al., 2016).

Overview

From a historical perspective, we have often had to reevaluate what we think we know about other planets; whether that be the canali on Mars, the origin of the lunar craters, or the geologic activity of Pluto. Each of the chapters of this work explains a unique scientific problem observed on separate planets: the size dependence of the lunar mascons (Chapter 1), the apparent lack of abundance of olivine around Isidis Planitia on Mars and the formation of the observed megabreccia (Chapter 2), and the existence of the Pluto polygons and their surprisingly young surface (Chapter 3). While each of these works seemingly have little relation to each other, the accumulation of this work provides an addition to our ever-evolving knowledge of the planets.

References

Andrews-Hannah (2013). The origins of the non-mare mascon gravity anomalies in lunar basins, *Icarus* **222**, 159-168.

Alderman, A. R. (1932). Meteorite craters at Henbury, Central Australia. *Mineralog. Mag.* **23**, 19.

Apollo 16 Mission Report (August, 1972). NASA, MR-11, Wash., D.C.

Baldwin, R. B. (1978). An overview of impact cratering. *Meteoritics* **13**, 364-379.

Barringer, D. M. (1906). Coon Mountain and Its Crater. *Acad. Nat. Sci. Philadelphia* **57**, 861.

Barringer Jr, D. M. (1928). A new meteor crater. *Proceedings of the academy of natural sciences of Philadelphia* **80**, 307-311.

Bogard, D. D., and Johnson, P. (1983). Martian gases in an Antarctic meteorite?. *Science* **221**, 651-654.

Boon, J.D., and Albritton, C. C., Jr. (1936). Meteorite craters and their possible relationships to "cryptovolcanic" structures. *Field and Lab*. **5**, 1-9.

Boon, J.D., and Albritton, C. C., Jr. (1938). The Impact of Large Meteorites, *Field and Laboratory* **6**, 56-64.

Buie, M. W., and Tholen, D. J. (1989). The surface albedo distribution of Pluto. *Icarus* **79**, 23-37.

Buie, M. W., Tholen, D. J., and Horn, K. (1992). Albedo maps of Pluto and Charon: Initial mutual event results. *Icarus* **97**, 211-227.

Buie, M. W., Grundy, W. M., Young, E. F., Young, L. A., and Stern, S. A. (2010). Pluto and Charon with the Hubble Space Telescope. II. Resolving changes on Pluto's surface and a map for Charon. *The Astronomical Journal* **139**, 1128-1143.

Christy, J. W, and Harrington, R. S. (1978). The satellite of Pluto. *Astron. J.* **83**, 1005-1008.

Cruikshank, D. P., Pilcher, C. B., and Morrison, D. (1976). Pluto: Evidence for methane frost. *Science* **194**, 835-837.

Cruikshank, D. P., and Silvaggio, P. M. (1980). The surface and atmosphere of Pluto. *Icarus* **41**, 96-102.

Cruikshank, D. P., Roush, T. L., Moore, J. M., Sykes, M. V., Owen, T. C., Bartholomew, M. J., ... and Tryka, K. A. (1997). The surfaces of Pluto and Charon. *Pluto and Charon* **1**, 221-268.

Dreibus, G., and Wanke, H. (1985). Mars, a volatile-rich planet. *Meteoritics* **20**, 367-381.

Elliot, J. L., Dunham, E. W., Bosh, A. S., Slivan, S. M., Young, L. A., Wasserman, L. H., and Millis, R. L. (1989). Pluto's atmosphere. *Icarus* **77**, 148-170.

Freed, A. M., et al. (2014), The formation of lunar mascon basins from impact to contemporary form, *J. Geophys. Res. Planets* **119**.

Galilei, G. (1609). *Sidereus Nuncius, or the sidereal messenger.* Reprint, Chicago: University of Chicago Press, 2016.

Gifford, A. C. (1924). The mountains of the moon. *New Zealand J. Sci. and Technol.* **7**, 129.

Gifford, A. C. (1931). The origin of the surface features of the moon. *Journal of the Royal Astronomical Society of Canada* **25**, 70.

Gilbert, G. K. (1893). *The Moon's face: A study of the origin of its features*. Philosophical Society of Washington.

Golombek, M. P., Cook, R. A., Economou, T., Folkner, W. M., Haldemann, A. F. C., Kallemeyn, P. H., ... and Rieder, R. (1997). Overview of the Mars Pathfinder mission and assessment of landing site predictions. *Science* **278**, 1743-1748.

Grosser, M. (1964). The search for a planet beyond Neptune. *Isis* **55**, 163-183.

Hoefen, T. M., Clark, R. N., Bandfield, J. L., Smith, M. D., Pearl, J. C., and Christensen, P. R. (2003). Discovery of olivine in the Nili Fossae region of Mars. *Science* **302**, 627-630.

Hooke, R. (1665). *Micrographia, or, Some physiological descriptions of minute bodies made by magnifying glasses: With observations and inquiries thereupon.* London: Printed by J. Martyn and J. Allestry.

Kemp, K. W. (1990). Pluto and the patterns of planetary discovery. *Astron. Quarterly* **7**, 19-33.

Kremer, C. H., Mustard, J. F. and Bramble, M. S. (2019). A widespread olivine-rich ash deposit on Mars. *Geology* **47**, 677–681.

Maclaren, M. (1931). Lake Bosumtwi, Ashanti. *Geog. J.* **78**, 270.

Marcialis, R. L. (1988). A two-spot albedo model for the surface of Pluto. *The Astronomical Journal* **95**, 941-947.

McCauley, J. F., Carr, M. H., Cutts, J. A., Hartmann, W. K., Masursky, H., Milton, D. J., ... and Wilhelms, D. E. (1972). Preliminary Mariner 9 report on the geology of Mars. *Icarus* **17**, 289-327.

Melosh, H. J. (1989). *Impact cratering: A geologic process.* New York, Oxford University Press.

Melosh, H. J., et al. (2013). The Origin of Lunar Mascon Basins, *Science* **340**, 1552-1555.

Millis, R. L. et al. (1988). Observations of the 9 June 1988 occultation by Pluto. *Bull. Amer. Astron. Soc.* **20**, 806.

Moore, J. M., Howard, A. D., Schenk, P. M., McKinnon, W. B., Pappalardo, R. T., Ewing, R. C., ... and Buratti, B. (2015). Geology before Pluto: Pre-encounter considerations. *Icarus* **246**, 65-81.

Muller, P. M., and Sjogren, W. L. (1968). Mascons: Lunar mass concentrations. *Science* **161**, 680-684.

Mustard, J. F., Murchie, S., Erard, S., and Sunshine, J. (1997). In situ compositions of Martian volcanics: Implications for the mantle. *Journal of Geophysical Research: Planets* **102**, 25605-25615.

Mustard, J. F., Poulet, F., Gendrin, A., Bibring, J. P., Langevin, Y., Gondet, B., ... and Altieri, F. (2005). Olivine and pyroxene diversity in the crust of Mars. *Science* **307**, 1594-1597.

Mustard, J. F. *et al.* (2007). Mineralogy of the Nili Fossae region with OMEGA/Mars Express data: 1. Ancient impact melt in the Isidis Basin and implications for the transition from the Noachian to Hesperian. *J. Geophys. Res. E Planets* **112**, 1–14.

Mustard, J. F., Ehlmann, B. L., Murchie, S. L., Poulet, F., Mangold, N., Head, J. W., ... and Roach, L. H. (2009). Composition, morphology, and stratigraphy of Noachian crust around the Isidis basin. *Journal of Geophysical Research: Planets* **114**.

Neumann, G. A., M. T. Zuber, D. E. Smith, and F. G. Lemoine (1996). The lunar crust: Global structure and signature of major basins. *J. Geophys. Res*. **101***,* 841–863.

Neumann, G. A. *et al.* (2004). Crustal structure of Mars from gravity and topography. *J. Geophys. Res. E Planets* **109**, 1–18.

Nier, A. O., and McElroy, M. B. (1977). Composition and structure of Mars' upper atmosphere: Results from the neutral mass spectrometers on Viking 1 and 2. *Journal of Geophysical Research* **82**, 4341-4349.

Ody, A., Poulet, F., Bibring, J. P., Loizeau, D., Carter, J., Gondet, B., and Langevin, Y. (2013). Global investigation of olivine on Mars: Insights into crust and mantle compositions. Journal of Geophysical Research: Planets **118**, 234-262.

Owen, T. C., Roush, T. L., Cruikshank, D. P., Elliot, J. L., Young, L. A., de Bergh, C., Schmitt, B., Geballe, T. R., Brown, R. H. and Bartholomew, M. J. (1993). Surface ices and atmospheric composition of Pluto. *Science* **261**, 745-748.

Owen, T. C., Cruikshank, D. P., Roush, T., Geballe, T. R., and de Bergh, C. (1996). Water ice on Pluto. In preparation.

Philby, H. St. J. (1933). Rub'al Khali: An account of exploration in the Great South Desert of Arabia. *Geog. J.* **81**, 1-26.

Reinvaldt, I. (1928). Bericht über geologische Untersuchungen am Kaalijärv (Krater von Sall) auf Ösel. Mit Beiträgen von A. Luha. Sitzungsber. Naturfors. *Gesell. Univ. Tartu* **35**, 30-70.

Sagdeev, R. Z., and Zakharov, A. V. (1989). Brief history of the Phobos mission. *Nature* **341**, 581-585.

Sanloup, C., Jambon, A., and Gillet, P. (1999). A simple chondritic model of Mars. *Physics of the Earth and Planetary Interiors* **112**, 43-54.

Sjogren, W. L., Wimberly, R. N., Wollenhaupt, W. R. (1974). *Moon* **9**, 115.

Shoemaker, E. M. (1960). Penetration mechanics of high velocity meteorites, illustrated by Meteor Crater, Arizona., In *Rept. of the Int. Geol. Congress*, XXI Session, Norden. Part XVIII, 418-434, Copenhagen.

Shoemaker, E. M. (1962). Interpretation of lunar craters. In Z. Kopal (Ed.), *Physics and Astronomy of the Moon*, Academic Press, New York and London, 283-359.

Shoemaker, E. M. (1963). Impact mechanics at Meteor Crater, Arizona. In: *The Moon, Meteorites, and Comets.* (The Solar System, **4**), 301-336, Univ. of Chicago Press, Chicago.

Smrekar, S. E., Lognonné, P., Spohn, T., Banerdt, W. B., Breuer, D., Christensen, U., ... and Garcia, R. F. (2019). Pre-mission InSights on the Interior of Mars. *Space Science Reviews* **215**, 3-75.

Spencer, L. J. (1933). Meteorite craters as topographical features on the earth's surface. *The Geographical Journal* **81**, 227-243.

Stern, S. A. (1989). Pluto: Comments on crustal composition, evidence for global differentiation. *Icarus* **81**, 14-23.

Stern, S. et al. (2015). The Pluto system: initial results from its exploration by New Horizons. *Science* **350**.

Treiman, A. H., Gleason, J. D., and Bogard, D. D. (2000). The SNC meteorites are from Mars. *Planetary and Space Science* **48**, 1213-1230.

Urey, H.C. (1952). On the early chemical history of the Earth and the origin of life. *Proc. Natl. Acad. Sci. USA* **38**, 351–363.

Weiss, B. *et al.* (2018). Megabreccia at Northeast Syrtis Major and its importance for mars science. in *49th Lunar and Planetary Science Conference*.

Wieczorek, M. A., and Phillips, R. J. (1999). *Icarus* **139**, 246.

Wylie, C. C. (1934). Meteoric Craters, Meteors, and Bullets. *Pop. Astr.* **42**, 469-71.

Zahnle, K. (2001). Decline and fall of the Martian empire. *Nature* **412**, 209–213.

Zuber, M. T., Solomon, S. C., Phillips, R. J., Smith, D. E., Tyler, G. L., Aharonson, O., ... and Lemoine, F. G. (2000). Internal structure and early thermal evolution of Mars from Mars Global Surveyor topography and gravity. *Science* **287**, 1788-1793.

Zuber, M. T., et al. (2013). Gravity field of the Moon from the Gravity Recovery and Interior Laboratory (GRAIL) mission, *Science* **339**, 668–671.

CHAPTER 1. WHY THE LUNAR SOUTH POLE-AITKEN BASIN IS NOT A MASCON

This chapter was accepted to the journal *Icarus* on 21 July 2020.

1.1 Abstract

Lunar mascon basins exist for only a range of observed diameters. We modeled the full formation of South-Pole Aitken basin using a sequential two-code (hydrocode and Finite Element Model) approach to understand why this range does not extend to the largest lunar basin. Similar to previous work, we found that the best-fit hydrocode had impact parameters of a 170 km in diameter dunite projectile striking at 10 km/s, with a pre-impact lunar thermal gradient of 50 K/km (until a depth of 28 km at which an adiabat is reached), and pre-impact crustal thickness value of 40 km. Unlike previous work, we matched the crustal distribution of the inner basin by utilizing a weaker melt rheology in our model. The crust in the models with a weaker melt rheology flowed inward after crater collapse to cover the basin center; therefore, not requiring an additional step of melt differentiation. Given the large diameter of South-Pole Aitken and the high geothermal gradient at impact, the steady state configuration after relaxation and thermal cooling was driven by isostatic forces without a significant contribution from lithospheric rigidity. Without lithospheric rigidity, South-Pole Aitken relaxed to a slightly negative free-air gravity signature, not forming a mascon, in agreement with GRAIL observations.

1.2 Introduction

From 1966 through 1967, the United States sent five unmanned lunar orbiter missions to map the Moon's surface, intending to help select an appropriate landing site for the Apollo missions (Muller and Sjogren, 1968). When a spacecraft orbits a body, the along-track velocity

increases as it approaches a mass surplus and decrease as it approaches a mass deficit; therefore, it was expected that the orbiting lunar probes would decrease in their along-track velocity as they approached the large impact basins that cover the lunar surface. The lunar orbiter probes, however, remarkably accelerated towards some of the bigger impact basins. This is because these basins contained mass concentrations, or mascons, as indicated by positive free-air gravity anomalies, denoting excess mass being supported superisostatically (Muller and Sjogren, 1968; Neumann et al., 1996). Later, on Apollo 16, orbital perturbations due to the mascons cut the life of its subsatellite to a mere 35 days (Apollo 16 Mission Report, 1972).

Mascon basins (e.g., A, C, and D in Figure 1.1) are characterized by a positive free-air gravity anomaly (~100 mGals for Freundlich-Sharonov; red in Figure 1.1) at their center, often surrounded by an annulus of negative free-air gravity anomalies (~-200 mGals for Freundlich-Sharonov; blue in Figure 1.1), which in turn is sometimes surrounded by an outer positive free-air gravity anomaly (~200 mGals for Freundlich-Sharonov) to form a bullseye pattern (Melosh et al., 2013, Freed et al., 2014). Mascon basins are generally found only within a limited range of diameters: basins smaller than ~200 km are not typically observed to be mascons (e.g., E in Figure 1.1), nor is the largest basin, South Pole-Aitken (SPA; 2500 km in diameter; B in Figure 1.1). In general, basins between these sizes can be divided into two groups: (1) mascon basins with significant mare fill, and (2) empty (or nearly empty) mascon basins. Empty basins develop into mascons because of lithospheric support of a superisostatic configuration of shallow mantle at the basin centers (Melosh et al., 2013; Freed et al., 2014). Mare basins are mascons due to lithospheric support of the mare fill (Baldwin, 1968; Neumann et al., 1996), and additionally, in some cases, lithospheric support of basin uplift as well (Melosh et al., 2013; Freed et al., 2014).

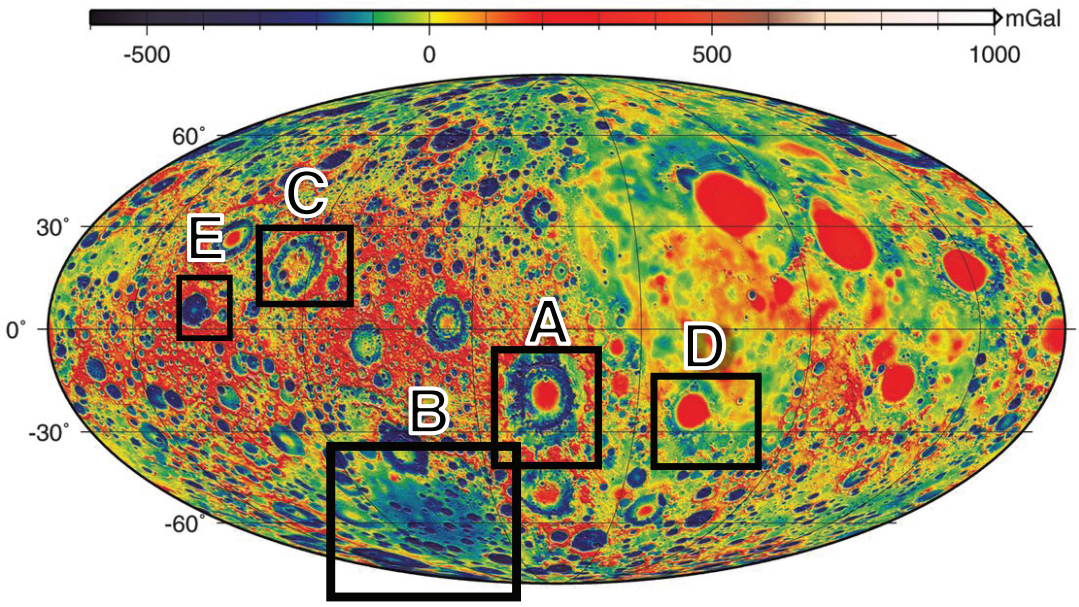

Figure 1.1. Lunar free-air gravity map showing various impact basins: (A) Orientale (a mascon); (B) South Pole-Aitken Basin (non-mascon); (C) Freundlich-Sharonov (empty mascon); (D) Humorum (mare filled mascon); and (E) Mendeleev crater (non-mascon). The gravity map is derived from GRAIL lunar gravity model GL0420A, to spherical harmonic degree and order 420, adapted from Fig. 1 of Zuber et al. (2013). The map is an Molleweide projection centered on 270°E longitude and show the nearside on the right and farside on the left. Gravity is plotted in units of milliGalileos, where 1000 mGal = 1 cm/s^2.

Recent numerical modeling studies (Melosh et al, 2013; Freed et al., 2014) of the complete evolution of two lunar mascons, the ~200 km diameter Freundlich-Sharonov and Humorum basins (C and D in Figure 1.1, respectively), demonstrated how a superisostatic basin can develop after a large impact. These studies showed that, following transient crater collapse, the basins emerged in a sub-isostatic state due to low topography at the basin center surrounded by a thickened crustal collar (Figure 1.2) (Andrews-Hannah et al., 2013). Low pressure associated with subisostatic topography and the buoyancy of the depressed crust result in mantle flow toward the basin center from outside the basin, resulting in uplift that brings basin topography closer to isostatic equilibrium. Since such pressure gradients will dissipate as isostatic equilibrium is approached,

there is generally no mechanism from which isostatic forces can raise topography beyond isostatic equilibrium and into a superisostatic state.

However, in the millions of years following impact, cooling of the basin leads to the development of a lithosphere inside the basin that mechanically couples the inner basin to the crustal collar region; this occurs prior to completion of isostatic adjustment. For mid-sized basins (Figure 1.2a), as the crustal collar continues to rise, this lithospheric bridge elevates the inner basin above isostatic equilibrium into a superisostatic state, or a mascon. This is possible because the crustal collar region represents a significantly larger volume of initially subisostatic material, thus enabling the rising crustal collar to easily lift the inner basin above isostatic equilibrium. Lithospheric stiffness, however, prevents the depressed crustal collar itself from reaching isostatic equilibrium. Therefore, it remains in the subisostatic state we observe today. In the outer basin, lithospheric support of higher topography associated with impact ejecta surrounds the entire basin with a superisostatic annulus. Thus, a mid-sized lunar basin is associated with a bullseye free-air gravity anomaly, whether or not the basin is filled with mare.

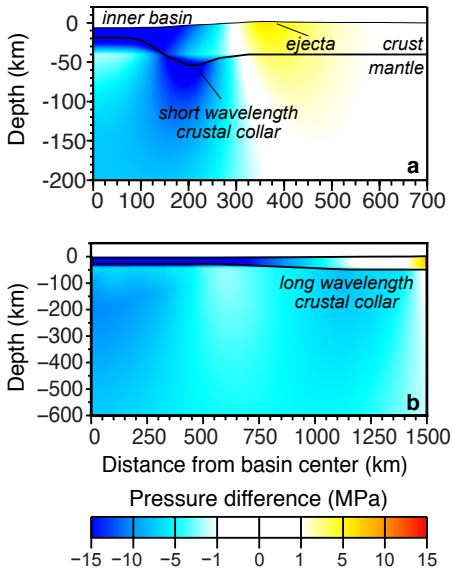

Figure 1.2. Overburden pressure difference between pressure at depth for the post transient crater collapse crustal structure of Freundlich-Sharonov Basin (a) and South-Pole Aitken Basin (b) and a nominal pressure at distance. Note the change of scales between the two panels. Cool colors correlate with areas that are subisostatic while warm colors are superisostatic.

The key to the development of a non-mare mascon basin is that the time-frame of cooling and formation of a lithosphere following crater formation is shorter than the time-frame of isostatic

uplift. The rate of lithospheric deformation is controlled primarily by the cooling rate of the remnant thermal anomaly after impact. The larger the impact energy, the hotter and broader the remnant thermal anomaly and associated melt pool, which will increase the cooling time. It has been shown that the amount of impact melt scales as the square of the impact velocity or the impactor mass (Pierazzo et al. 1997). The rate these thermal anomalies cool conductively depends on the thermal diffusivity. Additionally, convection within the melt pool can decrease the cooling rate and the time to form a strong lithosphere.

The time frame of isostatic uplift is controlled by the viscosity of the mantle and the diameter of the basin. A lower viscosity mantle will result in a faster isostatic response and vice-versa (Melosh, 2011). In addition, isostatic forces acting on an undercompensated ring of thicker crust surrounding the basin center need to exceed lithospheric resistance to raise the inner basin above isostatic equilibrium. This latter requirement explains why basins smaller than 200 km in diameter do not form mascons (E in Figure 1.1). In these cases, lithospheric strength immediately after impact is much stronger than the isostatic forces generated by the lesser degree of excavation. However, this does not explain why SPA is not a mascon. Here we test the idea that this may be due to some combination of higher thermal gradients at the time of impact, larger basin size, or differences in the distribution of post-crater collapse crustal thickness. We employ the same numerical approach of modeling the entire evolution of an impact basin from initial impact through transient crater collapse, then cooling and isostatic adjustment (Melosh et al., 2013; Freed et al., 2014) to understand how SPA achieved its modern configuration as constrained by topography and gravity, and explore why it did not develop into a mascon basin.

1.3 South Pole-Aitken

South Pole-Aitken (B in Figure 1.1) is the largest (~2300 km, calculated average diameter from outer topography fit; Garrick-Bethell & Zuber, 2009) confirmed impact structure on the Moon and underlies all other basins, making it the oldest. It is likely 100s of millions of years older than most other recognized basins (Fassett et al., 2012; Spudis et al., 1994; Wilhelms, 1987). Observations of average topography, gravity, and crustal thickness of SPA as a function of distance from its center (-53°, 191°E) are shown as black lines in Figure 1.3. This figure also shows the 1-sigma deviations from the mean as a function of distance from basin center (vertical lines) which arises from a combination of the basin's elliptical shape (semi-major axis of ~970 km, semi-minor axis of ~720 km, and a westward tilt of 18.8°; Garrick-Bethell & Zuber, 2009), uneven distribution of ejecta, and variations in pre-impact regional topography and crustal thickness. Measurements from the Lunar Orbiter Laser Altimeter (LOLA) instrument at a resolution of 16 pixels per degree show an approximate average elevation of -5.5 km for the centermost 500 km (in respect to a spherical Moon with a radius of 1737 km; Smith et al., 2010). The free-air gravity and Bouguer gravity anomalies are taken from the spherical harmonic expansions of GRAIL measurements to degree and order 660 (Lemoine et al., 2014, model GGGRX_1200A). The free-air gravity shows a slight free-air negative anomaly to around 1100 km from basin center where the highlands are associated with a slight free-air positive anomaly. The Bouguer gravity is associated with a large positive (~400 mGals) and falls off to zero at 1100 km. Crustal thickness values taken from an inversion of Bouguer gravity anomaly data (Wieczorek et al., 2013) show the northern section outside of SPA contains thicker crustal material (>50 km) compared to the southern section (~40 km) (Wieczorek et al., 2013); this difference has been attributed to the deposition of ejecta from an oblique SPA forming impact (Melosh et al., 2017). Additionally, there is no significant low-

wavelength thickened crust (i.e. crustal collar) normally associated with large lunar basins; other work (Potter et al., 2012a) has attributed this to the higher geothermal gradient at the time of impact (50 K/km) compared to younger smaller basins like Orientale (14 K/km; Johnson et al. 2016).

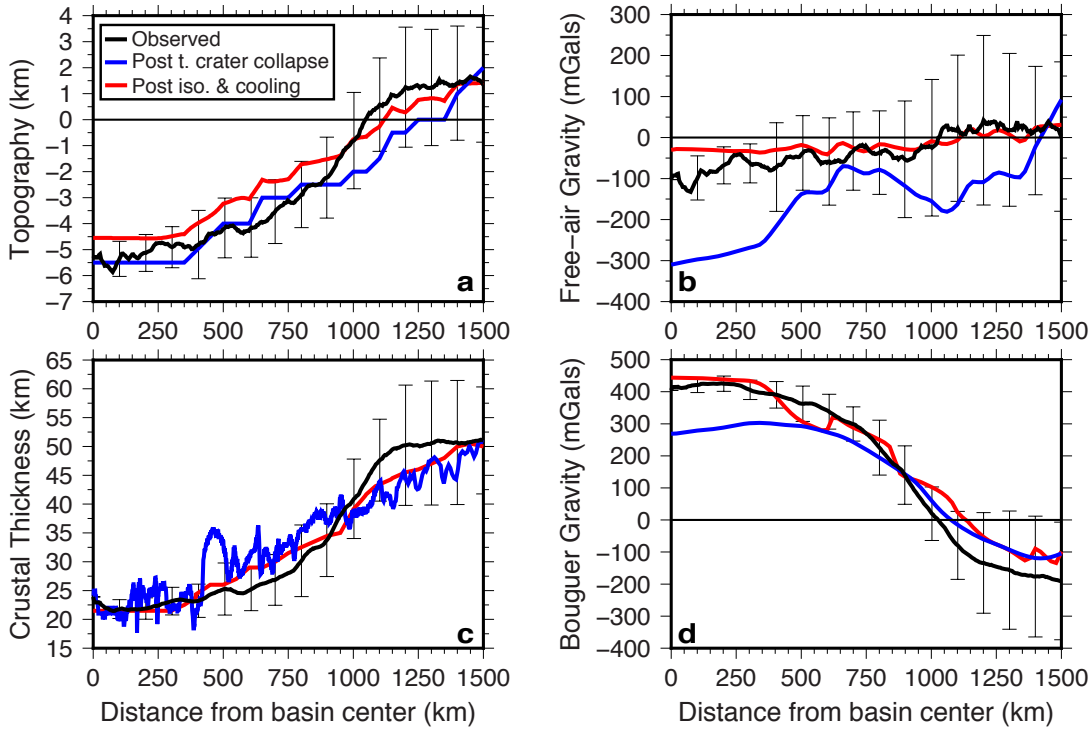

Figure 1.3. Observational constraints and comparison with best-fitting numerical model for SPA. The blue line in panel c shows the hydrocode results immediately following transient crater collapse for a 170-km-diameter projectile impacting at 10 km/s, assuming a 50 K/km geotherm, and a melt viscosity $<10^9$ PaS (see Figure 1.6). All other blue lines are the initial conditions for the finite element calculation. The red line in panel c shows the initial and final crustal thickness used in the finite element calculation. All other red lines show the results of the finite element calculation for the state of the basin after thermal cooling and isostatic adjustment. This assumes that the distance from basin center is radial distance for incorporation in our flat runs. Observed values are the mean and 1 standard deviation (1σ) variations with radial distance from elliptical averages of samples taken at 360 intervals around the basin center. Free-air and Bouguer gravity anomalies from a spherical harmonic expansion of the GRAIL-derived gravity field to degree and order 660 (Lemoine et al., 2014) and LOLA-derived topography (Smith et al., 2010). LOLA-derived topography at 16 pixels per degree (Smith et al., 2010). Crustal thickness calculated from GRAIL observations (Wieczorek et al., 2013, model 4).

Previous work (e.g. Potter et al., 2012a; Melosh et al., 2017) have modeled various aspects of the formation of SPA. Potter et al. (2012a) conducted 2D iSALE runs and determined that a 170

km diameter dunite projectile impacting at 10 km/s with a geothermal gradient of 50 K/km was the best fit to reproduce SPA's innermost outcrop of anorthosite. However, this study only modeled impact through transient crater collapse and did not attempt to match basin depth, inferred crustal thickness, nor gravity constraints that require consideration of cooling and isostatic adjustment. Melosh et al. (2017) used oblique ejecta constraints with a 3D version of iSALE to ascertain that a 200-km-diameter dunite object impacting at 45 degrees could reproduce the transient crater diameter suggested by Potter et al. (2012a) and explain the ejecta pattern. However, this study only considered the time-frame prior to transient crater collapse and did not consider the influence of a chemically distinct crust.

1.4 Modeling Approach

We numerically simulated the entire evolution of SPA from initial impact, through transient crater collapse, then cooling and isostatic adjustment, to arrive at its present-day form using the techniques developed by Melosh et al. (2013) and Freed et al. (2014). This was accomplished by using a hydrocode to simulate the initial impact through transient crater collapse (i.e. the first couple of hours) followed by a finite element code (FEM) to simulate cooling and isostatic adjustment (the next 10s to 100s of millions of years). The key to this sequential-code approach is to maintain self-consistency by applying the resulting geometry, thermal structure, and density structure at the end of the hydrocode simulation as initial conditions for the finite element model. Upon achieving steady state at completion of the FEM model, we compared the calculated topography, free-air and Bouguer gravity anomalies, and crustal thickness to modern observations to evaluate each model.

1.4.1 Hydrocode Modeling

We used the iSALE hydrocode (Amsden et al., 1980; Collins et al., 2004; Wünnemann et al., 2006) to simulate the process of crater excavation and transient crater collapse. Although SPA likely resulted from an oblique impact, we assumed axisymmetry due to the necessity for high resolution meshes to match the resolution of the topographic and gravity constraints. While this approach likely underestimates the size of the impactor (the 3-D models of Melosh et al., 2017 inferred a larger impactor than that of the axisymmetric models of Potter et al., 2012a), it is not likely to alter conclusions as to why SPA did not form as a mascon. Our models contained a high-resolution zone with 1-km cell sizes extending 1000 km from basin center and to a depth of 500 km. Outside this zone, the cells increased to a maximum size of 10 km. For an impact basin the size of SPA, the curvature of the Moon is influential. Thus, we treated the Moon as a spherical target with a central gravity field.

We use the similar model setup and the same material parameter as used by Johnson et al. (2018). The lunar crust and mantle were modeled assuming granite and dunite ANEOS equations of state, respectively. The difference between using an anorthosite and granite equation of state has a negligible effect on the post-transient crater configuration (Johnson et al., 2016). Additionally, the crustal strength parameters are fit to gabbro appropriate for the lunar crust (Potter et al., 2012a). To simulate the rheology of lunar crust and mantle, we incorporated a rock-like strength model (Collins et al., 2004), a damage model with an exponential dependence on plastic strain (Collins et al., 2004, Johnson et al., 2016), a dilatancy model (Collins, 2014), tensile failure model (Collins et al., 2004), and a thermal weakening model (Ohnaka, 1995) corresponding to the temperature- and pressure-dependence of the material. Additionally, we incorporated a visco-elastic-plastic rheology for the mantle (Elbeshausen and Melosh, 2018). To produce a realistic surface gravity, we incorporated an iron core into our lunar spherical target with a radius of 350

km (Johnson et al., 2016). The core was modeled using the iron ANEOS equation of state (Thompson, 1990) and the Johnson-Cook strength model with inputs for Armco Iron (Bowling et al., 2013; Johnson et al., 1983). Although the impactor was likely differentiated (Wieczorek et al., 2012; James et al, 2019), we use a homogenous dunite impactor rather than a differentiated one with a nickel-iron core to be consistent with Potter et al. (2012a). The use of a homogenous projectile overestimates the overall diameter of the impactor, but does not alter our best-fit hydrocode model as the kinetic energy of the impact rather than the mass distribution within the projectile controls the impact process (Melosh, 1989). The incorporation of core material within the mantle after impact will affect the gravity and isostatic adjustment of the FEM run (James et al., 2019), which we discuss in more detail in Section 1.4.2. Assumed hydrocode values and corresponding references are listed in Table 1.1.

Table 1.1. iSALE Parameters for SPA.

Description	Crust Values	Mantle Values	Core Values	References
Equation of State	ANEOS granite	ANEOS dunite	ANEOS iron	Benz et al. (1989); Pierazzo et al. (1997); Thompson (1990)
Melting temperature	1513 K	1373 K	1811 K	Davison et al. (2010); Potter et al. (2012a); Bowling et al. (2013); Johnson et al. (1983)
Thermal softening parameter	1.2	1.1	1.2	Davison et al. (2010); Potter et al. (2012a); Bowling et al. (2013); Johnson et al. (1983)
Simon A parameter	1840 MPa	1520 MPa	6000 MPa	Davison et al. (2010); Potter et al. (2012a); Bowling et al. (2013); Johnson et al. (1983)
Simon C parameter	7.27	4.05	3	Davison et al. (2010); Potter et al. (2012a); Bowling et al. (2013); Johnson et al. (1983)
Poisson's ratio	0.25	0.25	0.29	Davison et al. (2010); Potter et al. (2012a); Bowling et al. (2013); Johnson et al. (1983)
Frictional coefficient (damaged)	0.71	0.63	N/A	Davison et al. (2010); Potter et al. (2012a)
Frictional coefficient (undamaged)	1.1	1.58	N/A	Davison et al. (2010); Potter et al. (2012a)

Table 1.1. continued

Strength at infinite pressure	2.49 GPa	3.26 GPa	N/A	Davison et al. (2010); Potter et al. (2012a)
Cohesion (damaged)	0.01 MPa	0.01 MPa	N/A	Davison et al. (2010); Potter et al. (2012a)
Cohesion (undamaged)	31.9 MPa	5.07 MPa	100 MPa	Davison et al. (2010); Potter et al. (2012a); Bowling et al. (2013); Johnson et al. (1983)
Brittle ductile transition	1.23 GPa	1.23 GPa	N/A	(Collins et al., 2004)
Brittle plastic transition	2.35 GPa	2.35 GPa	N/A	(Collins et al., 2014)
Initial tensile strength	10 MPa	10 MPa	10 MPa	(Collins et al., 2014)
Maximum distension	1.2	1.2	N/A	(Collins et al., 2014)
Maximum dilatancy coefficient	0.045	0.045	N/A	(Collins et al., 2014)
Dilatancy pressure limit	200 MPa	200 MPa	N/A	(Collins et al., 2014)
Frictional coefficient (maximum distension)	0.4	0.4	N/A	(Collins et al., 2014)
Minimum Pressure	N/A	N/A	-2440 MPa	Bowling et al. (2013); Johnson et al. (1983)
Johnson-Cook A parameter	N/A	N/A	101 MPa	Bowling et al. (2013); Johnson et al. (1983)
Johnson-Cook B parameter	N/A	N/A	219 MPa	Bowling et al. (2013); Johnson et al. (1983)
Johnson-Cook N parameter	N/A	N/A	0.32	Bowling et al. (2013); Johnson et al. (1983)
Johnson-Cook C parameter	N/A	N/A	0	Bowling et al. (2013); Johnson et al. (1983)
Johnson-Cook M parameter	N/A	N/A	0.55	Bowling et al. (2013); Johnson et al. (1983)
Johnson-Cook reference temperature	N/A	N/A	2.5	Bowling et al. (2013); Johnson et al. (1983)

The zero-pressure melting temperature for dunite was chosen to be 1373 K, the approximate peridotite solidus; however, the material between the liquidus and solidus temperatures has some shear strength as super-solidus material will contain a slurry of melt, hot and cold clasts with the clasts providing some resistance to shear (Stewart, 2011). To address this discrepancy in material strength, we incorporated a melt viscosity as an input variable (Potter et al., 2012b; Potter et al., 2015), which gives the dunite some shear resistance above the melting temperature. Using the peridotite solidus as the melting temperature for dunite is suitable if the thermal profiles follow an adiabatic gradient (0.05 K/km) at subsolidus temperatures more than 1400 K (Freed et al., 2014); thereby, suppressing the presence of abundant supersolidus material.

We explored a range of model parameter space consistent with plausible uncertainties in several parameters. We varied the assumed impactor diameter (150-220 km; Potter et al., 2012a) and impact velocity (10-15 km/s; Le Feuvre and Wieczorek, 2008, 2011). We varied the average pre-impact crustal thickness from 40-60 km, based on uncertainties in estimated lunar crustal thicknesses (Wieczorek et al., 2013). Due to the uncertainty in the early lunar thermal structure, we varied the geothermal gradient from 30-50 K/km (Solomon and Head, 1980; Crosby and McKenzie, 2005). Additional uncertainties in the rheologic structure of the lunar mantle led us to vary the temperatures at which our thermal gradient transitioned to an adiabatic gradient between 1300 K and 1400 K, and the melt viscosity between 0 and 10^{10} Pa s.

A typical iSALE run begins immediately before contact between the target and projectile. It models the subsequent few hours following impact, encompassing the ejections of material and formation and collapse of the transient crater, incorporating several cycles of central uplift and collapse. For large lunar basins, the crust can migrate back in towards the basin center during

transient crater collapses (e.g. Freed et al., 2014; Johnson et al., 2018). We run our models until this crustal migration stops and topography reaches a stable, steady state condition. Once this state is reached, the crustal thickness, topography, and thermal and density structures of iSALE are output as initial conditions for the finite element models.

1.4.2 Finite Element Modeling

For each iSALE model run close to our best fit, an associated FEM model was developed to continue the calculation following transient crater collapse with a self-consistent thermal, density, and geometric structure. The output from iSALE was simplified in a manner that characterized the basic thermal, mechanical, and density structure. We varied these inputs into the FEM to characterize the uncertainty in that process as well as the iSALE results themselves. To model the evolution of SPA through cooling and isostatic adjustment, we used the finite element code Abaqus. Consistent with the hydrocode models, our FEM models assume axisymmetry. Since SPA spans ~80 degrees of lunar arc and thus may be influenced by spherical curvature, we developed and compared flat and spherical geometry models. As with the hydrocode mesh, the FEM mesh utilizes a high-resolution zone containing 1 km size elements that increase to a maximum of 10 km. Resolution tests showed that a more detailed mesh did not substantially influence model results. The model extends a lateral distance of 5500 km and to a depth of 1400 km, with the far-field boundaries fixed. Tests with alternate boundaries verify that these distances are sufficient to be beyond the influence of the impact scenarios we consider. The mesh is built with predominately linear quadrilateral elements with some triangular elements necessary to conform to the geometry of the basin. We assumed FEM thermal and elastic parameters for typical lunar crust and mantle (Melosh, 2011): Young's modulus, 100 GPa; Poisson's ratio, 0.25; thermal conductivity, 2.5 W/(m·K). These values were held fixed throughout each run.

The FEM modeling consists of four stages as outlined in Freed et al. (2014; refer to this paper for details not present here). The first stage reproduces the pre-impact thermal gradient assumed in iSALE. The second step introduces the iSALE thermal structure following crater collapse into the model and performs a conductive cooling calculation until the original background thermal structure is achieved. This serves to control the temperature-dependent rheology and density structures as a function of time in the viscoelastic analysis step. To consider the effects of thermal convection on melt pool cooling rates, we varied the thermal properties of the modeled materials; however, these variations did not have a significant effect on the final basin topography or gravity signatures. The third step introduces the starting density structure to the model, consistent with that from the iSALE model following transient crater collapse. This also enables us to calculate and apply an initial prestress state based on overburden pressure, which is required to prevent self-compression of the mesh due to gravity at the beginning of the viscoelastic analysis step. Since initial and final temperature structures are known, so are the initial and final temperature-dependent density structures. This information is used to assign an appropriate volumetric coefficient of thermal expansion to ensure that the final density structure is achieved at the completion of the viscoelastic step.

The final step is a viscoelastic run in which pressure gradients associated with the initial stress state drive viscoelastic flow while the viscoelastic rheology and density structures are modified by the cooling thermal structure. Cooling will cause thermal contraction and the growth of a lithosphere with time within the previously hot and weak inner basin. The main unknown variable in the viscoelastic run is the viscosity structure and its dependence on temperature. The viscosity structure strongly influences the evolution of the basin during cooling because it controls the rate of isostatic adjustment relative to the rate of lithospheric development (i.e. growth of

lithospheric thickness). To be consistent with Freed et al. (2014), we used a non-Newtonian approximation with an assumed uniform strain rate field derived from the temperature-dependent creep data of dry Maryland diabase,

$$\eta = \dot{\varepsilon}^{(1-\mathrm{n})/\mathrm{n}} e^{Q/(\mathrm{n}RT)} / \left(2A^{1/\mathrm{n}}\right) \quad (1)$$

where η is the viscosity, $\dot{\varepsilon}$ is the strain rate, n is the strain (or stress) exponent, Q is an activation energy, R is the universal gas constant, T is absolute temperature, and A is an experimental constant. For dry Maryland diabase, A = 8 $MPa^{4.7}$ s^{-1}, Q = 485 kJ/mol, and n = 4.7 (Mackwell et al., 1998). Temperature-dependent viscosity structures for four assumed strain rates (10^{-24}, 10^{-20}, 10^{-19}, 10^{-18} s^{-1}; rheologies 1–4, respectively) are shown in Figure 1.4. The crust is sufficiently cool such that its effective viscosity is high enough to preclude significant relaxation during cooling and isostatic adjustment. Thus, for simplicity, we used the same temperature-dependent viscosity for both the crust and mantle. The viscoelastic model is run until the basin topography and thermal structure reaches steady state, at which point we calculate the gravity field for the FEM using the equation for gravitational acceleration over a ring from Turtle and Melosh (1997). We then build and calculate the gravity field for an FEM run with the crustal thickness and density structure that corresponds to the Moon before impact. Comparing the gravity signatures between this undisturbed (geoid) FEM and the steady state output of our post-isostatic/conductive cooling model allows us to calculate the free-air and Bouguer gravity anomalies for comparison to the observed constraints.

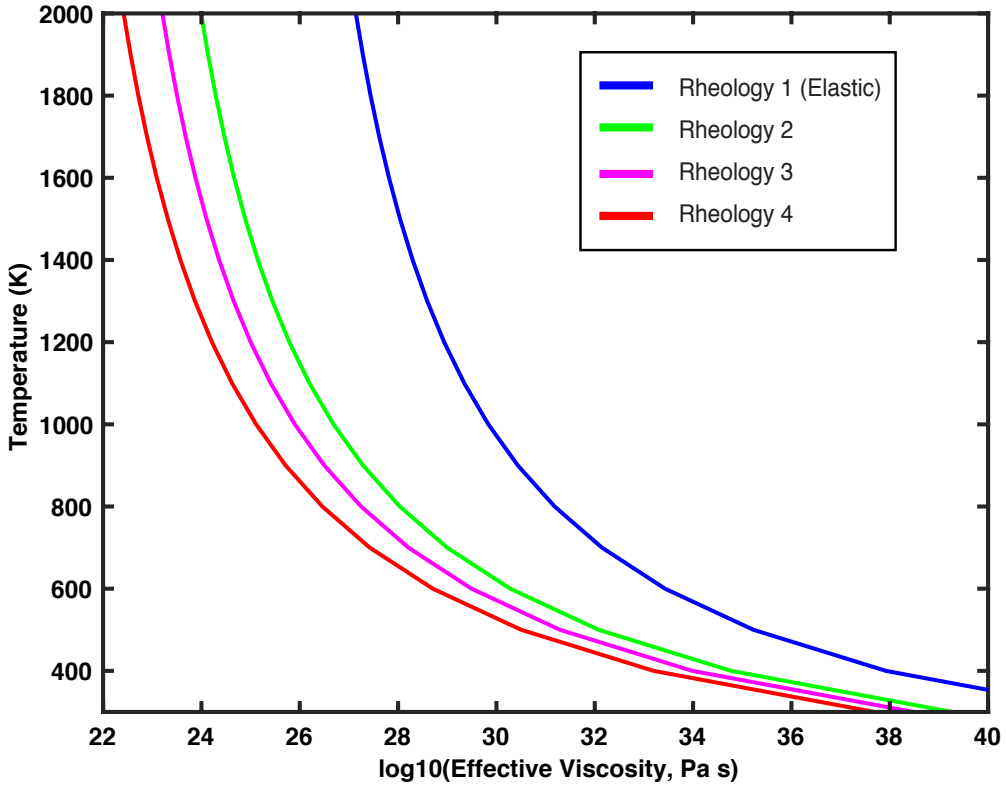

Figure 1.4. Temperature-dependent viscosity models used in the finite element analysis. Temperature-dependent viscosities based on experimentally determined values for dry Maryland diabase (Mackwell et al., 1998). Rheology 1 has such a high viscosity value it is essentially elastic.

1.5 Results of Hydrocode Modeling

Time frames from our best-fit hydrocode calculated evolution of SPA are shown in Figure 1.5. After ~10 min (Figure 1.5b) following impact, the transient cavity (800-1000 km in diameter) forms as the basin floor begins to rise. Approximately 20 min after transient crater formation this central uplift begins to collapse. The collapse of the central uplift (composed predominately of mantle material) generates the outward flow of material over the crust (currently ~400 km from basin center; Figure 1.5c). The collapse of the central uplift continues resulting in a central cavity (~50 min after impact). The cavity is subsequently filled by the inward flow of mantle material generating a secondary central uplift (~62.5 min after impact). The rise of this secondary uplift

pulls the crust inward, generating intense boudinage of the crust (Figure 1.5d). The rise and fall of subsequent central uplift continue to migrate the crust inward until it reaches a steady state at ~285 minutes from impact (Figure 1.5d). The inner basin crust settles at a depth of 5 to 6 km, close to modern observed topography; however, >1000 km from basin center the topography is about 2 km under observed values (Figure 1.5d). Furthermore, the free-air and bouguer gravity signatures are below observed values for the entire basin indicating some post-transient crater isostatic adjustment is needed (Figure 1.3b and 1.3d; see blue lines).

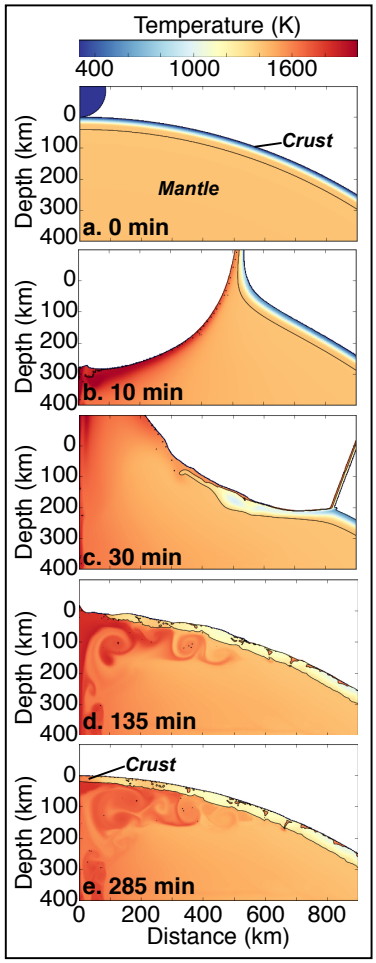

Figure 1.5. Calculated iSALE temperatures at five time-slices during the evolution of the South Pole-Aitken basin based on a 170-km-diameter projectile impacting at 10 km/s, assuming a 50 K/km geotherm. Black line shows boundary between crust and mantle.

The diameter of the SPA basin and the variation of crustal thickness with distance from the basin center are set at the end of transient crater collapse. They are not influenced by subsequent cooling and isostatic adjustment. These observational constraints are best matched by a projectile of about 170 km in diameter striking at 10 km/s, with a pre-impact lunar thermal gradient of 50 K/km (until a depth of 28 km at which an adiabat is reached), and pre-impact crustal thickness values of 40 km, similar conditions to the best-fit of Potter et al., 2012a.

Figure 1.6 illustrates the variations in crustal thickness with varying impact parameters. Increasing the size of the impactor from 170 km in diameter without changing the impact velocity

results in a larger basin diameter and thinner crust than observed; smaller impactors yield the opposite result (Figure 1.6a). Models in which the pre-impact crustal thickness is assumed to be greater or less than 40 km leads to the crust at the basin center to be either too thick or too thin, respectively (Figure 1.6b.) As the geothermal gradient is lowered below 50 K/km, a short-wavelength crustal bulge at approximately 600 km from the basin center begins to form (Figure 1.6c). This large crustal collar is a clear deviation from observations. Similarly, if we decrease the temperature our geothermal gradient rolls-over to an adiabatic gradient from 1400 K to 1300 K, another large crustal bulge forms between 500-1250 km (Figure 1.6d). Notice, though, that regardless of the roll-over temperature crust migrates into the basin center.

Potter et al. (2012a) concluded that SPA should have zero crustal thickness at the basin center following transient crater collapse. They appealed to differentiation of a melt pool to explain the currently observed crustal thickness of ~21.5 km at the basin center (Figure 1.3c; model 4 from Wieczorek et al., 2013). Our model can explain the observed crustal thickness at the center of SPA without appealing to differentiation, instead as resulting from the migration of crustal material back to the basin center during transient crater collapse (Figure 1.5). The difference in our respective models is the assumed melt viscosity (Figure 1.6e). Potter et al. (2012a) assumed a viscosity of 10^{10} Pa s while our best fitting model assumes the melt has a negligible viscosity ($<10^{9}$ Pa s). The reduction in melt viscosity reduced the drag at the base of the crust during transient crater collapse and allowed for more crustal migration (Figure 1.6e). Similar basin center crustal migrations have been simulated in the evolution of large lunar basins using this lower melt viscosity (e.g. Freed et al., 2014; Johnson et al., 2018).

Previous work (Vaughan and Head, 2014) showed differentiation can produce the crust within the center of SPA. Petrological modeling of SPA's melt sheet (Hurwitz and Kring, 2014)

noted that the Al content of the initial melt needs to be sufficient to generate the shallow noritic materials from the differentiation of the impact melt sheet. They noted that sources for the Al in the impact melt could reside in the crust or a pre-lunar overturn upper mantle. Since previous impact models completely removed the crust during impact, they concluded that SPA must have formed before the lunar mantle overturn. Observations can match crustal formation from differentiation for only a post lunar overturn case (Uemoto et al. 2017). This might suggest that the inward flow of crust during the impact process might actually be necessary for SPA to form additional crust after transient crater collapse, via melt differentiation. In this case, either our best-fitting hydrocode run would change to a cooler/stronger early Moon (see Figure 1.6) or the additional crust formed by melt differentiation is not substantial.

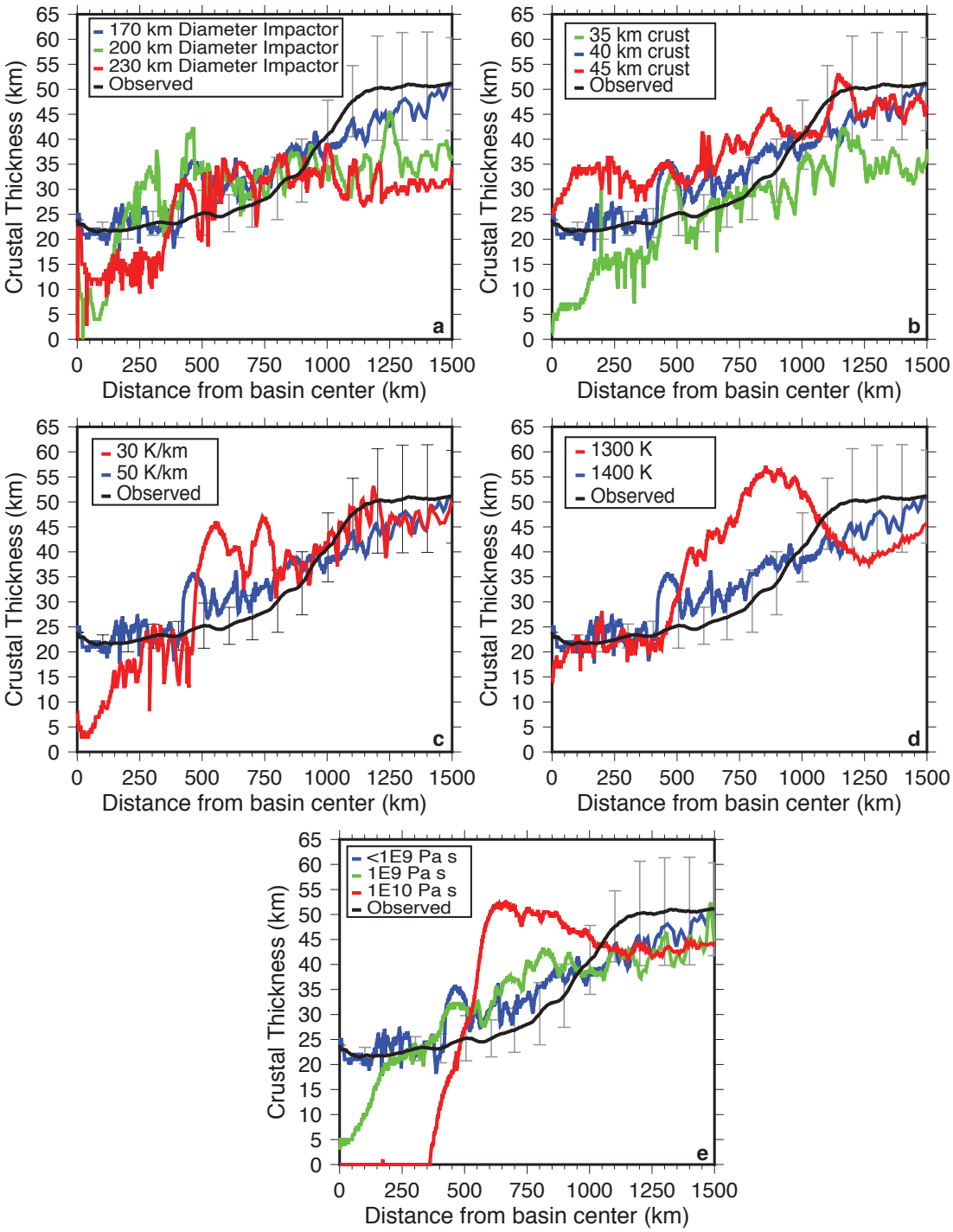

Figure 1.6. iSALE results with various impact parameters: (a) impactor diameter; (b) pre-impact crustal thickness; (c) geothermal gradient; (d) temperatures the thermal gradient rolls-over to adiabatic gradient; and (e) melt viscosity. Thin lines donate exact iSALE output while solid lines are the running average taken over increments of 200. Observed values are the mean and 1 standard deviation (1σ) variations with radial distance from elliptical averages of samples taken at 360 intervals around the basin center.

1.6 Results of Finite Element Modeling

For our best-fitting iSALE run (blue line in Figure 1.6), the FEM calculated the evolution of cooling following the collapse of the transient crater at several timesteps is shown in Figure 1.7. The crust cools relatively quickly allowing the lithosphere to reach its maximum thickness of 0-15 km (depending on the rheology model chosen, see Figure 1.4) within ~15 million years. The lithosphere within our model is defined as the depth at which the viscosity equals 1×10^{27} Pa·s; i.e. the value the material will not flow over the 300 million year run time of the models (see dotted line in Figure 1.7), the time required for the mantle to cool to the pre-impact background temperatures. Upon reaching final steady-state temperatures and basin morphology, the basin topography associated with each of the candidate rheology models (Figure 1.4) was compared to the observed topography of SPA (Figure 1.8).

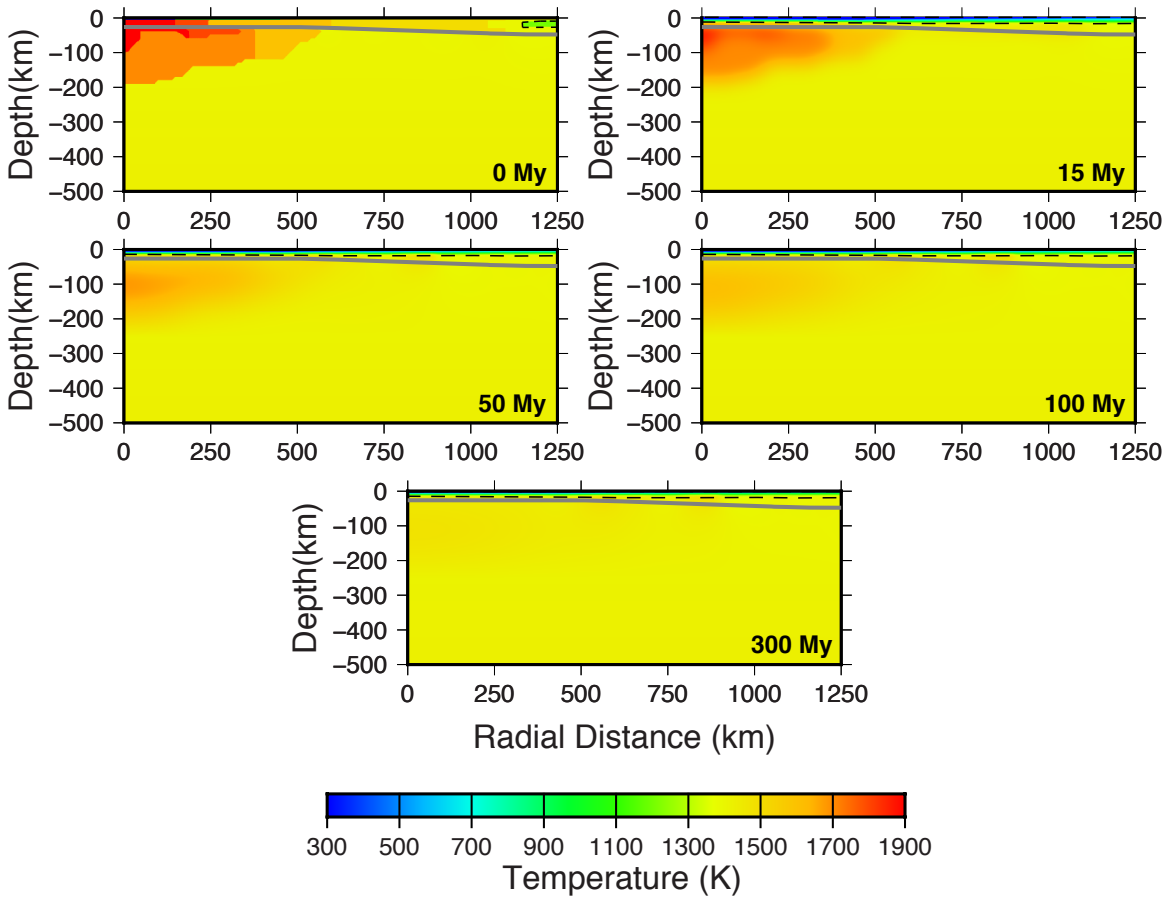

Figure 1.7. Conductive cooling time series for mantle thermal anomaly after impact. Grey line illustrates crust/mantle boundary. Black dotted line illustrates lithospheric thickness for rheology 2, see Figure 1.4. The lithosphere within our model is defined as the depth at which the viscosity equals $1x10^{27}$ Pa•s; i.e. the value the material will not flow over the run time of the models.

The highest viscosity model (rheology 1), which is essentially an elastic crust and mantle, results in topography that is ~1 km deeper than the observed topography throughout the entire inner basin (burgundy line in Figure 1.8). Without mantle flow responding to isostatic forces, the only mechanism from which deformation following transient crater collapse can occur is cooling and thermal contraction.

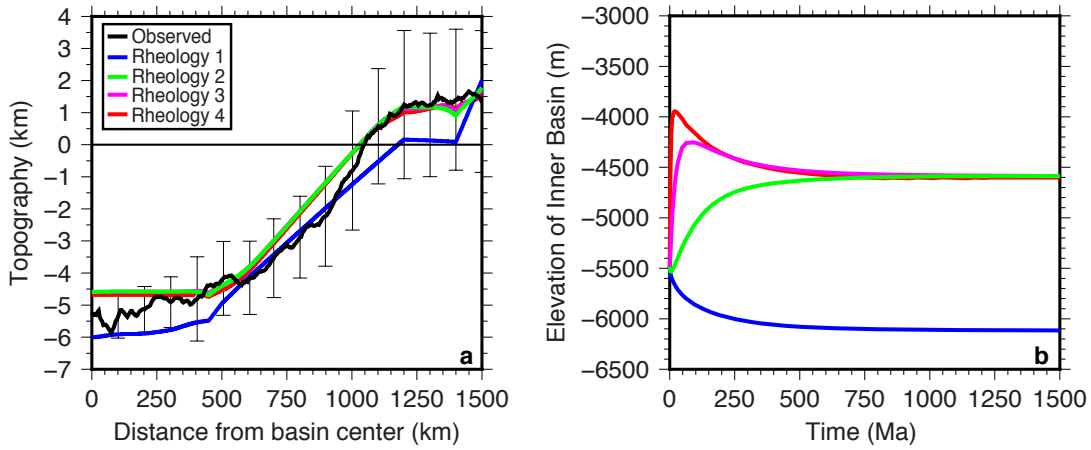

Figure 1.8. Comparison of Abaqus results for various temperature-dependent rheologies shown in Figure 1.4: (a) topography; (b) and elevation of inner basin due to isostatic adjustment with time.

The best fitting models (rheologies 2-4 in Figure 1.4) had a negligible difference between their final steady-state profiles; however, the displacements with time varied due the differences in viscosity (see Figure 1.8b). Rheology 3 (purple line in Figure 1.8) has a low enough mantle viscosity to allow the isostatic response to occur before the thermal anomaly cools completely, resulting in the basin reaching its highest topography at ~50 Ma, and slowly descends afterwards. Once the thermal anomaly completely cools, the inner basin reaches steady state. In contrast, the rheology 2 run (green line in Figure 1.8) has a mantle viscosity chosen so the thermal anomaly will begin to cool before the isostatic response. In this case, the inner basin starts to subside initially with thermal cooling. Once the isostatic response begins to occur, the basin rises to steady state. At steady state, both basins have the same final topography and gravity signatures indicating that the viscosity of the mantle does not play a significant role in the final configuration of SPA.

In addition to varying the viscosity structure, we investigated changing the starting topography from -5.5 km to -6.5 km, the uncertainty ranges from the best-fitting hydrocode models,

keeping all other variables constant. The final configuration of the basin was independent of the starting topography.

To account for how convective heat transfer (not incorporated within our modeling) affects the evolution of SPA, we varied the conductivity of the material within the melt pool from 2.5 W/(m·K) to 250 W/(m·K) (see Figure 1.9). Due to modeling constraints, the conductivity values are held constant throughout the run. As the conductivity increased, the maximum topography of the inner basin reached during the run decreased. Regardless of the conductivity of the run, all the runs reached the same final configuration independent of the cooling rate of the melt pool. In the run with a conductivity of 250 W/(m·K), cooling of the melt pool occurred quicker than the isostatic response. For this case, the inner basin rose to a steady state without subsiding from a previously higher topography.

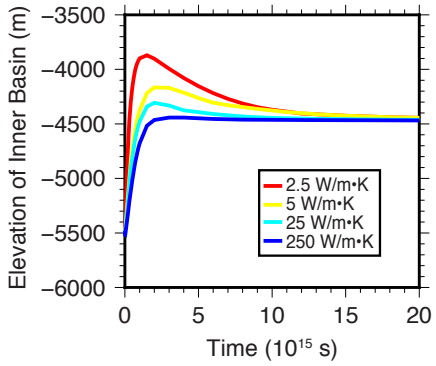

Figure 1.9. Elevation of the inner basin with time for various conductivity values of the melt pool.

Due to the uncertainty in the lunar density structure, we explored the effect of uniformly increasing the mantle density determined by iSALE's dunite equation of state by 80 kg/m^3. This increased mantle density, used in our best-fit model, resulted in a better match to observed topography for a majority of the basin with a negligible effect on the free-air and bouguer gravity signature. Increasing the crustal density for the inner basin would generate a similar outcome as increasing the mantle density; however, since the observed crustal thickness is based upon a bouguer gravity inversion that assumes a constant crustal density of 2550 kg/m^3 (Wieczorek et al., 2013), we kept the crust at a constant density of 2550 kg/m^3. Additionally, we varied the thickness

of the crust at distance: 40 km, 43 km, and 45 km. With increasing crustal thickness at distance, the entire basin settled to a lower final topography without changing the gravity anomalies. The 45 km crust produced the best-fit to the topography in the inner basin, but under-predicted topography between 1150-1500 km. The 40 km crust produced the worst fit, over-predicting much of the basin topography. The 43 km crust was chosen as our best-fit run given its close fit to the inner and outer basin topography.

In summary, our best fit models of SPA lead to a reasonable match with the observed topography, free-air and Bouguer gravity anomalies, and the inferred crustal thickness. Modest deviations from these observational constraints include an over-prediction of topography of ~0.5 km compared to the observed topography of the innermost basin (~ 200 km from basin center), and a corresponding ~70 mGal over-prediction of the free-air gravity anomaly (comparison of red and black lines in Figures 1.3a, 1.3b, and 1.3d, respectively). Our calculations suggest that from a starting position following transient crater collapse from the hydrocode runs, the final configuration of SPA after isostatic adjustment and thermal cooling is independent of the starting topography (within the uncertainty of the hydrocode runs), lunar curvature, mantle viscosity (as long as the crust is strong enough to not flow and the mantle is weak enough to flow within a few billion years), and cooling rate of the melt pool. The parameters that dictate the final topography of SPA post-transient crater collapse are the assumed thickness of the inner basin crust following transient crater collapse, the pre-impact crustal thickness, and the mantle and crustal densities.

The deviation between the modeled free-air gravity (red line in Figure 1.3b) and observed (black line in Figure 1.3b) could potentially be due to a mass anomaly at depth beneath the basin (James et al., 2019). The observed inner part of the basin is deeper than our models predict, which

is consistent with the James et al. (2019) conclusion that this inner depression is not formed by cooling and contraction, but is weighted down by a deep mantle mass anomaly.

1.7 The Different Evolution of Mid- vs Large-diameter Basins

Our result for the large SPA basin suggests that the evolution of this basin, and the parameters that influence this evolution, significantly differ from the evolution of mid-sized basins (Freed et al., 2014). These differences are primarily due to the differences in basin size and the magnitude and distribution of isostatic forces following transient crater collapse. Figure 1.2a and 1.2b shows the pressure differences (pressure outside basin minus pressure inside basin) following transient crater collapse for a mid-sized basin (e.g. Freundlich-Sharonov) and a large basin (e.g. SPA), respectively. These pressure differences mean that the isostatic forces of the mid-size basin are more narrowly focused compared to the large basin. This leads to isostatic forces being a much more dominant process compared to lithospheric rigidity during the evolution of large compared to mid-sized basins.

The relative influence of isostatic forces and lithospheric rigidity in the evolution of different sized basins can be understood by a general case study of how a lithospheric shell responds to a topography load. We will approximate a basin by assuming that topography of height h is removed from the shell. Assuming the shell is not infinitely rigid, the removal of this mass will cause an upward displacement of the shell (w). When a topographic load is applied to a spherical shell, the lithosphere can support the load via bending stresses and/or membrane stresses. We can write the ratio of the shell's displacement to the height of material removed as a function of these two methods of lithospheric support (Turcotte et al., 1981),

$$\frac{w}{h} = \frac{\rho_c}{\rho_m - \rho_c}\left\{1 - \frac{3\rho_m}{(2n+1)\bar{\rho}}\right\} \cdot \left\{\frac{\sigma\left[n^3(n+1)^3 - 4n^2(n+1)^2\right] + \tau[n(n+1) - 2] + n(n+1) - (1-\nu)}{n(n+1) - (1-\nu)} - \frac{3\rho_m}{(2n+1)\bar{\rho}}\right\}^{-1} \quad (2)$$

where ρ_c is crustal density, ρ_m is mantle density, $\bar{\rho}$ is the average density of the planet, ν is the Poisson's ratio, n is the degree of the spherical harmonic expansion of w and h (the lower the value of n corresponds with a longer wavelength topographic load). In equation 2, σ and τ represent the dimensionless parameters that measure the rigidity of the spherical shell to bending stresses and membrane stresses respectfully; these parameters are defined as follows,

$$\tau \equiv \frac{Ed}{R^2 \mathrm{g}(\rho_m - \rho_c)} \tag{3}$$

$$\sigma \equiv \frac{D}{R^4 \mathrm{g}(\rho_m - \rho_c)} \tag{4}$$

where E is the Young's modulus, R is the planet's radius, g is the surface gravity, d is the lithospheric thickness, and D is the flexural rigidity (defined $D = \frac{Ed^3}{12(1-\upsilon^2)}$). For short wavelength loads, $n(n+1) \gg 1$, equation 2 reduces to

$$\frac{w}{h} = \frac{\rho_c}{\rho_m - \rho_c}\left(\frac{1}{1+\sigma[n^2(n+1)^2]}\right) \tag{5}$$

where bending stresses dominate (Turcotte et al., 1981). For relatively small planetary bodies and large wavelength loads, it is appropriate to neglect bending stresses ($\sigma \to 0$), and equation 2 becomes

$$\frac{w}{h} = \frac{\rho_c}{\rho_m - \rho_c}\left\{1 - \frac{3\rho_m}{(2n+1)\bar{\rho}}\right\} \cdot \left\{1 - \frac{3\rho_m}{(2n+1)\bar{\rho}} + \tau\frac{n(n+1)-2}{n(n+1)-(1-\nu)}\right\}^{-1} \tag{6}$$

where membrane stresses dominate (Turcotte et al., 1981).

For the FEM models corresponding with our best-fit iSALE result (thermal lithospheric geotherm of 50 K/km, Figure 1.6c), there is a negligible difference between the models with a curved (with membrane stresses) and flat (without membrane stresses) geometries for a lithosphere that is 10-14 km thick (see rheologies 2-4 in Figure 1.4). From equation 3, the ability for a shell to support a topographic load is dependent on the lithospheric thickness, such that if $d \to 0$ then $\tau \to 0$, and equation 6 reduces to

$$\left(\frac{w}{h}\right)_0 = \frac{\rho_c}{\rho_m - \rho_c} \quad (7)$$

meaning the load is isostatically compensated. For SPA, during cooling the lithosphere never grows to a thickness sufficient to provide significant support of basin topography via membrane or bending stresses. This explains why the choice of viscoelastic rheology, starting topography, and cooling rate does not influence the final state of SPA — the final state is dominated by isostatic equilibrium.

In contrast to the isostatic dominated evolution of SPA during cooling and isostatic adjustment, the younger, mid-sized Freundlich-Sharonov basin took a different path in which isostatic forces and lithospheric rigidity (in terms of bending stresses) both played significant roles in its evolution. At only 200-km in crustal anulus diameter, and forming at a time when the Moon was cool enough to have a pre-impact lithosphere of the order of 40 km thick (Freed et al., 2014), *n* (isostatic forces) was of comparable magnitude to *D* (lithospheric rigidity) after only a few millions of years of cooling. This is why the choice of viscoelastic rheology, which controls the pace of mantle flow associated with isostatic forces, significantly influenced the evolution of the Freundlich-Sharonov basin.

The ability of the rapidly forming lithosphere during cooling of Freundlich-Sharonov enabled a mechanical bridge to form via bending stresses between the inner basin and the crustal collar (Figure 1.2a). This enabled the isostatically rising crustal collar to uplift the inner basin into a superisostatic configuration, leading to the formation of a mascon basin (Andrews-Hanna, 2013; Melosh et al., 2013; Freed et al., 2014). Not only did a strong mechanical bridge not form during cooling of SPA, but the long-wavelength crustal collar associated with SPA was much farther away from the basin center and experienced significantly lower isostatic forces (small pressure

difference compared to the outer basin as shown in Figure 1.2b). Thus, SPA did not evolve into a mascon basin.

The alternate evolutionary paths of mid-sized and large basins, and reason why the latter does not form a mascon, is illustrated in Figure 1.10, which shows the displacement history of the two basin centers. Isostatic uplift dominates the initial evolution of both basins. At ~50 Ma, as isostatic equilibrium of the large basin is approached, continued cooling and thermal contraction cause topography to decrease, resulting in the slightly negative free-air gravity anomaly of SPA observed today. In contrast, cooling of the mid-sized basin enables a lithosphere with significant bending rigidity to develop, arresting further changes in topography due to thermal contraction, and locking in the positive free-air gravity anomaly observed within Freundlich-Sharonov today.

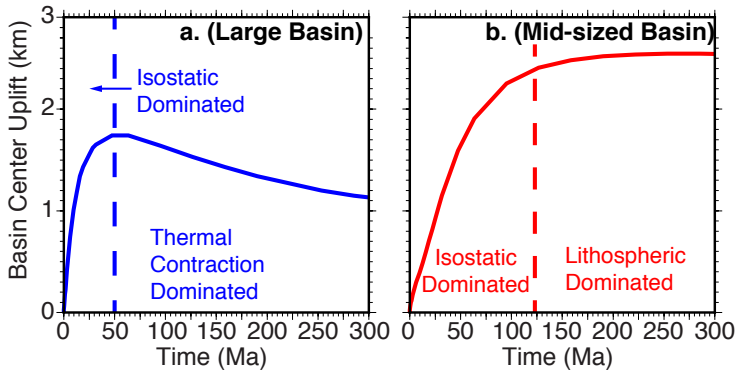

Figure 1.10. Time evolution of inner basin for a large basin (blue; e.g. SPA) and a mid-sized basin (red; e.g. Freundlich-Sharonov). Dashed lines indicate change in dominate forces for each basin size.

It is important to isolate the effect membrane stresses have on basin evolution to fully understand larger sized mascons and if SPA would have become a mascon if the impact occurred later in lunar history when the lithosphere grew in thickness. Figure 1.11 shows a comparison of runs conducted for a Freundlich-Sharonov sized basin for a curved geometry (thin lithosphere, 50 K/km, light blue line; and thick lithosphere, 30 K/km, dark blue line) and a flat geometry (thick lithosphere, 30 K/km, purple line). Comparing the flat run (purple line in Figure 1.11) with the curved run for the same lithospheric thickness (dark blue line in Figure 1.11) shows that membrane

stresses result in a lower topography and decrease the free-air gravity anomaly for the basin center. This is because membrane stresses depend much less on lithospheric thickness than bending stresses. With less dependence on lithospheric thickness, membrane stresses can affect the evolution of the basin much earlier than bending stresses. As stated above, basins start off subisostatic and rise up during relaxation. The membrane stresses act to resist this upward displacement resulting in a less positive free-air gravity anomaly. The evolution of the basin center for the curved run with a 50 K/km geothermal gradient (light blue line in Figure 1.11) is very similar to the evolution of SPA, shown in Figure 1.10. Without a strong lithosphere, the 50 K/km curved run is not able to sustain a superisostatic configuration and results in a negative free-air gravity anomaly; reconfirming that bending stresses are the key factor in lithospheric rigidity to form a mascon. For SPA models with a geothermal gradient of 30 K/km, the membrane stresses with the curved geometry begin to influence the evolution of the basin, resulting in a lower topography for the entire basin as well as a more negative free-air gravity anomaly. However, this model begins to deviate from the observed topography and gravity signature of the basin, indicating a less plausible lithospheric thickness. Thicker assumed lithospheres lead to a greater influence of membrane stresses in curved model results, but a poorer fit to the observational constraints. Thus, though membrane stresses can influence the evolution of the SPA basin, for the thin lithosphere inferred in our best-fitting models, the effect is negligible. This does not rule out the possibility SPA could have become a mascon due to infilling of volcanic deposits at a later time when the lithosphere grew in thickness and membrane stresses were strong enough to support a denser load, which might be important for similarly large basins on other planets that have been proposed to be free-air gravity positives (i.e. Sputnik Planitia on Pluto; Nimmo et al., 2016; Keane et al., 2016).

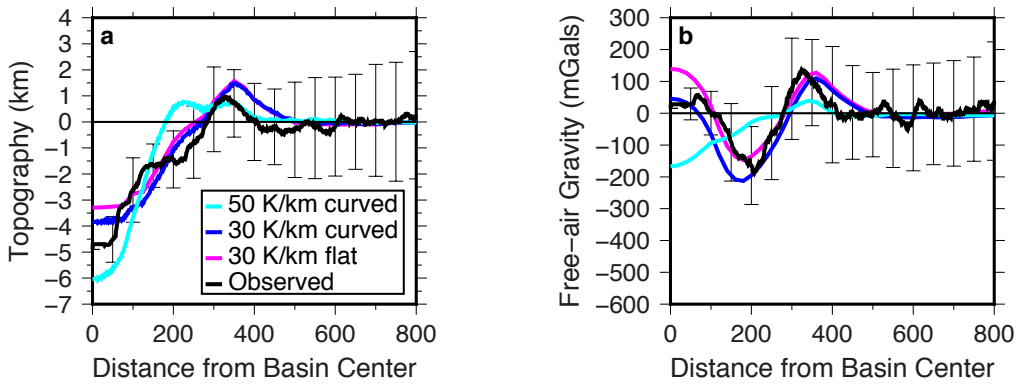

Figure 1.11. Topography (a) and Free-air gravity anomaly (b) results for post conductive cooling and isostatic adjustment of a Freudlich-Sharnov sized basin. The models consist of curved versus flat geometries for varying lithospheric thicknesses for the curved runs. Observed values are the mean and 1 standard deviation (1σ) variations with radial distance from azimuthal averages of samples taken at 360 intervals around the basin center. Free-air gravity anomalies from a spherical harmonic expansion of the GRAIL-derived gravity field to degree and order 660 (Lemoine et al., 2014) and LOLA-derived topography (Smith et al., 2010). LOLA-derived topography at 16 pixels per degree (Smith et al., 2010).

1.8 Conclusion

We used a sequence of hydrocode and finite element modeling constrained by topography and gravity observations to explore the complete evolution of the SPA basin, from impact through transient crater collapse, then through conductive cooling and isostatic adjustments. Our results suggest that the diameter of the basin is best explained by a projectile of 170 km in diameter striking at 10 km/s (with trade-offs between diameter and velocity, and remembering that this is for a vertical impact: Accounting for the obliquity would require a larger impactor), with a pre-impact lunar thermal gradient of 50 K/km (until a depth at which an adiabat is reached), and pre-impact crustal thickness of 40 km. The final crustal thickness of ~21.5 km at the basin center can be explained as being due to the migration of crust back to the basin center during transient crater collapse, based on an assumption of a negligible melt viscosity.

Following transient crater collapse, cooling and isostatic adjustment shallow the SPA basin to the topography observed today, as well as the observed free-air and Bouguer gravity anomalies.

Predominately due to the great size of the basin (long wavelength), this process was controlled by isostatic adjustment with minimal influence from lithospheric rigidity. Given the high geothermal gradient at impact, membrane stresses had a minimal effect on the final basin configuration. Thus, the choice of mantle viscosity had little influence on the final state of the SPA basin, though it did influence the pace of post-transient crater collapse deformation. This is in contrast to mid-sized basins (short wavelength of order 200 km in diameter), where lithospheric rigidity played a significant role in the evolution of the basin following transient crater collapse, resulting in the formation of mascon basins. Influenced primarily by isostatic forces, SPA could not evolve into a mascon basin.

1.9 References

Amsden, A., Ruppel, H., and Hirt, C. (1980). SALE: A simplified ALE computer program for fluid flow at all speeds. *Los Alamos National Laboratories Report*, LA-8095:101p. Los Alamos, New Mexico: LANL.

Andrews-Hannah (2013). The origins of the non-mare mascon gravity anomalies in lunar basins, *Icarus* **222**, 159-168.

Apollo 16 Mission Report (August, 1972). NASA, MR-11, Wash., D.C.,

Baldwin, R. B. (1968). Mascons: Another interpretation. *Science* **162**, 1407-1408.

Benz, W., Cameron, A. G., Melosh, H. J. (1989). The origin of the moon and the single-impact hypothesis III. *Icarus* **81**, 113–131 Medline doi:10.1016/0019-1035(89)90129-2

Bowling, T. J., Johnson, B. C., Melosh, H. J., Ivanov, B. A., O'Brien, D. P., Gaskell, R., Marchi, S. (2013). Antipodal terrains created by the Rheasilvia basin forming impact on asteroid 4 Vesta. J. *Geophys. Res. Planets* **118**, 1821–1834 10.1002/jgre.20123 doi:10.1002/jgre.20123

Collins, G. S. (2014). Numerical simulations of impact crater formation with dilatancy. *J. Geophys. Res. Planets* **119**, 2600–2619. doi:10.1002/2014JE004708

Collins, G. S., Melosh, H. J., and Ivanov, B. A. (2004), Modeling damage and deformation in impact simulations. *Meteoritics and Planetary Science* **39**, 217-231.

Crosby, A., & McKenzie, D. (2005), Measurements of the elastic thickness under ancient lunar terrain. *Icarus* **173**, 100-107.

Davison, T.M., Collins, G.S., Ciesla, F.J., (2010), Numerical modelling of heating in porous planetesimal collisions. *Icarus* **208**, 468–481.

Elbeshausen, D., and Melosh, J. (2018). A nonlinear and time-dependent visco-elasto-plastic rheology model for studying shock-physics phenomena. *arXiv preprint arXiv:1805.06453*.

Fassett, C. I., J. W. Head, S. J. Kadish, E. Mazarico, G. A. Neumann, D. E. Smith, and M. T. Zuber (2012), Lunar impact basins: Stratigraphy, sequence and ages from superposed impact crater populations measured from Lunar Orbiter Laser Altimeter (LOLA) data, *J. Geophys. Res.* **117**.

Freed, A. M., et al. (2014), The formation of lunar mascon basins from impact to contemporary form, *J. Geophys. Res. Planets* **119**.

Garrick-Bethell, I., & Zuber, M. T. (2009), Elliptical structure of the lunar South Pole-Aitken basin. *Icarus* **204**, 399-408.

Hurwitz, D. M., & Kring, D. A. (2014). Differentiation of the South Pole–Aitken basin impact melt sheet: Implications for lunar exploration. *Journal of Geophysical Research: Planets* **119**, 1110-1133.

Ivanov, B.A., D. Deniem, and G. Neukum (1997), Implementation of Dynamic Strength Models into 2D Hydrocodes: Applications for Atmospheric Breakup and Impact Cratering." *Inter. J. Impact Engr* **20**, 411-30.

James, P. B., D. E. Smith, P. K. Byrne, J. Kendall, H.J. Melosh, and M. T. Zuber. (2019), Deep Structure of the Lunar South Pole-Aitken Basin. *Geophysical Research Letters.* 10.1029/2019GL082252.

Johnson, B. C., Blair, D. M., Collins, G. S., Melosh, H. J., Freed, A. M., Taylor, G. J., ... & Keane, J. T. (2016), Formation of the Orientale lunar multiring basin. *Science* **354**, 441-444.

Johnson, B. C., Andrews-Hanna, J. C., Collins, G. S., Freed, A. M., Melosh, H. J., & Zuber, M. T. (2018). Controls on the formation of lunar multiring basins. *Journal of Geophysical Research: Planets* **123**, 3035-3050.

Johnson, G. R., Cook, W. H. (1983), A constitutive model and data for metals subjected to large strains, high strain rates and high temperatures. *7th International Symposium on Ballistics*; www.lajss.org/HistoricalArticles/A%20constitutive%20model%20and%20data%20for%20metals.pdf.

Keane, J. T., Matsuyama, I., Kamata, S., & Steckloff, J. K. (2016). Reorientation and faulting of Pluto due to volatile loading within Sputnik Planitia. *Nature* **540**, 90-93.

Le Feuvre, M., and M. A. Wieczorek (2008), Nonuniform cratering of the terrestrial planets, *Icarus* **197**, 291–306.

Lemoine, F., Goossens, S., Sabaka, T., Nicholas, J., Mazarico, E., Rowlands, D. et al. (2014). GRGM900C: A degree-900 lunar gravity model from GRAIL primary and extended mission data. *Geophysical Research Letters* **41**, 10.1002/2014GL060027.

Mackwell, S. J., M. E. Zimmerman, and D. L. Kohlstedt (1998). High-temperature deformation of dry diabase with application to tectonics on Venus. *J. Geophys. Res.* **103**, 975–984, doi:10.1029/97JB02671.

Melosh, H. J. (1989), *Impact cratering: A geologic process.* New York, Oxford University Press, p. 74.

Melosh, H. J. (2011), *Planetary surface processes* (Vol. 13). Cambridge University Press.

Melosh, H. J., et al. (2013), The Origin of Lunar Mascon Basins, *Science* **340**, 1552-1555.

Melosh, H. J., Kendall, J., Horgan, B., Johnson, B. C., Bowling, T., Lucey, P. G., & Taylor, G. J. (2017), South Pole–Aitken basin ejecta reveal the Moon's upper mantle. *Geology* **45**, 1063-1066.

Muller, P. M., & Sjogren, W. L. (1968), Mascons: Lunar mass concentrations. *Science* **161**, 680-684.

Neumann, G. A., M. T. Zuber, D. E. Smith, and F. G. Lemoine (1996), The lunar crust: Global structure and signature of major basins. *J. Geophys. Res*. **101**, 841–863.

Nimmo, F., Hamilton, D. P., McKinnon, W. B., Schenk, P. M., Binzel, R. P., Bierson, C. J., ... & Olkin, C. B. (2016). Reorientation of Sputnik Planitia implies a subsurface ocean on Pluto. *Nature* **540**, 94-96.

Ohnaka, M. (1995), A shear failure strength law of rock in the brittle-plastic transition regime. *Geophysical Research Letters* **22**, 25-28.

Pierazzo, E., Vickery, A. M., & Melosh, H. J. (1997), A reevaluation of impact melt production. *Icarus* **127**, 408-423.

Potter, R. W. K., G. S. Collins, W. S. Kiefer, P. J. McGovern, and D. A. Kring. (2012a), Constraining the size of the South Pole-Aitken basin impact. *Icarus* **220**, 730–743.

Potter, R. W. K., D. A. Kring, G. S. Collins, W. S. Kiefer, P. J. McGovern. (2012b). Estimating transient crater size using the crustal annular bulge: Insights from numerical modeling of lunar basin-scale impacts. *Geophys. Res. Lett.* **39**, L18203.

Potter, R. W. K., D. A. Kring, G. S. Collins (2015). Scaling of basin-sized impacts and the influence of target temperature, in *Large Meteorite Impacts and Planetary Evolution V: Geological Society of America Special Paper* **518**, G. R. Osinski, D. A. Kring, Eds., pp. 99–113.

Smith, D. E., et al. (2010). Initial observations from the Lunar Orbiter Laser Altimeter (LOLA). *Geophys. Res. Lett.* **37***,* L18204.

Solomon, S. C., & Head, J. W. (1980). Lunar mascon basins: Lava filling, tectonics, and evolution of the lithosphere. *Reviews of Geophysics* **18**, 107-141.

Spudis, P. D., Reisse, R. A., & Gillis, J. J. (1994). Ancient multiring basins on the Moon revealed by Clementine laser altimetry. *Science* **266**, 1848–1851.

Stewart, S. T. (2011). Impact basin formation: The mantle excavation paradox resolved, *Lunar Planet. Sci.* **42**, abstract #1633.

Thompson, S. L. (1990). ANEOS analytic equations of state for shock physics codes input manual. *Sandia Report SAND89-2951 UC-404*.

Turcotte, D. L., Willemann, R. J., Haxby, W. F., & Norberry, J. (1981). Role of membrane stresses in the support of planetary topography. *Journal of Geophysical Research: Solid Earth* **86**, 3951-3959.

Turtle, E., Melosh, H. (1997). Stress and flexural modeling of the Martian lithospheric response to Alba Patera. *Icarus* **126**, 197.

Uemoto, K., Ohtake, M., Haruyama, J., Matsunaga, T., Yamamoto, S., Nakamura, R., ... & Iwata, T. (2017). Evidence of impact melt sheet differentiation of the lunar South Pole-Aitken basin. *Journal of Geophysical Research: Planets* **122**, 1672-1686.

Vaughan, W. M., & Head, J. W. (2014). Impact melt differentiation in the South Pole-Aitken basin: Some observations and speculations. *Planetary and Space Science* **91**, 101-106.

Wieczorek, M. A., Weiss, B. P., & Stewart, S. T. (2012). An impactor origin for lunar magnetic anomalies. *Science* **335**, 1212-1215.

Wieczorek, M. A., et al. (2013), The crust of the Moon as seen by GRAIL. *Science* **339**, 671–675.

Wilhelms, D. E. (1987), *The Geologic History of the Moon*, USGS Prof. Pap. 1348.

Wünnemann, K., G.S. Collins, and H. J. Melosh (2006), A Strain-based Porosity Model for Use in Hydrocode Simulations of Impacts and Implications for Transient Crater Growth in Porous Targets. *Icarus* **180**, 514-27.

Zuber, M. T., et al. (2013), Gravity field of the Moon from the Gravity Recovery and Interior Laboratory (GRAIL) mission, *Science* **339**, 668–671.

CHAPTER 2. EXCAVATION OF LOWER CRUST AND MANTLE MATERIAL BY THE ISIDIS IMPACT, MARS

A version of this chapter will be submitted to the journal *Nature Geoscience*.

2.1 Introduction

Isidis Planitia (topography[1]: Figure 1a; crustal thickness[2]: Figure 1b) is a ~1900 km diameter impact basin centered at 13°N, 87.0°E and is thought to be the last major impact basin that formed on Mars, approximately 3.8-3.9 billion years ago[3,4]. The impact excavated material and deposited it in the northeast Syrtis region on which the 49-km Jezero crater subsequently formed. The NASA Mars 2020 rover, named Perseverance, will land in Jezero crater[5] (Figure 1a), ~620 km from the center of Isidis Planitia, to seek out signs of past life and collect samples for eventual return to Earth. These samples will offer a unique opportunity to study the composition and mineralogy of the martian interior because of the deep excavation by the Isidis impact. Here, we simulate the Isidis impact using the shock physics hydrocode iSALE to determine the provenance and shock state of ejecta within Isidis' basin rim. We find that the Isidis-forming impact is capable of excavating the Martian upper mantle and depositing it near Jezero. Spectral and High Resolution Imaging Science Experiment (HiRISE) imaging has observed megabreccia with a dominate orthopyroxene (Opx) signature 500-1000 km radially from the center of Isidis (see Figure 1c and 1d), which is consistent with the modeled excavated mantle material. We find that the mantle-derived ejecta in this region are predominately shocked between 50-80 GPa, which will enable the Perseverance rover to distinguish it from other materials. Additionally, we find that ejecta in this region also contains abundant unshocked (<4 GPa) crustal materials excavated via impact. Thus, the circum-Isidis region will be a critical area for in-situ and returned sample

investigations of basin-forming impact processes as well as determining the composition of the Martian interior[6].

A previous study[7] conducted a suite of impact simulations and admittance modeling of Isidis Planitia in which the geothermal gradient, pre-impact crustal thickness, and projectile diameter were treated as variable parameters. The authors concluded that the best-fit impact parameters for Isidis are a 100-120 km diameter dunite impactor striking a 60 km thick basalt crust with a 20 K/km geothermal gradient at 12 km/s; however, they did not constrain the provenance depth of ejected material. Additionally, their models[7] did not cap the thermal gradient within the crust at its melting temperature, resulting in a molten, strengthless pre-impact lower crust. In this work, we seek to improve the modeling of the excavation depth and evolution of the Isidis Planitia forming impact by conducting a parameter search of impactor parameters for impactor diameters ranging from 120 to 180 km as well as varying thermal profiles (see Section 2.5).

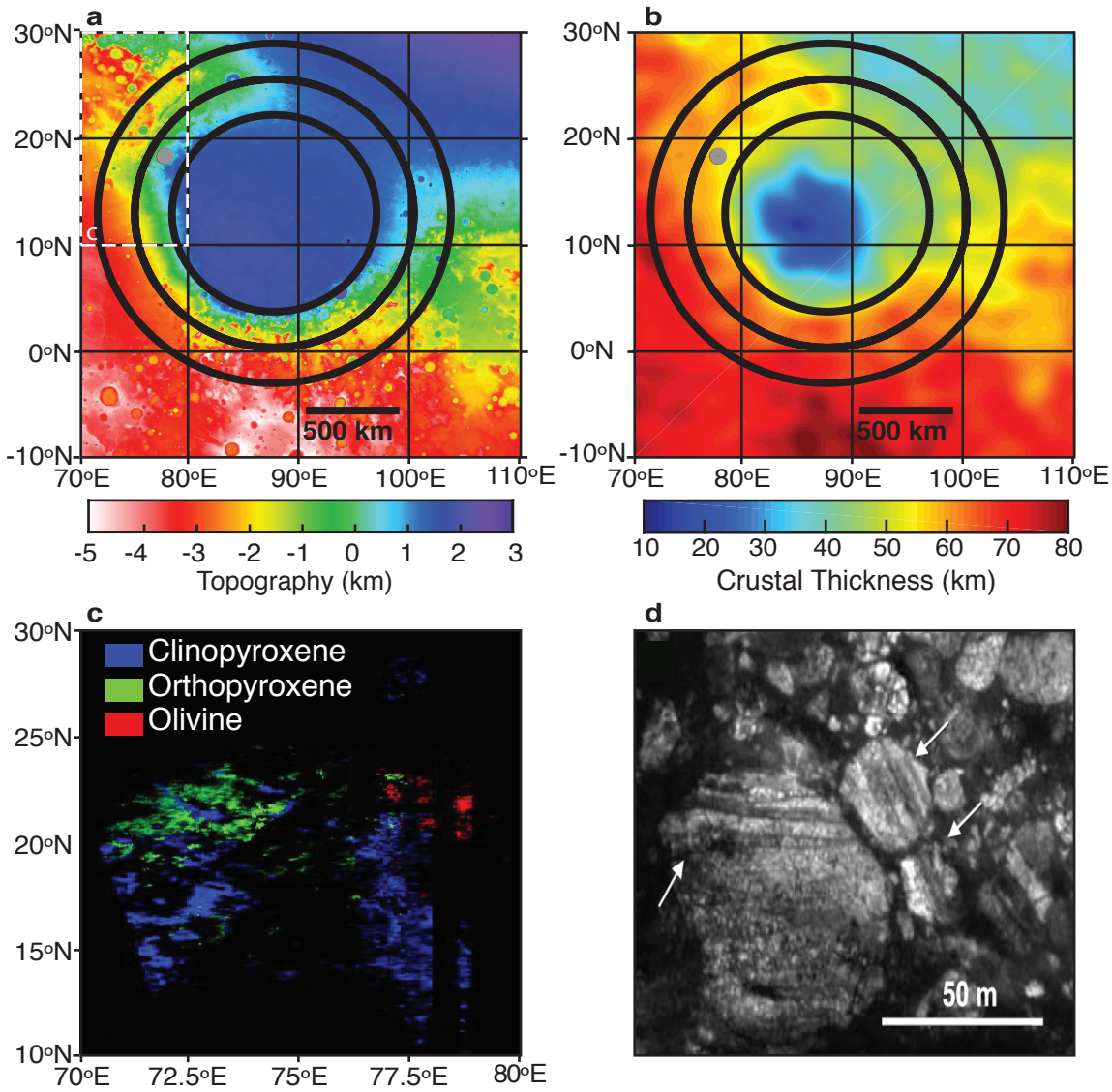

Figure 2.1. Topography (a); crustal thickness (b); mafic spectral signature map of Isidis Planitia taken from (ref. 23) (c); and HiRISE image of megabreccia blocks taken from (ref. 24) (d). The white dotted line in (a) outlines (c). Black rings label the inner (radius of ~550 km), middle (radius of ~750 km), and outer rim (radius of ~950 km) from (ref. 3). The gray circle shows the location of Jezero. Topography map of Mars at a resolution of 128 pixels per degree, based on altimetry data acquired by the Mars Global Surveyor MOLA instrument[1]. Crustal thickness map derived from spherical harmonic coefficients (from degree 2 up to degree 100) of the NASA GMM-3 gravity field with a resolution of 16 pixels per degree[2].

2.2 Excavation of Martian Upper Mantle Observed in Hydrocode Models

Figure 2.3 shows a comparison between our best-fit model (red line) and the observed azimuthally averaged crustal profile of Isidis Planitia (black line). With a 180 km diameter dunite

projectile impacting at 12 km/s into a 40 km thick average basalt crust with a geothermal gradient of 30 K/km, we reproduce the crustal profile of Isidis within the one sigma measurement uncertainties of the mean profile. Our best-fit impactor diameter is 60 km larger than previous estimates[7]. In addition, our work predicts a warmer thermal gradient and thinner average crust for earlier Mars. Our thinner preimpact crustal thickness of 40 km better matches the 45 km value predicted by gravity calculations of previous work[8]. Using a 45 km thick crust, this previous work calculated the gravity anomaly generated by the flexural loading of Isidis and matched the observed free-air gravity anomaly of Isidis Plainita. As for the thermal gradient, our geothermal gradient is within the 17-32 K/km range for the late Noachian-early Hesperian period determined from modeling thrust faults in the southern Thaumasia region[9]. Given the geothermal gradient predicted by our modeling is on the higher bound of this geothermal range might imply that Isidis Planitia is closer to the 3.9-billion-year-old age rather than the 3.8 billion.

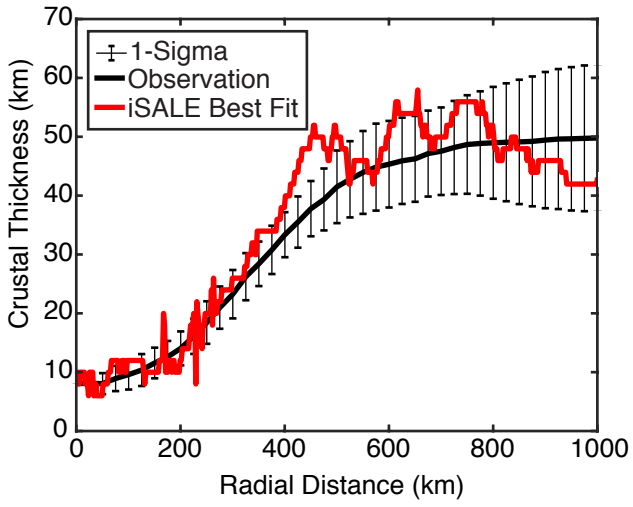

Figure 2.2. Comparison of observed crustal thickness with iSALE best-fit. The solid red line shows the result of a hydrocode simulation with a 180 km in diameter dunite projectile impacting at 12 km/s into a 40 km thick pre-impact crustal thickness with a geothermal gradient of 30 K/km and a melt viscosity of $2x10^9$ Pa s. The black solid line is the azimuthal averaged observed crustal thickness taken at 360 intervals around the basin center while the error bars represent one standard deviation (1σ) from that average. Observed crustal thickness anomalies derived from the inversion of the NASA GMM-3 Bouguer gravity field from degree 2 up to degree 100[2].

Figure 2.3 shows the evolution of the Lagrangian tracer particles over the course of a run for a 180 km impactor. At ~4 minutes after impact (Figure 2.3b), a transient crater ~350 km in

diameter forms. As the transient crater becomes unstable, it collapses in on itself resulting in a large central uplift (Figure 2.3c). At approximately this time (~15 minutes after impact), mantle material contained within the ejecta curtain is deposited at a radial distance of ~620 km (Figure 2.3c). The central peak rises until it becomes gravitationally unstable and collapses, resulting in both vertical and lateral flow of mantle and crustal material (Figure 2.3d). Over the next 7 minutes, this material is transported to distances greater than or equal to 620 km from the basin center. Material originating from the central uplift is deposited atop ejecta in the vicinity of Jezero as its subsequent collapse demarks the last significant dynamic step in the impact process. The physical processes responsible for this displacement are similar to those that have been suggested to form peak rings within impact craters of smaller diameters[10]. A comparison of the amount of mantle material deposited via the central peak collapse versus the ejecta curtain can be found in Section 2.7.

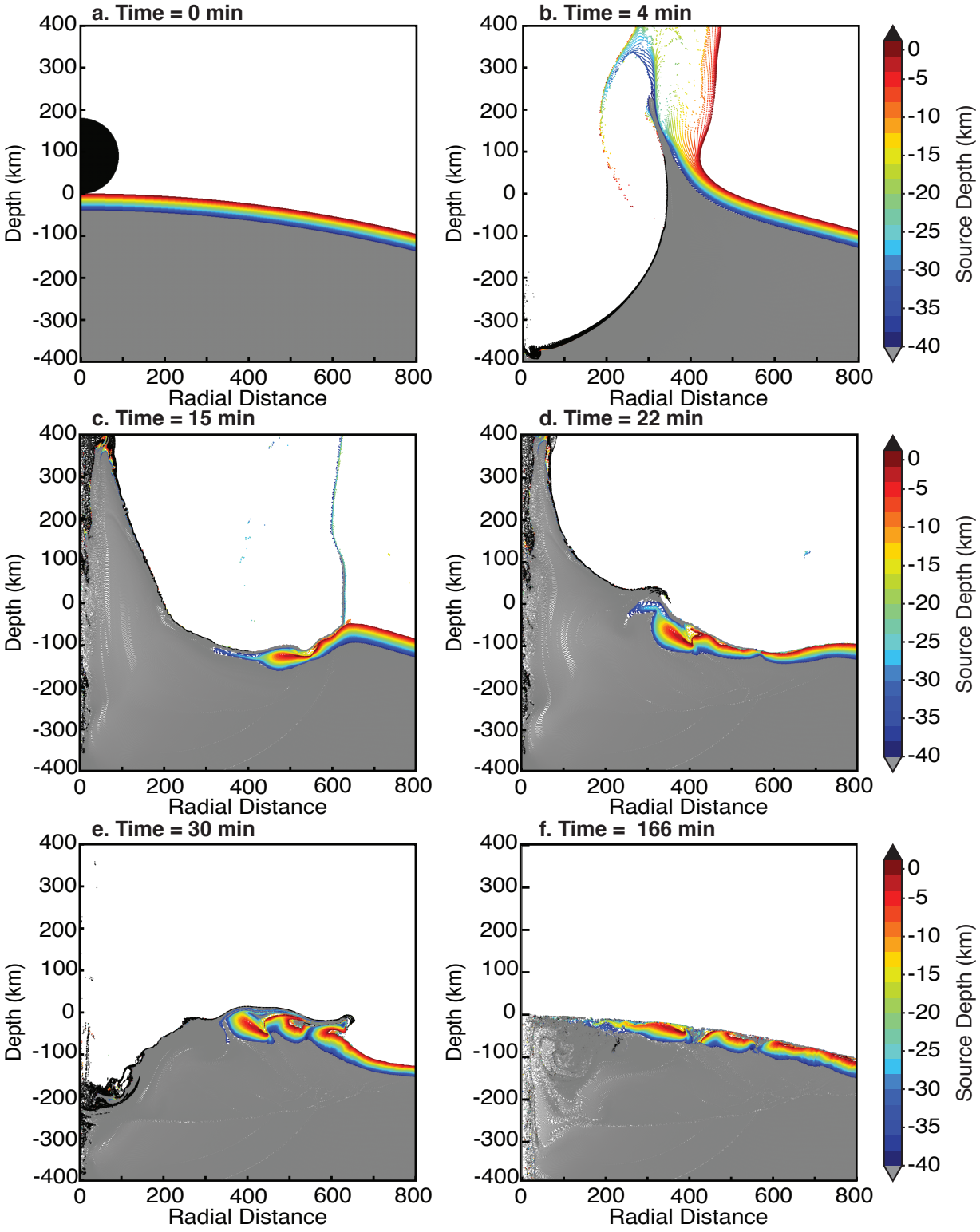

Figure 2.3. iSALE lagrangian tracer particle results for a 180 km diameter dunite impactor with an impact velocity of 12 km/s, a geothermal gradient of 30 K/km, and a pre-impact crustal thickness of 40 km. Color indicates crust while gray indicates mantle material. Tracers from >40 km depth are plotted on top of other tracers to emphasize mantle material.

For our hydrocode run that best matches the crustal thickness of Isidis, the top 15 km of ejecta at the location of Jezero crater consists of a mixture of material with variable depth of provenance: 1% remnant projectile material, 52% upper crustal material (0-20 km depth), 27% lower crustal material (20-40 km depth), and 20% mantle material (>40 km depth) (see Figure 2.4). Figure 4 shows a histogram plot for a variety of impactor diameters, including our best-fit of 180 km (Figure 2.4d). For every run, upper mantle material (blue in Figure 2.4) is deposited at the radial distance of Jezero crater (outlined in black dotted lines) with the larger impactors depositing mantle material over greater radial distance. This is consistent with crater scaling laws that suggest a maximum excavation depth of 60–100 km for a basin the size of Isidis[8]. Given that mantle material is deposited at Jezero regardless of the projectile diameter range (120-180 km), we expect mantle material to be excavated during the Isidis forming impact regardless of the discrepancy between ours and previous work's[7] best-fit impactor diameter.

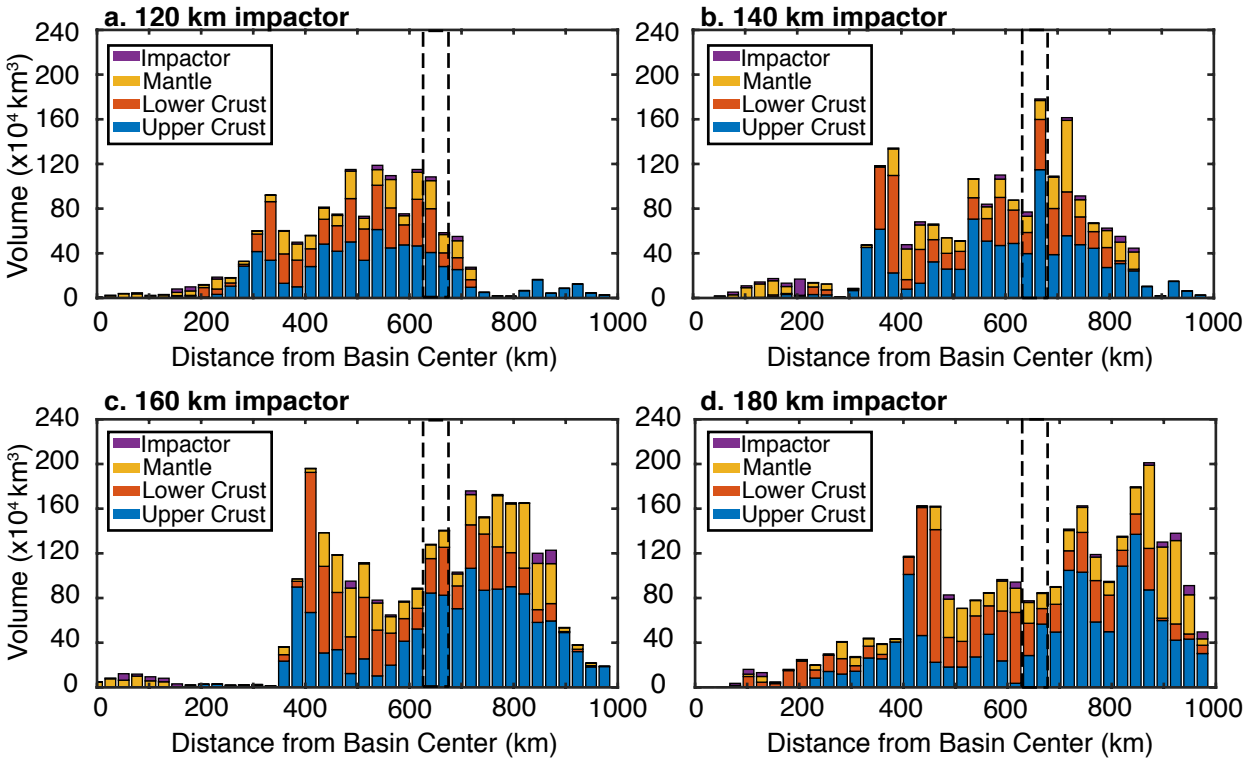

Figure 2.4. Histogram plot showing the volume of lagrangian tracer particles per bin outputted for a series of iSALE runs with varying impactor diameter. The tracers shown here are picked from above 15 km final depth. Tracer particles were binned with a width of 10 km. Impactor (purple) corresponds to the remnent impactor, upper crust (blue) corresponds to tracer particles with pre-impact depth above 20 km, lower crust (orange) corresponds to tracer particles with pre-impact depth between 20-40 km, and mantle (tan) corresponds to all material with pre-impact depth below 40 km. The black dotted lines outline the location where Jezero crater will form (~620 km from the center of Isidis Planitia). *Note: the stack plot does not represent distinct uniform layers or stratigraphy.*

Uncertainties within our modeling approach include the uncertainties in the thicknesses of the average Martian crust prior to impact as well as in the observed crustal thickness of Isidis Planitia due to volcanic infilling[8], mass wasting of the northeast rim[11] (Figure 2.1a), and the formation of Syrtis Major planum[3,12]; a volcanic province southwest of Isidis Planitia (see Figure 2.1a). Additionally, there has been some speculation as to whether the central peak collapse (Figures 4c-d) is a real phenomenon or an artifact of a weak mantle melt rheology invoked within the hydrocode[13]. In this case, for the modeled projectile diameter range, mantle material was

deposited at Jezero regardless of the pre-impact crustal thickness (40-60 km). In fact, every modeled run excavated mantle material. However, when the melt viscosity was greater than 10^{10} Pa·s for a 120 km impactor, the mantle material did not flow radially far enough to reach Jezero. Previous work showed that the assumed mantle melt rheology can significantly change the final calculated crustal thickness profile within the basin after transient crater collapse, with higher melt viscosities resulting in a thinner crust in the basin center and a thicker annulus of crust outside the rim[14]. For the runs that failed to produce mantle material at Jezero, a melt viscosity $\geq 10^{10}$ Pa·s resulted in an annulus of thickened crust ~50 km greater than observed. Thus, we conclude that the deposition of mantle material at the radial distance of Jezero occurs regardless of these uncertainties in our modeling.

For constant impact energy, a shallower angle of impact generates a smaller transient crater volume, decreases the depth of excavation, forms a more elliptical basin, decreases and partitions more peak shock pressure at shallower depths, and produces a more asymmetric distribution of ejecta[15–17]. Since we assumed a vertical impact within our simulations, it is necessary to discuss how impact angle may affect our results. The impact angle of the Isidis forming impact is unknown; however, given that only impact angles <15º result in elliptical basins[18] and Isidis is not elliptical (see Figure 2.1), we can rule out these most extreme impact angles. The most probable angle of impact is 45º with only 1 in 15 impacts occurring steeper than 75º[19]. Ejecta deposits begins to become axisymmetric for impact angles <45º with a ‘‘forbidden’’ azimuthal zones appear for impact angles <30º, first uprange, and then downrange of the crater for small (<15º) impact angles[15]. Orbital imagery shows a diverse suite of layered ejecta deposits as well as large megabreccia blocks predominately throughout the Nili Fossae region[20], which implies that if the impact angle was <30º than Jezero crater is downrange.

Crater volume roughly scales as $\sin^2(\theta)$, where θ is the impact angle[18], while peak shock pressure scales as $\sin(\theta)$[17]. At 45º, material will be shocked to ~70% the values reported in our results and the crater volume will be ~50%. Assuming a parabolic shape for the transient crater and a constant depth/diameter ratio between vertical and oblique impact, the transient crater depth scales as $\sin^{2/3}(\theta)$ (see Section 2.5). For our best-fit case, we had a transient crater depth of ~360 km, which is roughly three times the maximum depth of excavation[21]. With an impact angle of 45º and the same impact energy as our best-fit, the maximum excavation depth would be ~95 km, which would still excavate well below the determined pre-impact crustal thickness of 45 km[8].

Additionally, due to the decrease in cratering efficiency with decreasing impact angle, the impact diameter needs to be larger to produce the same transient crater volume. With everything else held constant, crater scaling would predict an impactor ~210 km in diameter to produce the same transient crater volume as our best-fit case (see Section 2.5).

2.3 Implications for the Martian Mantle

Our results indicate that the Perseverance rover should encounter a mix of crustal and upper mantle materials in the Isidis ejecta deposits that form the basement unit outside of Jezero crater. The large megabreccia blocks throughout the region dominantly exhibit an Opx spectral signature and variable alteration signatures[20]. There are discovered Martian orthopyroxenite samples from Mars, namely the Martian meteorite ALH84001, which is believed to be a crustal cumulate[22]. This might imply that the Opx-rich megabreccia in the circum-Isidis region originated from the lower crust; however, this raises the question of where the 20% of mantle material predicted by the models resides.

Olivine deposits in the Nili Fossae region were originally interpreted as mantle-derived impact melt deposited by the Isidis-forming impact[23,24]; however, recent work has reinterpreted

the deposits as post-Isidis volcanic tephra[25]. Therefore, we hypothesize that the unaltered Opx signatures observed in many megabreccia-bearing units could be consistent with a Martian upper mantle with a large, perhaps dominant, component of Opx (in addition to olivine). This would suggest that Mars has a distinct mantle mineralogy compared to the Earth, where the upper mantle contains a high fraction of olivine along with Opx. However, Opx-rich upper mantles may be common across the solar system. In particular, hydrocode modeling[26] of the Moon's ~2500 km diameter South Pole-Aitken impact basin showed that the lunar upper mantle was excavated and deposited on the adjacent surface. This analysis was consistent with the observed enrichment in mafic minerals such as orthopyroxene (Opx) contained within lunar highlands ejecta deposits[26]. Furthermore, hydrocode models of the formation of Phobos and Deimos (Mars' moons) show that >50% of their bulk composition coming from the Martian upper mantle[27]. The Infrared Spectrometer for Mars (ISM) instrument on board PHOBOS 2 recently observed spectra data on Phobos compatible with a mixture of low-calcium pyroxene and olivine[28]. Our hypothesis of an Opx-rich upper mantle is also consistent with recent analyses suggesting that asteroid 4 Vesta's mantle may also be dominated by Opx[29,30]. The Mars 2020 mission can use the results from the models presented here (Figure 2.4) to determine if the Opx-rich megabreccia is sourced from the Martian mantle.

2.4 Identifying Excavated Mantle Material with Mars 2020 Rover

The uppermost 15 km of Isidis ejecta deposited 600-700 km from the basin's center are shocked to variable peak pressures (Figure 2.5). The upper crust is either lightly shocked (0-4 GPa) or shocked above 100 GPa depending upon if the material was deposited via the ejecta curtain or the central peak collapse, respectively. At peak shock pressures above 100 GPa, whole rock melts[31]; therefore, figure 5 shows up to this point. The material derived from the lower crust exhibits a

more evenly distributed range of peak shock pressures with the majority (96%) of material shocked <50 GPa. The majority of mantle material has been shocked above >50 GPa. Additionally, 95% of the total ejected material shocked between 50-100 GPa in this area is mantle material. Experimentally derived effects of the progressive stages of shock metamorphism on basaltic achondrites have shown that shock pressures of >50 GPa demark the incipient edge-melting of pyroxene[31]. Using the Mastcam-Z, SuperCam, PIXL and SHERLOC, the Perseverance rover should be able to detect this unique characteristic in-situ within the observed megabreccia. For example, the Mg-rich composition of opx-rich rocks could be identified using the X-ray fluorescence high resolution spectrometry using PIXL, while impact melt vein textured could potentially be imaged at the outcrop scale using Mastcam-Z and sub-mm scales using PIXL and SHERLOC.

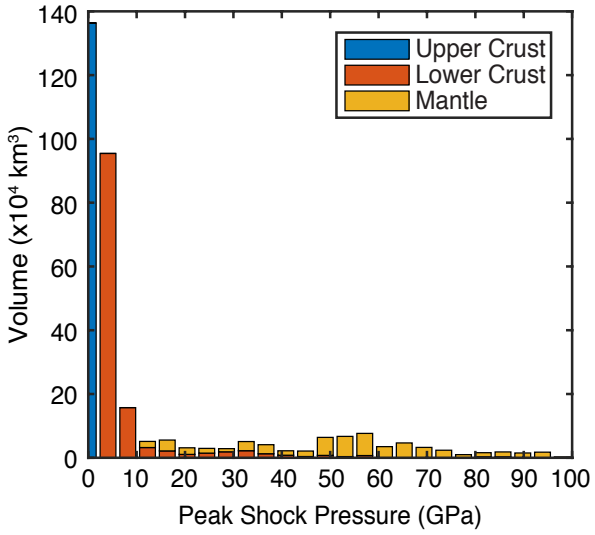

Figure 2.5. Histogram showing the volume of Lagrangian tracer particles shocked to various peak pressures for material deposited 600-700 km from basin center. The tracers shown here are picked from the top 15 km of deposited material. Tracer particles were binned with a width of 4 GPa. Upper crust (blue) corresponds to tracer particles with pre-impact depth above 20 km, lower crust (orange) corresponds to tracer particles with pre-impact depth between 20-40 km, and mantle (tan) corresponds to all material with pre-impact depth below 40 km.

Therefore, the results of this study can be used as a guide for the Mars 2020 science team to identify the composition of the lower crust/mantle at the time of the Isidis impact and select these critical petrologic samples for future sample return. In-situ and laboratory studies of Isidis basin ejecta will further enable a link to be established between ejecta composition, alteration, and depth

of provenance as well as better constrain early crustal modification processes. Lastly, if the Martian upper mantle is indeed Opx-rich, it raises the question why is the Earth seemingly unique with its olivine-rich mantle.

2.5 Methodology

2.5.1 Hydrocode Modeling

We used the shock physics hydrocode iSALE[32–34] to simulate the impact process. Our fiducial models assumed an axisymmetric planar target with 2 km grid spacing within a high-resolution zone spanning 1000 km and 450 km in the horizontal and vertical directions, respectively. Cells outside this region gradually increased to a maximum cell size of 10 km. Surface gravity was set at a constant value of 3.69 m/s^2. Given the size of Isidis relative to Mars' radius, the curvature of the Mars can influence the dynamics of the impact process[7]. Thus, we also conducted a spherical simulation (Figure 2.3) that, with the exception of grid spacing, utilized identical parameters from our best-fit fiducial model (Figure 2.2) which assumed a flat geometry; the spherical model used 1 km cells in the high-resolution zone. In addition, the spherical model used self-consistent gravity, pressure, and density fields that were calculated according to the thermal profile.

The Martian crust and mantle were modeled using basalt[35] and dunite[36] ANEOS equations of state, respectively. Additionally, the crustal strength parameters are fit to basalt which is appropriate for the Martian crust[37]. The rheologic properties of the Martian crust and mantle were simulated by incorporating a rock-like strength model[33], a damage model with an exponential dependence on plastic strain[14,33], a dilatancy model[38], tensile failure model[33], and a thermal weakening model[39] corresponding to the temperature- and pressure-dependence of the material.

Additionally, we incorporated a visco-elastic-plastic rheology for the mantle[16]. To produce a realistic surface gravity, we incorporated an iron core into our Martian spherical target with a radius of 1522 km. Core material was modeled using the iron ANEOS equation of state[40] and the Johnson-Cook strength model with inputs for Armco Iron[41,42]. Assumed hydrocode values and their corresponding references are listed in Table 2.1.

Table 2.1. iSALE Parameters for Isidis.

Description	Crust Values	Mantle Values	Core Values	References
Equation of State	ANEOS basalt	ANEOS dunite	ANEOS iron	[35; 36; 40]
Melting temperature	1393 K	1373 K	1811 K	[37; 46; 41, 42]
Thermal softening parameter	1.2	1.1	1.2	[37; 46; 41, 42]
Simon A parameter	6000 MPa	1520 MPa	6000 MPa	[37; 46; 41, 42]
Simon C parameter	3	4.05	3	[37; 46; 41, 42]
Poisson's ratio	0.25	0.25	0.29	[37; 46; 41, 42]
Frictional coefficient (damaged)	0.6	0.63	N/A	[37; 46]
Frictional coefficient (undamaged)	1.2	1.58	N/A	[37; 46]
Strength at infinite pressure	3.5 GPa	3.26 GPa	N/A	[37; 46]
Cohesion (damaged)	0.01 MPa	0.01 MPa	N/A	[37; 46]
Cohesion (undamaged)	10 MPa	5.07 MPa	100 MPa	[37; 46; 41, 42]
Brittle ductile transition	1.23 GPa	1.23 GPa	N/A	[38]
Brittle plastic transition	2.35 GPa	2.35 GPa	N/A	[38]
Initial tensile strength	10 MPa	10 MPa	10 MPa	[38]
Maximum distension	1.2	1.2	N/A	[38]
Maximum dilatancy coefficient	0.045	0.045	N/A	[38]
Dilatancy pressure limit	200 MPa	200 MPa	N/A	[38]
Frictional coefficient (maximum distension)	0.4	0.4	N/A	[38]
Minimum Pressure	N/A	N/A	-2440 MPa	[41, 42]
Johnson-Cook A parameter	N/A	N/A	101 MPa	[41, 42]

Table 2.1. continued

Johnson-Cook B parameter	N/A	N/A	219 MPa	[41, 42]
Johnson-Cook N parameter	N/A	N/A	0.32	[41, 42]
Johnson-Cook C parameter	N/A	N/A	0	[41, 42]
Johnson-Cook M parameter	N/A	N/A	0.55	[41, 42]
Johnson-Cook reference temperature	N/A	N/A	2.5	[41, 42]

Note: References are in order corresponding to column values.

The zero-pressure melting temperature for dunite was chosen to be 1373 K; the approximate peridotite solidus. Using the peridotite solidus as the melting temperature for dunite is suitable if the thermal profiles follow an adiabatic gradient (0.5 K/km) at sub-solidus temperatures more than 1400 K[43]; thereby, suppressing the presence of abundant super-solidus material. Material between the liquidus and solidus temperatures, however, retains some residual shear strength as super-solidus material contains a combination of melt, and hot and cold clasts of rock where the clasts provide some resistance to shear[13]. To address this discrepancy in material strength, we invoked a melt viscosity as an input variable[44–46] which provides the dunite some shear resistance above the melting temperature.

We explored a range of model parameters that were consistent with recent estimates and their respective uncertainties. We varied the assumed impactor diameter (120-200 km[7]) while maintaining a constant impact velocity of 12 km/s[47]. Given the range of current estimates for Mars' crustal thickness distribution[8], we varied the average pre-impact crustal thickness between 40-60 km. Near surface thermal gradients between 20-30 K/km were considered due to the uncertainty in the early Martian thermal structure[48]. Additional uncertainties in the rheologic structure of the Martian mantle led us to vary the temperatures (1300 K and 1400 K) at which our near surface thermal gradient transitioned to an adiabatic gradient, and utilize melt viscosities between 0 and 10^{10} Pa s.

A typical iSALE run begins immediately before contact between the target and projectile. It models the subsequent few hours following impact, which include the excavation and ejection of material, the formation and collapse of the transient crater, and one or more cycles of central uplift and collapse. For large basins, the crust can migrate back in towards the basin center during transient crater collapses. Our models are run until crustal migration stops and the basin topography achieves a stable, steady state condition.

2.5.2 Scaling Maximum Depth of Excavation with Impact Angle

Crater volume roughly scales as $\sin^2(\theta)$, where θ is the impact angle[18]. The transient crater is approximately paraboloid in shape[21] with a volume (V) defined as:

$$V = \pi H_{at} {D_{at}}^2/8 \quad (1)$$

where H_{at} is the transient crater depth and D_{at} is the transient crater diameter. If V_{45} (transient crater volume for impact angle of 45°) is equal to $V_{90}\sin^2(45°)$ (transient crater volume for impact angle of 90°) then the transient crater depth and diameter are related as follows from equation 1:

$$(H_{at})_{90}{(D_{at})_{90}}^2 \sin^2(45°) = (H_{at})_{45}{(D_{at})_{45}}^2 \quad (2)$$

where $(H_{at})_{90}$ is the transient depth for an impact angle of 90°,$(D_{at})_{90}$ is the transient diameter for an impact angle at 90°, $(H_{at})_{45}$ is the transient depth for an impact angle of 45°, and $(D_{at})_{45}$ is the transient diameter for an impact angle at 45°. Assuming a constant depth to diameter ratio (i.e. $(H_{at})_{90}/(D_{at})_{90} = (H_{at})_{45}/(D_{at})_{45}$; which experimentally has been shown to be a close approximate for impact angles 90-30º)[49], then equation 2 simplifies to:

$$(H_{at})_{45} = (H_{at})_{90} \sin^{2/3}(45°) \quad (3)$$

Plugging our value for $(H_{at})_{90}$ from our best-fit run (~360 km) gives a transient crater depth of ~286 km. The transient crater depth is roughly three times the maximum depth of excavation[21]; therefore, the maximum depth of excavation for the same impact energy at 45° is ~95 km.

2.5.3 Scaling Impactor Diameter with Impact Angle

The volume of the transient crater excavated by the impact can be solved with impact angle using the following equation[17]:

$$V = 0.28\frac{\rho_p}{\rho_t}{D_p}^{2.35}\mathrm{g}^{-0.65}{v_i}^{1.3}\sin^{1.3}(\theta) \quad (4)$$

where ρ_p is projectile density, ρ_t is target density, g is surface gravit, v_i is impact velocity, and D_p is projectile diameter. At an impact angle of 45°, equation 1, and values from our best-fit iSALE simulation (ρ_p=3300 kg/m^3, ρ_t=2900 kg/m^3, g=3.69 m/s^2, v_i=12 km/s, D_{at}=660 km, and V=6.1581*10^{16} m^3), we can calculate the projectile diameter needed to produce the same transient crater volume at 45° as a vertical impact: ~210 km.

2.6 References

1. Smith, D. E. *et al.* Mars Orbiter Laser Altimeter: Experiment summary after the first year of global mapping of Mars. *J. Geophys. Res. E Planets* **106**, 23689–23722 (2001).
2. Genova, A. *et al.* Seasonal and static gravity field of Mars from MGS, Mars Odyssey and MRO radio science. *Icarus* **272**, 228–245 (2016).
3. Ivanov, M. A., Hiesinger, H., Erkeling, G., Hielscher, F. J. & Reiss, D. Major episodes of geologic history of Isidis Planitia on Mars. *Icarus* **218**, 24–46 (2012).
4. Frey, H. V. Impact constraints on, and a chronology for, major events in early Mars history. *J. Geophys. Res. E Planets* **111**, 1–11 (2006).

5. Horgan, B. H. N., Anderson, R. B., Dromart, G., Amador, E. S. & Rice, M. S. The mineral diversity of Jezero crater: Evidence for possible lacustrine carbonates on Mars. *Icarus* **339**, 113526 (2020).

6. Weiss, B. *et al.* Megabreccia at Northeast Syrtis Major and its importance for mars science. in *49th Lunar and Planetary Science Conference* (2018).

7. Mancinelli, P., Mondini, A. C., Pauselli, C. & Federico, C. Impact and admittance modeling of the Isidis Planitia, Mars. *Planet. Space Sci.* **117**, 73–81 (2015).

8. Neumann, G. A. *et al.* Crustal structure of Mars from gravity and topography. *J. Geophys. Res. E Planets* **109**, 1–18 (2004).

9. Grott, M., Hauber, E., Werner, S. C., Kronberg, P. & Neukum, G. Mechanical modeling of thrust faults in the Thaumasia region, Mars, and implications for the Noachian heat flux. *Icarus* **186**, 517–526 (2007).

10. Morgan, J. V *et al.* The formation ofpeak rings in large impact craters. *Geology* **354**, 878–883 (2016).

11. Tanaka, K. L., Kargel, J. S., MacKinnon, D. J., Hare, T. M. & Hoffman, N. Catastrophic erosion of Hellas basin rim on Mars induced by magmatic intrusion into volatile-rich rocks. *Geophys. Res. Lett.* **29**, 37-1-37-4 (2002).

12. Hiesinger, H. & Head, J. W. The Syrtis Major volcanic province, Mars: Synthesis from Mars Global Surveyor data. *J. Geophys. Res. E Planets* **109**, (2004).

13. Stewart, S. T. Impact basin formation: The mantle excavation paradox resolved. *Lunar Planet. Inst. Sci. Conf. Abstr.* **42**, 1633 (2011).

14. Johnson, B. C. *et al.* Formation of the Orientale lunar multiring basin. *Science* **354**, 441–444 (2016).

15. Gault, D. E. & Wedekin, J. A. Experimental studies of oblique impact*. in *Proc. Lunar Planet. Sci. Conf. 9th*, 3843–3875 (1978).

16. Elbeshausen, D. & Melosh, H. J. A nonlinear and time---dependent visco--- elasto---plastic rheology model for studying shock---physics phenomena. *arXiv preprint arXiv:1805.06453* (2018).

17. Pierrazzo, E. & Melosh, H. J. Understanding oblique impacts from experiments, observations, and modeling. *Annu. Rev. Earth Planet. Sci.* 1–39 (2000).

18. Elbeshausen, D., Wünnemann, K. & Collins, G. S. Scaling of oblique impacts in frictional targets: Implications for crater size and formation mechanisms. *Icarus* **204**, 716–731 (2009).

19. Gilbert, G. K. The Moon's Face: A Study of the Origin of Its Features. *Bull. Philos. Soc. Wash.* **12**, 241–292 (1893).

20. Scheller, E. L. & Ehlmann, B. L. Composition, Stratigraphy, and Geological History of the Noachian Basement Surrounding the Isidis Impact Basin. *J. Geophys. Res. Planets* **125**, 0–2 (2020).

21. Melosh, H. J. *Impact Cratering A Geologic Process*. (Oxford University Press., 1989).

22. Mittlefehldt, D. W. ALH84001, a Cumulate Orthopyroxenite Member of the martian meteorite clan. **221**, 214–221 (1994).

23. Mustard, J. F. *et al.* Mineralogy of the Nili Fossae region with OMEGA/Mars Express data: 1. Ancient impact melt in the Isidis Basin and implications for the transition from the Noachian to Hesperian. *J. Geophys. Res. E Planets* **112**, 1–14 (2007).

24. Mustard, J. F. *et al.* Composition, morphology, and stratigraphy of Noachian crust around the Isidis basin. *J. Geophys. Res. E Planets* **114**, 1–18 (2009).

25. Kremer, C. H., Mustard, J. F. & Bramble, M. S. A widespread olivine-rich ash deposit on Mars. *Geology* **47**, 677–681 (2019).

26. Melosh, H. J. *et al.* South Pole-Aitken basin ejecta reveal the Moon's upper mantle. *Geology* **45**, 1063–1066 (2017).

27. Hyodo, R., Genda, H., Charnoz, S. & Rosenblatt, P. On the Impact Origin of Phobos and Deimos. I. Thermodynamic and Physical Aspects. *Astrophys. J.* **845**, 125 (2017).

28. Gendrin, A., Langevin, Y. & Erard, S. ISM observation of Phobos reinvestigated: Identification of a mixture of olivine and low-calcium pyroxene. *J. Geophys. Res. E Planets* **110**, 1–13 (2005).

29. McCord, T. B. & Scully, J. E. C. The composition of Vesta from the Dawn mission. *Icarus* **259**, 1–9 (2015).

30. McSween, H. Y. *et al.* Composition of the Rheasilvia basin, a window into Vesta's interior. *J. Geophys. Res. E Planets* **118**, 335–346 (2013).

31. Bischoff, A. & Stoffler, D. Shock metamorphism as a fundamental process in the evolution of planetary bodies: information from meteorites. *Eur. J. Mineral.* **4**, 707–755 (1992).

32. Amsden, A. A., Ruppel, H. M. & Hirt, C. W. SALE: A Simplified ALE computer program for fluid flow at all speeds. *Los Alamos Natl. Lab. Rep.* **LA-8095**, 101p (1980).

33. Collins, G. S., Melosh, H. J. & Ivanov, B. A. Modeling damage and deformation in impact simulations. *Meteorit. Planet. Sci.* **39**, 217–231 (2004).

34. Wünnemann, K., Collins, G. S. & Melosh, H. J. A strain-based porosity model for use in hydrocode simulations of impacts and implications for transient crater growth in porous targets. *Icarus* **180**, 514–527 (2006).

35. Pierazzo, E., Artemieva, N. A. & Ivanov, B. A. Starting conditions for hydrothermal systems underneath Martian craters: Hydrocode modeling. *Spec. Pap. Geol. Soc. Am.* **384**, 443–457 (2005).

36. Benz, W.; Cameron, A. G.W.; Melosh, H. J. The origin of the Moon and the single impact hypothesis V. *Icarus* **81**, 113–131 (1989).

37. Potter, R. W. K. & Head, J. W. Basin formation on Mercury: Caloris and the origin of its low-reflectance material. *Earth Planet. Sci. Lett.* **474**, 427–435 (2017).

38. Collins, G. S. Numerical simulations of impact crater formation with dilatancy. *J. Geophys. Res. Planets* **119**, 2600–2619 (2014).

39. Ohnaka, M. A shear failure strength law of rock in the brittle-plastic transition regime. *Geophys. Res. Lett.* **22**, 25–28 (1995).

40. Thompson, S. ANEOS Analytic Equations of State for shock physics codes input manual. *SANDIA Rep.* **3**, (1990).

41. Bowling, T. J. *et al.* Antipodal terrains created by the Rheasilvia basin forming impact on asteroid 4 Vesta. *J. Geophys. Res. Planets* **118**, 1821–1834 (2013).

42. Johnson, G.R. and Cook, W. H. A Constitutive Model and Data for Metals Subjected to Large Strains, High Strain Rates, and High Temperatures. in *Seventh International Symposium on Ballistics* 541–547 (1983).

43. Freed, A. M. *et al.* The formation of lunar mascon basins from impact to contemporary form. *J. Geophys. Res. E Planets* **119**, 2378–2397 (2014).

44. Potter, R. W. K., Kring, D. A., Collins, G. S., Kiefer, W. S. & McGovern, P. J. Estimating transient crater size using the crustal annular bulge: Insights from numerical modeling of lunar basin-scale impacts. *Geophys. Res. Lett.* **39**, 1–5 (2012).

45. Potter, R. W. K., Kring, D. A. & Collins, G. S. Scaling of basin-sized impacts and the influence of target temperature. **2518**, 99–113 (2015).

46. Potter, R. W. K., Collins, G. S., Kiefer, W. S., McGovern, P. J. & Kring, D. A. Constraining the size of the South Pole-Aitken basin impact. *Icarus* **220**, 730–743 (2012).

47. Bottke, W.F., Nolan, M.C., Greenberg, R. , Kolvoord, R. A. Collisional life time and impact statistics of near-Earth asteroid. in *Gehrels,T., Matthews,M.S. (Eds.), Hazards due to Comets and Asteroids.* 337–357 (UoA Press, 1994).

48. Schubert, G. & Spohn, T. Thermal history of Mars and the sulfur content of its core. *J. Geophys. Res.* **95**, (1990).

49. Gault, D. E. Displaced mass, depth, diameter, and effects of oblique trajectories for impact craters formed in dense crystalline rocks. the *moon* **6**, 32–44 (1973).

2.7 Supplementary Material

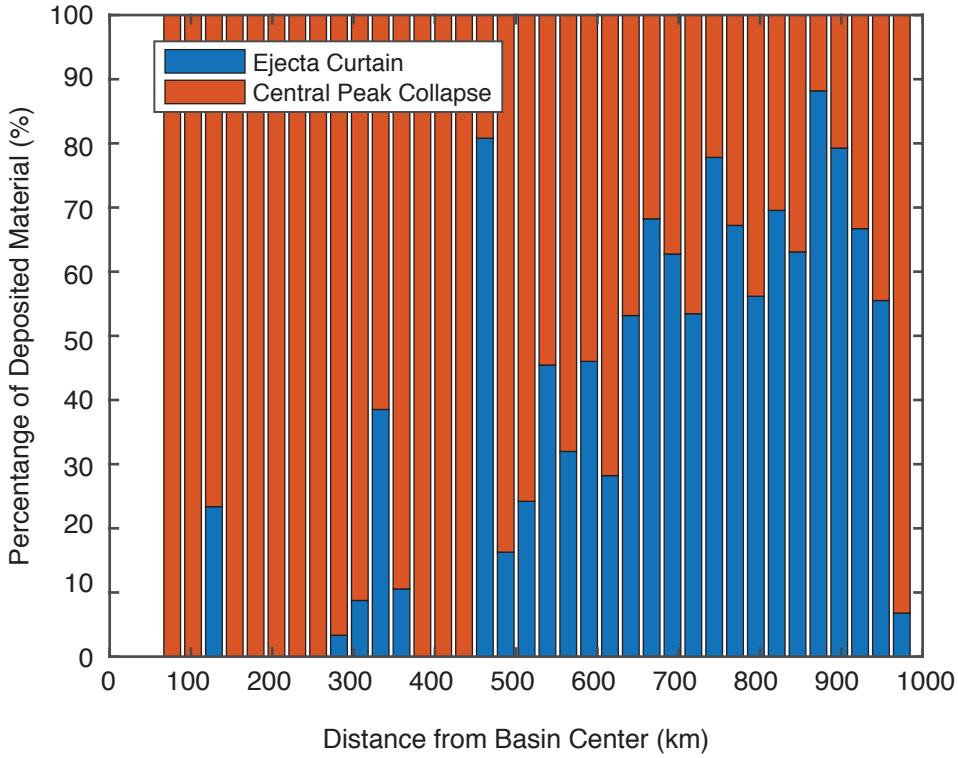

Figure 2.6. Histogram plot showing the percentange of all excavated mantle material deposited via the central peak collapse (orange); or ejecta curtain (blue). The results are from our best-fit run.

CHAPTER 3. VIGOROUS CONVECTION AS THE EXPLANATION FOR PLUTO'S POLYGONAL TERRAIN

This chapter was published in the journal *Nature* on 01 June 2016; DOI: 10.1038/nature18016

3.1 Main Text

Pluto's surface is surprisingly young and geologically active[1]. One of its youngest terrains is the near-equatorial region informally named Sputnik Planitia, which is a topographic basin filled by nitrogen (N_2) ice mixed with minor amounts of CH_4 and CO ices[1]. Nearly the entire surface of the region is divided into irregular polygons about 20–30 kilometres in diameter, whose centres rise tens of metres above their sides. The edges of this region exhibit bulk flow features without polygons[1]. Both thermal contraction and convection have been proposed to explain this terrain[1], but polygons formed from thermal contraction (analogous to ice-wedges or mud-crack networks)[2,3] of N_2 are inconsistent with the observations on Pluto of non-brittle deformation within the N_2-ice sheet. Here we report a parameterized convection model to compute the Rayleigh number of the N_2 ice and show that it is vigorously convecting, making Rayleigh–Bénard convection the most likely explanation for these polygons. The diameter of Sputnik Planitia's polygons and the dimensions of its 'floating mountains' of water ice suggest that its N_2 ice is about ten kilometres thick. The estimated convection velocity of 1.5 centimetres a year indicates a surface age of only around a million years.

Previous work first proposed that convection or thermal contraction could have formed the polygons on Sputnik Planitia[1] (see Figure 3.1). However, we find contraction unlikely: studies of Arctic ice-wedges show that the spacing of thermal contraction polygons is typically about five times the annual thermal skin depth (the depth to which the summer–winter thermal wave penetrates into the surface of a planetary body)[4]. Using reasonable values for the thermal diffusivity of N_2 ice[5], we compute the annual thermal skin depth to be about 100 m, corresponding to thermal contraction polygons around 500 m across; this is nearly two orders of magnitude smaller than the observed polygons. Furthermore, contractional polygons require brittle failure of the ice. However, viscoelastic deformation of N_2 ice over annual (248 years) and diurnal (153 Earth hours cycles[6] on Pluto can easily relax differential stresses on this timescale, preventing

brittle failure in response to such slowly building stress (the Maxwell time, over which stresses relax by 1/e, of N_2 ice at 40 K is about 4 min at a stress of 0.1 MPa). Although there are other ways to generate polygons (such as compaction of sediments over heavily cratered terrain[7], extensional tectonic processes[8] or contraction of cooling cryovolcanic flows[9,10], these processes occur on timescales that are significantly longer than the Maxwell time, and are inconsistent with the lack of craters and observed flow features within this region.

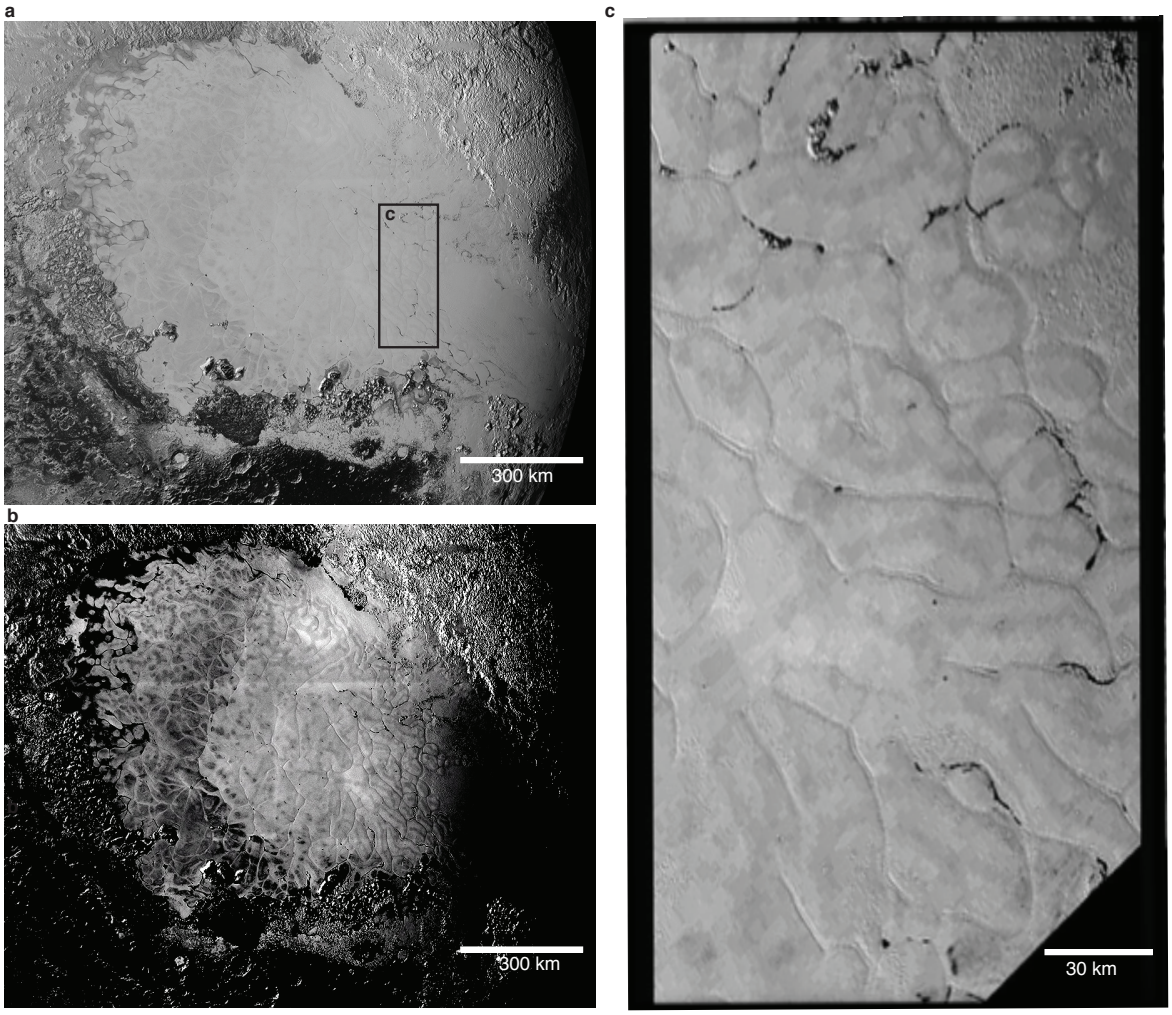

Figure 3.1. New Horizon's image of Sputnik Planitia on Pluto. A mosaic image of Sputnik Planitia is shown in a. Within the centre of the ice field, where the ice is presumably thickest, the polygons are approximately 30 km across[1]. Close to the edge, the average polygon diameter decreases to 20 km and then vanishes, leaving a smooth surface. A contrast-enhanced version of a is given in b to better illuminate the polygons. The 'floating mountains' are observable within the edges of these polygons, and can be seen in c, the zoom of the rectangle in a. Image credit: NASA/John Hopkins University-Applied Physics Laboratory/Southwest Research Institute (2015).

The viability of Rayleigh–Bénard convection as an explanation for Pluto's polygons depends critically on the thickness of the convecting layer[11]. Because spectral data only probes micrometres into the surface, the N_2 ice could be only a thin surface veneer, making convection impossible. However, this possibility is unlikely given the high exchange rates of the N_2 atmosphere with a surface ice reservoir[12]. We can estimate the ice thickness from the observed polygon size and the depth to diameter ratio (that is, aspect ratio) for a Rayleigh–Bénard convection cell. Laboratory convection experiments and three-dimensional numerical modelling almost invariably predict aspect ratios near 3:1 (ref. 11), as assumed here. Some two-dimensional numerical simulations of convection in fluids with strongly temperature-dependent viscosity predict larger aspect ratios[13,14], implying a thinner ice layer. However, N_2-ice viscosity depends only weakly on temperature[15] and falls in the small-viscosity-contrast regime that precludes aspect ratios larger than 3:1[14,16]. Moreover, if Pluto's 'floating mountains' (see Figure 3.1c) are truly supported by their buoyancy, then their heights, widths and the small density contrast between N_2 ice and water ice requires an N_2-ice thickness of at least 5 km, as shown in the Methods. We account for uncertainty in the layer thickness by widely varying the depth of the convection cell within our model (Figure 3.2).

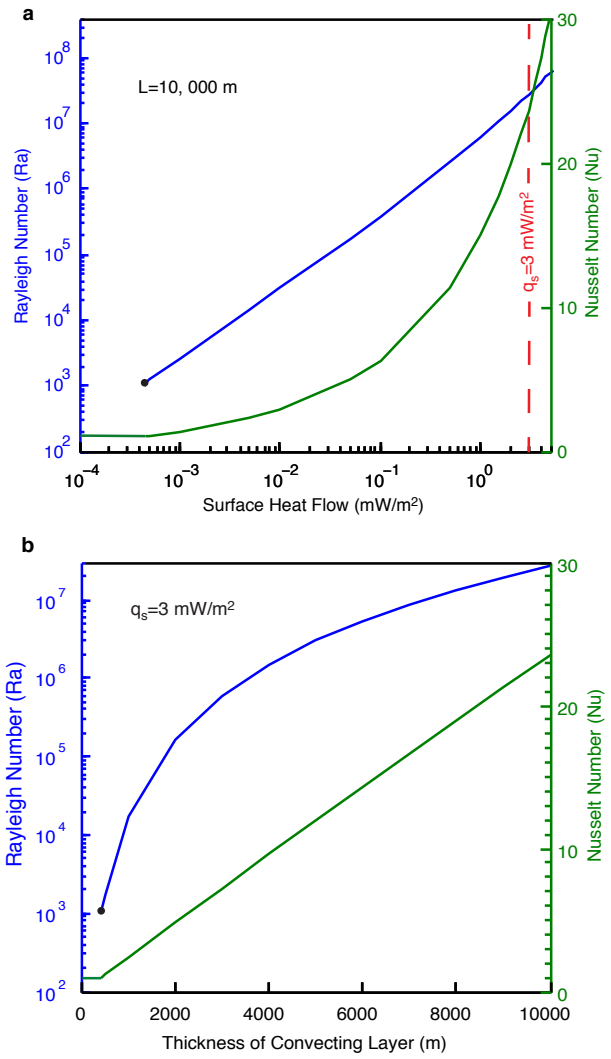

Figure 3.2. Calculated convection for Sputnik Planitia polygons. The calculated Rayleigh number and Nusselt number as a function of surface heat flow (q_s) and thickness of convecting layer (L) are shown in **a** and **b**, respectively. The blue line is the Rayleigh number; the green line is the Nusselt number (see Methods section). The black dot marks the point where the Rayleigh number reaches the critical value (~1,000) at which convection just begins. At this point, the Nusselt number equals 1, and heat is transferred entirely by conduction. The calculation as a function of surface heat flow(**a**) shows that the Rayleigh number remains above ~1,000 for surface heat flows down to 4×10^{-4} mW m^{-2} for a 10-km-thick N_2 layer. The estimated surface heat flow for Pluto (3 mW m^{-2}) is marked by a vertical red line. For a constant heat flow of 3 mW m^{-2}, the Rayleigh number (**b**) decreases with the thickness of the convecting layer until convection stops at a thickness of 425 m.

Using extrapolated N_2-ice rheology measurements[15] (see Methods and parameters used in equation (9)), a surface temperature of 33 K, and a surface heat flux q_s of ~3 mW m^{-2} (consistent

with the radiogenic heat generated by a carbonaceous chondrite core about 900 km in radius, as suggested by Pluto's mean density), we calculate a Rayleigh number $Ra > 10^6$ and an interior temperature of approximately 40 K for the N_2-ice layer. This Rayleigh number is four orders of magnitude greater than the critical value that denotes the onset of convection ($R_{crit} \approx 1{,}000$), suggesting that Sputnik Planitia's N_2 ice is vigorously convecting.

Figure 3.2 shows the calculated Rayleigh and Nusselt numbers for a range of surface heat flows and N_2 thicknesses. In our model, the Rayleigh number remains above the critical value for surface heat flows as small as 4×10^{-4} mW m^{-2} (see Figure 3.2a), suggesting that our results are robust against uncertainties in our estimated surface heat flux. Figure 3.2b illustrates that as the thickness of the convecting cell decreases, the Rayleigh number drops until convection ceases at a thickness of 425 m (at a nominal heat flow of 3 mW m^{-2}). Thus, the observed decrease in polygon size away from the centre of Sputnik Planitia, and their absence at its edges (see Figure 3.1), both suggest that the depth of Sputnik Planitia's N_2 ice is thickest at the centre, and thins to around 400 m near the edges, where the polygons are absent. This result is consistent with the hypothesis that N_2 ice in Sputnik Planitia fills a topographic basin.

From the calculated average velocity of convection, ~1.5 cm yr^{-1} (the equation for velocity is given in the Methods), we compute the time needed for the ice surface to renew itself, and therefore the maximum age of the surface of Sputnik Planitia, to be about one million years. This is consistent with the lack of significant cratering, and further constrains the existing age estimates of a few hundred million years[1] by two orders of magnitude. The convection model also correctly predicts the topography of the polygons. The centres of convection cells are underlain by (relatively) warm rising currents and should therefore stand higher than the edges of the cells, where cooler ice descends. The elevation of the polygon centres above their edges is estimated from the convective buoyancy stress to be about 80 m (explained in more detail in the Methods), in good agreement with observations[1].

Our predicted central temperature of the convection cell of 40 K is close to the α-to-β phase transition temperature for N_2 ice ($T_{\alpha\beta}$ = 35.61 K for pure N_2 ice, but is higher if CO is dissolved in the N_2 ice[17]). Furthermore, high concentrations of CO have been reported on Sputnik Planitia[1]. CO and N_2 form a complete solid solution series (see Figure 3.3). With a CO concentration >10%, both phases are stable at 40 K (ref. 17), which may allow a phase-change-induced mixed

convection system to develop at these high Rayleigh numbers[18]. This may consist of warm β N_2 upwelling in the convection cells, while α N_2 concentrates in the troughs of the polygonal features, which might explain the albedo difference between the polygons and troughs. Owing to the difference in absorption spectra for each phase, New Horizon's Alice instrument can test this prediction of our convection model. Alternatively, a two-layer convection system may develop, in which the β N_2 forms the lower, deeper convection cell while α N_2 forms a layer of convection cells above. However, the positive Clapeyron slope of the α-to-β phase transition makes this alternative scenario unlikely[19]. Rather, the exothermic β-to-α phase change encourages single-cell overturn and produces a more stable convecting regime[20].

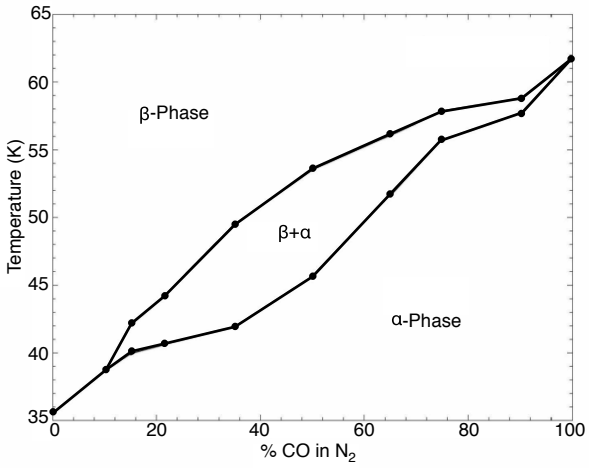

Figure 3.3. The N_2–CO phase diagram. We obtained this figure by collecting data points from experimental measurements[17]. Above CO concentrations of 10%, both phases of nitrogen are stable at Pluto's surface temperatures.

The N_2-ice mass in Sputnik Planitia seems to be the largest concentration of N_2 on Pluto. If our estimated 10-km ice thickness were evenly spread across the planet it would form a layer about 350 m thick that would, if converted to vapour, produce an atmosphere with a surface pressure of about one bar instead of the currently observed pressure of approximately ten microbars. The present atmospheric pressure is presumably controlled by the vapour pressure of N_2 ice at the current surface temperature, so that N_2 can either evaporate or condense onto the ice reservoir in Sputnik Planitia as temperature varies during Pluto's seasons and annual excursions from the Sun. N_2 ice should thus be mobile across Pluto's surface. We do not at present understand why most of the N_2 on Pluto is concentrated in what appears to be the basin of a large ancient

impact crater: that must be the subject of future climatological studies. However, it is an observational fact that most of the N_2 on Pluto is concentrated in a single large mass that lies in a basin nearly at the equator rather than at its poles, which is perhaps related to Pluto's large obliquity. The polygonal surface features of this mass indicate that it is vigorously convecting, and this convection is driven by the small amount of heat conducted through Pluto's lithosphere.

3.2 Methods

3.2.1 Maxwell time calculation

The Maxwell time (τ_m in units of seconds) is determined from the ratio of the effective viscosity (η_{eff}) to the shear modulus (μ in units of pascals):

$$\tau_m = \frac{\eta_{eff}}{\mu} \tag{1}$$

The shear modulus was determined from experimentally measured shear wave velocities (v_s in units of metres per second)[21], while the viscosity parameters are derived from the measurements of N_2 ice[15] and are listed below. The shear modulus is related to the shear wave velocity and material density (ρ in units of kilograms per cubic metre) by the expression:

$$\mu = v_s^2 \rho \tag{2}$$

At a differential stress of 0.1 MPa and a temperature of 30 K, the Maxwell time is about 4 min.

3.2.2 Temperature-dependent parameters for Rayleigh number

The temperature dependent values for the parameters in equation (7) (see below) were determined from best-fitting experimental data[21]. The density, volume coefficient of expansion (α in units of per kelvin), and thermal conductivity (k in units of watts per millikelvin) of N_2 are expressed by the following equations:

$$\rho = 0.0134T^2 - 0.6981T + 1038.1 \tag{3}$$

$$\alpha = \left(2\times10^{-6}\right)T^2 - 0.0002T + 0.006 \tag{4}$$

$$k = 0.1802T^{0.1041} \tag{5}$$

The thermal diffusivity (κ in units of metres squared per second) was determined from the relationship $\kappa = k/(\rho c_p)$, where c_p (in units of joules per kilogram per kelvin) is the specific heat at constant pressure. Specific heat is also temperature dependent, which is expressed as:

$$c_p = 926.91 e^{0.0093 T} \tag{6}$$

3.2.3 Parameterized convection model

The Rayleigh number assesses the vigour of thermal convection in a fluid layer under the influence of gravity and a downward-increasing thermal gradient. The Rayleigh number is essentially the ratio between the timescales for conductive cooling and buoyancy-driven overturn of the viscous fluid. The conventional expression for the Rayleigh number[11] requires knowledge of the temperature difference between the top and bottom of the convecting layer. Because the thermal gradient in Pluto is a priori unknown, but the surface heat flow can be estimated at about 3 mW m^{-2} from internal heat production, we employ a version of the Rayleigh number based on surface heat flow[11]:

$$\mathrm{Ra}_q = \frac{\alpha \rho g L^4}{\kappa k \eta} q_s \tag{7}$$

where α is the volume coefficient of thermal expansion, ρ is the fluid density, g is the acceleration of gravity (0.62 m s^{-2} on Pluto), L is the depth of the convecting layer, κ is thermal diffusivity, k the thermal conductivity, η is the viscosity (which depends strongly on temperature and stress in N_2 ice), and q_s is the surface heat flow. For Sputnik Planitia, we used (see below) measured, temperature-dependent values for N_2 in equation (1). This version of the Rayleigh number is related to the more standard version Ra through the Nusselt number Nu as follows: $\mathrm{Ra}_q = \mathrm{Nu}\ \mathrm{Ra}$. The Nusselt number is the ratio between the total heat transported by both convection and conduction to conductive heat transport only and is given by:

$$\mathrm{Nu} = \left(\frac{\mathrm{Ra}}{\mathrm{Ra}_c} \right)^{\beta} \text{ for } \mathrm{Ra} > \mathrm{Ra}_c \tag{8}$$

The critical Rayleigh number Ra_c I had difficulty changing Eq. 8 is of order 10^3, depending on detailed boundary conditions, and β has been measured to be 0.31 over a wide range of Ra values[22].

Equations (7) and (8) together define the "parameterized convection model", which has been widely used to model heat transport in planetary mantles[11,23]. The most important variable in these models is the viscosity of the convecting fluid. Unlike ideal liquids, the viscosity of a hot, creeping solid is a sensitive function of both deviatoric stress σ and temperature T and is often parameterized by the form[24]

$$\eta_{\text{eff}} = \frac{\sigma^{1-n}}{3A} \mathrm{e}^{Q/RT} \tag{9}$$

where Q is the activation enthalpy for creep and R is the gas constant. A is a constant that, along with Q, and n, must be determined experimentally. The viscosity parameters we determined for N_2 ice used in equation (9) are activation energy $Q = 3.5$ kJ mol^{-1}, $n = 2.2$ and $A = 3.5 \times 10^{-12}$ Pa^{-n} s^{-1}.

Because the effective viscosity depends strongly on both the temperature and stress in the convecting system, which vary widely from place to place, it is important to understand how to define them in a meaningful way. Previous work[11] has shown that the best choice is to use the mean temperature and buoyancy stress for accurate estimates of convective vigour, a choice that we follow here.

As they stand, equations (7) to (9) do not define a closed system and more information is required to compute Ra, even given the heat flow and material properties of the convecting fluid. The system can, however, be closed by recognizing that the mean temperature in a convecting fluid is determined by the surface temperature, T_s, heat flow q_s and the temperature drop across the cold (surface) boundary layer, whose thickness is itself determined from the Nusselt number. We thus set:

$$T = T_s + \frac{q_s L}{k\left(\mathrm{Nu}+1\right)} \tag{10}$$

where we have ignored the adiabatic increase of temperature in the convecting layer, a valid approximation for a thin layer, such as the N_2-ice deposit in Sputnik Planitia. Further adding an equation for the average deviatoric stress in convecting plumes:

$$\sigma = \left(\frac{\mathrm{Ra}_q}{\mathrm{Nu}^2}\right)\frac{\eta_{\text{eff}}\kappa}{L^2} \tag{11}$$

Equations (7) to (11) now define a closed, if highly nonlinear, system that can readily be solved by numerical methods to define most of the properties of the convecting layer from the properties of the fluid, the surface heat flow and surface temperature.

3.2.4 Viscosity of N_2 ice

Equations (3) to (6) describe all the material parameters in equation (7) except for viscosity. The stress-dependent parameters, A and n, within the viscosity equation are directly quoted from previous works[15]. The temperature-dependent parameter was determined from existing stress and strain rate (that is, viscosity) measurements for solid N_2 at 45 K and 56 K (ref. 15). By matching data points for the viscosity measured at two temperatures under the same applied strain rate, we can solve for the temperature-dependent parameter, Q, in equation (9).

3.2.5 Velocity of the convecting fluid

The mean velocity of a convecting layer is computed by comparing the surface heat flow to the rate at which warm fluid moves towards the surface and deposits its thermal energy. This equality can be written in terms of the Nusselt number Nu as:

$$\bar{v}_{\text{conv}} = \frac{\kappa}{L}\left(\text{Nu} - 1\right) \tag{12}$$

where κ is the thermal diffusivity and L is the depth of the convective cell.

3.2.6 Topographic relief of convecting terrain

We estimate the difference in elevation h between the upwelling centres of the polygons and their sinking margins by equating the buoyancy stress in the convecting fluid to the stress generated by topography, $\rho g h$, where ρ is the density and g is Pluto's surface acceleration of gravity (0.62 m s^{-2}). The buoyancy stress is equal to the density deficit of the warm, rising fluid, $\rho\alpha\Delta T$, where α is the volume coefficient of expansion and ΔT is the temperature difference between the hot and cold boundaries of the convecting system. The density deficit is multiplied by the height of the convection cell, L, times g to define the convective stress, from which we deduce:

$$h = \alpha \Delta T L \tag{13}$$

However, ΔT is not known a priori. We can define it more precisely in terms of quantities better defined in convecting systems by exploiting the conventional definition of the Rayleigh number Ra to solve for ΔT and write:

$$h = \frac{\mathrm{Ra}}{\rho g}\frac{\kappa \eta_{\mathrm{eff}}}{L^2} \tag{14}$$

Inserting this expression into our system of parameterized convection equations for the nominal case of a heat flow of 3 mW m^{-2} and L = 10 km yields an estimate of about 80 m for the difference in elevation between the centre and edges of the Rayleigh–Bénard cells, as we report in the text.

3.2.7 The 'floating mountains' of Sputnik Planitia

Dark material congregates along the edges of the Sputnik Planitia polygons (see Figure 3.1). These hills (currently called 'floating mountains' in NASA press releases) rise hundreds of metres[1] above the surrounding terrain within Sputnik Planitia. Owing to the albedo contrast with the surrounding N_2 ice, the material seems likely to be composed of water ice[1]. Nearly all of the ice chunks are located at the edges of the cells rather than the centres, and are arranged in lines and arcs that do not resemble the rims of submerged impact craters. Because such a non-random distribution of mountains is unlikely, we conclude that the mountains are afloat and are moved by N_2 convection to the edges of the polygons. Although it is possible that the downwelling limbs of the convection cells have aligned with grounded water-ice mountains in thinner N_2 ice, the polygonal arrangement of these mountains at the same distance scale as the mountain-free polygons strongly suggest that their arrangement was determined by the dynamics of the convection cells rather than vice versa.

If these 'floating mountains' are icebergs, then we can calculate the minimum depth of N_2 ice beneath each one that is needed to generate strong enough buoyancy forces to keep it afloat. According to Archimedes' principle, the depth of the bottom of an iceberg of height h above the surface is:

$$d = \frac{h}{\left(\frac{\rho_{\mathrm{N}}}{\rho_{\mathrm{W}}} - 1\right)} \tag{15}$$

where ρ_{N} is the density of N_2 and ρ_{W} is the density of H_2O. At Pluto temperatures of ~37 K (ref. 1), the density of water ice is ~930 kg m^{-3} (refs 25, 26), and the density of N_2 is ~1,030 kg m^{-3}

(from equation (3)). Using these densities and a height of 500 m for the iceberg topography, we calculate a minimum depth of 5 km.

The horizontal extent of these mountains also gives clues to their depths because tall, narrow cylindrical masses of ice are not stable: they would tilt to achieve a minimum gravitational energy configuration. The largest observed masses are about 5 km across (see Figure 3.1), suggesting a minimum N_2 ice depth comparable to, or greater, than this distance.

3.2.8 Effect of r_η on the aspect ratio for Rayleigh–Bénard convection

The temperature-dependence parameter r_η is a non-dimensional ratio between the viscosities at the top and bottom of the convection cell that determines the regime of convection (transitional mode, stagnant-lid mode or small-viscosity-contrast mode) as well as the aspect ratio[14]. We show below that the N_2-ice layer in Sputnik Planitia is well within the small-viscosity-contrast regime, so special considerations for convection in strongly temperature-dependent fluids do not apply.

The ratio r_η is given by the following formula[14]:

$$r_\eta = \mathrm{e}^{E(T_\mathrm{b}-T_\mathrm{s})} \tag{16}$$

where T_b is the temperature at the bottom of the convection and T_s is the surface temperature. Within equation (16), the E term is a constant determined by fitting the temperature-dependent viscosity equation to rheologic measurements for a material[15] and is defined as:

$$E = \frac{Q}{RT^2} \tag{17}$$

where T is the mean temperature, R is the gas constant, and Q is the activation energy (see equation (9)). Using a temperature of 45 K and the activation energy for N_2 (see parameters used in equation (9)), we calculated an E constant of ~0.2, corresponding to an r_η value of ~15 for N_2. This value for r_η places the N_2 in Sputnik Planitia within the small-viscosity-contrast convection regime, where an aspect ratio greater than 3:1 is not possible for $\mathrm{Ra} > 10^6$ (refs 6, 14).

3.3 References

1. Stern, S. et al. The Pluto system: initial results from its exploration by New Horizons. *Science* **350**, http://dx.doi.org/10.1126/science.aad1815 (2015).

2. Harry, D. & Gozdzik, J. Ice wedges: growth, thaw transformation, and palaeoenvironmental significance. *J. Quat. Sci.* **3**, 39–55 (1988).

3. Kindle, E. Some factors affecting the development of mud-cracks. *J. Geol.* **25**, 135–144 (1917).

4. Lachenbruch, A. Mechanics of thermal contraction cracks and ice-wedge polygons in permafrost. *Geol. Soc. Am. Spec. Pap.* **70**, 1–66 (1962).

5. Stachowiak, P., Sumarokov, V., Mucha, J. & Jeżowski, A. Thermal conductivity of solid nitrogen. *Phys. Rev. B* **50**, 543–546 (1994).

6. Hansen, C. & Paige, D. Seasonal nitrogen cycles on Pluto. *Icarus* **120**, 247–265 (1996).

7. McGill, G. & Hills, L. Origin of giant Martian polygons. *J. Geophys. Res.* **97**, 2633–2647 (1992).

8. Pechmann, J. The origin of polygonal troughs on the Northern Plains of Mars. *Icarus* **42**, 185–210 (1980).

9. Freed, A. et al. On the origin of graben and ridges within and near volcanically buried craters and basins in Mercury's northern plains. *J. Geophys. Res.* **117**, E00L06 (2012).

10. Blair, D. et al. The origin of graben and ridges in Rachmaninoff, Raditladi, and Mozart basins, Mercury. *J. Geophys. Res. Planets* **118**, 47–58 (2013).

11. Schubert, G., Turcotte, D. & Olson, P. Mantle Convection in the Earth and Planets (Cambridge Univ. Press, 2001).

12. Stern, S., Porter, S. & Zangari, A. On the roles of escape erosion and the viscous relaxation of craters on Pluto. *Icarus* **250**, 287–293 (2015).

13. Barr, A. & Hammond, N. A common origin for ridge-and-trough terrain on icy satellites by sluggish lid convection. *Phys. Earth Planet. Inter.* **249**, 18–27 (2015).

14. Kameyama, M. & Ogawa, M. Transitions in thermal convection with strongly temperature-dependent viscosity in a wide box. *Earth Planet. Sci. Lett.* **180**, 355–367 (2000).

15. Yamashita, Y., Kato, M. & Arakawa, M. Experimental study on the rheological properties of polycrystalline solid nitrogen and methane: implications for tectonic processes on Triton. *Icarus* **207**, 972–977 (2010).

16. Moresi, L. & Solomatov, V. Numerical investigation of 2D convection with extremely large viscosity variations. *Phys. Fluids* **7**, 2154–2162 (1995).

17. Angwin, M. Nitrogen–carbon monoxide phase diagram. *J. Chem. Phys.* **44**, 417–418 (1966).

18. Zhao, W., Yuen, D. & Honda, S. Multiple phase transitions and the style of mantle convection. *Phys. Earth Planet. Inter.* **72**, 185–210 (1992).

19. Christensen, U. & Yuen, D. The interaction of a subducting lithospheric slab with a chemical or phase boundary. *J. Geophys. Res.* **89**, 4389–4402 (1984).

20. Schubert, G., Yuen, D. & Turcotte, D. Role of phase transitions in a dynamic mantle. *Geophys. J. Int.* **42**, 705–735 (1975).

21. Scott, T. Solid and liquid nitrogen. *Phys. Rep.* **27**, 89–157 (1976).

22. Niemela, J., Skrbek, L., Sreenivasan, K. & Donnelly, R. Turbulent convection at very high Rayleigh numbers. *Nature* **404**, 837–840 (2000).

23. Schubert, G. Numerical models of mantle convection. *Annu. Rev. Fluid Mech.* **24**, 359–394 (1992).

24. Karato, S. Deformation of Earth Materials: an Introduction to the Rheology of Solid Earth. 338-362 (Cambridge Univ. Press, 2012).

25. Eisenberg, D. S. & Kauzmann, W. The Structure and Properties of Water. p. 296 Clarendon Press, 1969).

26. Hobbs, P. Ice Physics. p. 346 (Oxford Univ. Press, 2010).